Molecular epidemiological study of multidrug-resistant *Mycobacterium tuberculosis* in Lusaka, Zambia

(ザンビア、ルサカ市における多剤耐性結核の分子疫学的研究)

Chizimu Yamweka Joseph

Division of Bioresources
International Institute for Zoonosis Control
Hokkaido University

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

DR	Direct repeat
EMB	Ethambutol
gyrA	Gene encoding DNA gyrase A subunit
gyrB	Gene encoding DNA gyrase B subunit
HIV	Human Immunodeficiency Virus
INH	Isoniazid
InhA	Enoyl-acyl carrier protein reductase
KatG	Catalase-peroxidase
KatG	Gene encoding catalase-peroxidase
Leu	Leucine
MDR	Multidrug-resistant
MGIT	Mycobacteria growth indicator tube
MIRU-VNTR	Mycobacterial interspersed repetitive units-variable number tandem repeats
PCR	Polymerase chain reaction
PE	Proline-glutamate
PPE	Proline-proline-glutamate
QUB	Queen's University Belfast
RIF	Rifampicin
RFLP	Restriction fragment length polymorphism
RNA	Ribose nucleic acid
RpoB	RNA polymerase β-subunit
RR	Rifampicin resistant
RRDR	Rifampicin resistance determining region
SIT	Spoligotype international type
SNP	Single nucleotide polymorphism
Spoligotyping	Spacer-oligonucleotide typing
STR	Streptomycin
TB	Tuberculosis
WHO	World Health Organisation
XDR	Extensively drug-resistant
ZN	Ziehl-Neelsen

NOTES

1. Chizimu J.Y, Solo E.S, Bwalya P, Kapalamula T.F, Akapelwa M.L, Lungu P, Shrestha D, Fukushima Y, Mukonka V, Thapa J, Nakajima C, Suzuki Y. Genetic Diversity and Transmission of Multidrug Resistant Mycobacterium tuberculosis strains in Lusaka, Zambia. Int. J. Infect. Dis. 2021

2. Chizimu J.Y, Solo E.S, Bwalya P, Tanomsridachchai W, Chambaro H, Shawa M, Kapalamula T.F, Lungu P, Fukushima Y, Mukonka V, Thapa J, Nakajima C, Suzuki Y. Whole-genome sequencing reveals recent transmission of multidrug-resistant *Mycobacterium tuberculosis* CAS1-Kili strains in Lusaka, Zambia Antibiot. 2021

PREFACE

Tuberculosis continues to be a major public health issue around the world. Over the previous years, progress has been recorded towards TB elimination, though the gains have been lost due to the negative impacts of the COVID-19 outbreak on tuberculosis. Thus, this has handicapped the WHO End TB Strategy by 2035. In 2020, 1.3 million HIV-negative people died from this curable and preventable disease, which was an 8% increase when compared to 2019 (WHO, 2021a). Therefore, the disease has remained a major cause of ill-health and mortality in many countries. The disease is worse in developing countries, especially those with a high prevalence of HIV/AIDS. In 2020, the WHO regions of South-East Asia, Africa, and the Western Pacific contributed 43%, 25%, and 18% of global TB cases, respectively. While the Eastern Mediterranean, the Americas, and Europe had lower percentages of 8.3%, 3%, and 2.3% of TB cases respectively. Besides, 86% of all estimated incident cases globally came from the 30 nations with the highest TB burden (Figure 1). Moreover, two-thirds of the global total were from eight countries that included India (26%), China (8.5%), Indonesia (8.4%), Philippines (6.0%), Pakistan (5.8%), Nigeria (4.6%), Bangladesh (3.6%) and South Africa (3.3%). Figure 2 shows the estimated TB incidence rates of 2020 by WHO.

In addition to the current negative impact of COVID-19 on tuberculosis management, drug-resistant tuberculosis has continued to be a threat to TB control globally. It's a major cause of death due to anti-microbial resistance. This is because its treatment involves the use of second-line anti-tuberculosis drugs that have more adverse events and expensive. Drug resistance develops as a result of the incorrect administration of antibiotics in the treatment of drug-susceptible tuberculosis. This improper usage is the consequence of a variety of behaviors, including the prescription of ineffective treatment regimens, poor compliance, and the failure to ensure that patients finish the entire course of treatment. The development of drug resistance in most cases is indicative of ineffective tuberculosis control programs (WHO, 2021b). WHO has defined cases of drug-resistant TB into five groups. These are isoniazid-resistant, rifampicin-resistant TB (RR-TB), multidrug-resistant TB (MDR-TB), pre-extensively drug-resistant TB (pre-XDR-TB), and extensively drug-resistant TB (XDR-TB). MDR-TB is when *M. tuberculosis* becomes resistant to both isoniazid and rifampicin. While when resistant to rifampicin and any fluoroquinolone (a class of second-line anti-TB drug) it's defined as pre-XDR-TB. In an event where *M. tuberculosis* becomes resistant to rifampicin, any fluoroquinolone, and at least one of the medicines bedaquiline and linezolid it is known as XDR-TB (WHO, 2021a), a more severe disease manifestation with complicated treatment.

In 2020, 157903 drug-resistant cases were recorded, including 132222 multidrug or rifampicin-resistant tuberculosis (MDR/RR-TB) cases and 25681 pre-XDR-TB or XDR-TB cases. This was a considerable decline (by 22%) from the total of 201997 drug-resistant tuberculosis cases diagnosed in 2019. Between 2019 and 2020, there were considerable decreases in the overall number of people newly diagnosed with TB (by 18%) and the total number of people diagnosed with bacteriologically proven pulmonary TB (by 17%). The reductions were attributed to the COVID-19 epidemic, which disturbed the systems for TB control. MDR/RR-TB continues to be an agonizing situation regarding its treatment success. Though an increment of 19% from 2012 was reported in 2018, the global treatment success of 59% was still low. This ranged from 56% in WHO European Regions to 69% in WHO African Regions.

Mycobacterium tuberculosis, the cause of tuberculosis, belongs to the order Actinomycetales. It was identified by Robert Koch in 1882. *Mycobacterium tuberculosis* is a non-motile, non-sporulating, rod-shaped, acid-fast bacillus with a slow growth rate. It is 1 to 4 μm in length, and 0.3 to 0.6 μm width. It is a member of the *Mycobacterium tuberculosis* complex (MTBC) grouped under *M. tuberculosis sensu stricto*. Other members of MTBC include *M. africanum, M. canetti, M. bovis, M. caprae, M. microti, M. pinnipedii, M. mungi,* and *M. orygis* (Figure 3a). MTBC members are closely related with 99.9% similarity in DNA sequence, but they are discriminated against each other by the existence of rare synonymous single-nucleotide polymorphisms (SNPs) (Smith, 2003). Several MTBC species have inherent host preferences, with some infecting particular animal species whilst others infect humans. For example, *M. bovis* has cattle, *M. caprae* (sheep and goats), and *M. pinnipedii* (seals and sea lions) as natural hosts. The primary causes of human TB are *M. tuberculosis sensu stricto* and *M. africanum*. These pathogenic mycobacteria are thought to have originated from early predecessors that thrived in a variety of environments and later gained the ability to persist intracellularly. MTBC, together with *M. canetti* and other "smooth strains," form a single and compact clonal group based on 16S rRNA sequencing. *M. canetti* and the other smooth strains have different features than the other members of the MTBC, such as producing smooth colonies with a shorter generation period. *M. prototuberculosis* is the term referring to these smooth strains, which are thought to represent the mycobacteria population that gave rise to MTBC (Gutierrez et al., 2005).

The hydrophobic cell wall is characteristic to *M. tuberculosis.* The cell wall has four main layers (Figure 4). Starting from inside to the outside of the cell, the layers are as follows; (a) inner plasma membrane, (b) the peptidoglycan–arabinogalactan complex (AGP), (c) an

asymmetrical outer membrane connected to AGP by mycolic acids, and (d) the external capsule (Kalscheuer et al., 2019).

The peptidoglycan–arabinogalactan complex (AGP) has mycolic acids which are unique to mycobacteria and have a role in the cell wall structure and function. While arabinogalactan forms a hydrophobic permeability barrier. The external capsule of the cell wall is composed of proteins, lipids, and polysaccharides. The cell wall of *M. tuberculosis* contributes to its specific characteristics regarding acid/alcohol fast nature, slow cell division, distinctive immunostimulatory capabilities, resistance to antimicrobial agents, detergents, and other environmental stressors (Daffé and Draper, 1997).

People get infected with *M. tuberculosis* through the inhalation of infected droplet nuclei from an infected individual. A single aerosol droplet can contain 1 to 400 bacilli which is more than the infectious dose of 1 to 200 bacilli (Balasubramanian et al., 1994). Once in the lungs, the bacilli are phagocytosed by alveolar macrophages. However, bacilli continue to multiply in macrophages and trigger a chemotactic response, which attracts more macrophages and other defense cells to the site. The inflammatory cells encompass the infection site and form a tubercle with infected macrophages in the core surrounded by epithelioid macrophages, foam cells, and multinucleated giant cells of the Langhans type, lymphocytes, and a fibrous capsule. The tubercle balances between infection containment in the host and *M. tuberculosis* protection from IFN-producing cells. The majority of the macrophages fail to eliminate bacteria and instead release enzymes that induce more inflammation that is harmful to the host. After some time, the tubercle undergoes caseous necrosis, and hence the dead macrophages release the bacilli, high lipid, and protein content. At this stage of the disease, symptoms may start to appear. However, the tubercle is hypoxic and aerobic bacilli do not grow well and hence remain dormant, a stage called latent tuberculosis. The bacilli survive in this state by metabolizing fatty acids, reducing active replication, and thickening cell walls (Boshoff and Barry, 2005). In individuals with competent immune systems, lesions become calcified and the infection is stopped at this stage or remains dormant for many years.

While in immune-compromised people the caseous region continues to grow by liquefaction resulting in an air-filled tuberculous cavity that supports bacilli growth. Liquefaction erodes the walls causing rupture of the tubercle releasing the bacilli into the bronchioles and bloodstream leading to disseminated tuberculosis called milliary tuberculosis. At this stage, the patient presents with symptoms. The disease usually affects the lungs but can involve other organs as in the case of extra-pulmonary tuberculosis. Symptoms include productive cough with blood-stained sputum due to damaged tissues, fever, night sweats,

fatigue, lack of appetite, and weight loss. On examination of the chest, they may have abnormal breath sounds and crepitations. It's worth noting that the bacilli can spread to the regional lymph nodes and hematogenous spread to other organs before the formation of these organized granulomas.

Usually, 5–10% of people with latent tuberculosis develop an active infection at some point in their lives (CDC, 2016) when the immunity gets suppressed for some reason. Therefore, the risk factors for having an active infection include comorbidities such as HIV/AIDS, diabetes, chronic renal disease, neoplastic disorders, undernutrition, and those receiving chemotherapy. Others are smoking and alcohol abuse disorders. Hence, patients presenting with such risk factors and symptoms related to tuberculosis should be screened for TB using any of the available diagnostic methods.

Tuberculosis has been diagnosed using a variety of techniques. Smear microscopy is widely used, especially in developing nations, for the detection of acid-fast bacilli despite having low sensitivity and specificity when compared to culture. In this method, the smear is fixed, stained with carbol-fuchsin, and decolorized with alcohol. Methylene blue is then used to counterstain the smear, giving the bacilli a pink appearance. These characteristics are due to the lipid-rich cell wall that does not readily take up the stain and resists decolorization when distained with an acid-alcohol wash. TB culture is considered the gold standard, though the technique takes weeks to months to get results. Therefore, rapid tests based on molecular markers such as Xpert MTB/RIF and GenoType MTBDRsl assay have been developed for rapid diagnosis (Kurbatova et al., 2013; Lacoma et al., 2012). Chest X-ray is another tool used in the diagnosis of tuberculosis. Nonetheless, even in patients with active infection, the chest X-ray result is sometimes negative. All patients diagnosed with TB are ideally started on the appropriate treatment with anti-tuberculosis drugs.

Anti-tuberculosis drug discovery, development, and treatment have been available since the 1940s, though the regimens have changed with time. The current treatment for susceptible tuberculosis is by the use of the first-line drugs, which are rifampicin, isoniazid, ethambutol, and pyrazinamide under the directly observed treatment short-course. The treatment is for 2 and 4 months for the intensive and continuation phases, respectively. Multiple drug therapies are applied to avoid the emergence of drug-resistant strains. In the event that *M. tuberculosis* becomes resistant to any of the first-line drugs, this is called drug-resistant tuberculosis.

The treatment of drug-resistant tuberculosis such as MDR/RR-TB includes the use of a shorter all-oral bedaquiline-containing regimen for 9–12 months duration. The regimen is

composed of 6 months of bedaquiline, together with 4 months of levofloxacin/moxifloxacin, pyrazinamide, ethambutol, isoniazid (high-dose), clofazimine, and ethionamide, and then followed by 5 months with levofloxacin/moxifloxacin, clofazimine, ethambutol and pyrazinamide. Patients must be fluoroquinolone susceptible and have not been exposed to the second-line TB medicines utilized in this regimen for more than one month. While in other patients longer treatments regimens are utilized. These include all three Group A antibiotics (levofloxacin or moxifloxacin, bedaquiline and linezolid) and at least one Group B antibiotic (clofazimine, and cycloserine or terizidone). Both Group B agents are included if only one or two Group A agents are used. Group C agents (ethambutol, delamanid, pyrazinamide, imipenem–cilastatin or meropenem, amikacin (or streptomycin), ethionamide or prothionamide, and p-aminosalicylic acid) are utilized in cases where Groups A and B drugs alone are unable to complete the treatment regimen (WHO, 2020).

In 1998 Cole ST et al. published the complete genome of *M. tuberculosis* H37Rv with approximately 4000 genes and 4.4 million base pairs (Cole ST et al.1998). This improved the understanding of drug resistance mechanisms to the most routinely prescribed anti-TB medications. It also revealed that specific gene changes were linked to drug resistance (Ramaswamy and Musser, 1998). When compared to other bacteria, *M. tuberculosis* acquires drug resistance mostly through spontaneous mutations in chromosomal genes, as opposed to horizontal transmission of mobile genetic elements such as transposons, plasmids, and integrons. Sub-optimal medication therapy facilitates the acquisition of drug resistance, resulting in a selection of resistant bacteria (Kochi et al., 1993). According to studies, most anti-tuberculosis medications have a mutation rate of 10^{-9} mutations per cell division. This rate, however, may vary based on the type of drug chosen. As a result, medications are administered in combination to lessen the chance of developing drug resistance. A mutant with two resistance mutations, for example, has a risk of 10^{-18} (Gillespie, 2007).

Resistance to isoniazid involves mutations in *katG* and *inhA* genes. Several mutations in those genes have been identified, though Ser 315 Thr amino acid change in *katG* has been the most prevalent, appearing in 50-93 percent of resistant clinical isolates. While mutations in an 81 bp drug resistance determining region (RRDR) of *rpoB* lead to varying levels of rifampicin resistance, with mutations like His 526 Asp, Asp 516 Val, and Ser 531 Leu being the most prevalent and seen worldwide. Other mutations in *rpoC* and outside the RRDR act as compensatory mutations.

Despite the discrepancy in reports concerning the mechanisms for ethambutol resistance, mutations in *embB*, especially at condon 306, have been found to be responsible for

50% of ethambutol resistant cases (Sreevatsan et al., 1997; Telenti et al., 1997). Some *embB* mutations have been linked to high levels of resistance, while others have been linked to low levels of resistance. Whereas, studies have reported up to 20% of susceptible strains with mutations in *embB* (Lee et al., 2004). The last but not the least first-line drug is pyrazinamide. The resistance mechanism to this drug involves mutations in *pncA*, mainly in the 561 bp region of the open reading frame or an 82 bp region of its putative promoter (Scorpio et al., 1997).

Patients with resistant strains to the first-line anti-TB drugs are treated with second-line drugs, which include the use of fluoroquinolones, aminoglycosides etc. Fluoroquinolones target *gyrA* and *gyrB*, which encode the DNA gyrase involved in DNA replication. Mutations in the quinolone resistance determining region (QRDR) mainly at codons 90, 91, and 94 of *gyrA* are implicated in drug resistance (Maruri et al., 2012; Palomino and Martin, 2014). Additionally, *M. tuberculosis* strains become resistant to aminoglycosides (Kanamycin and Amikacin) by acquiring mutations at positions 1400, 1401 and 1483 of the *rrs* gene (Palomino and Martin, 2014; Suzuki et al., 1998).

However, some of the drug-resistance mutations in genes such as KatG (Ser 315 Thr) and rpoB (Ser 450 Leu) of *M. tuberculosis* have an impact on its fitness. Reports have hypothesized a reduced ability to spread and cause disease among such strains. Therefore, *M. tuberculosis* acquires compensatory mutations to mitigate the negative effects of drug resistance mutations and hence preserve the ability of the bacteria to spread (Brandis et al., 2012; Comas et al., 2011). It is cardinal to know the drug resistance patterns of the circulating strains in a particular area to guide the appropriate treatment regimens and control measures. Hence, the combination of drug susceptibility testing and molecular epidemiology plays an important role in drug resistance surveillance.

Molecular epidemiological studies facilitate the understanding of transmission dynamics, characteristics of the circulating genotypes, and effectiveness of the TB control measures (Besser, 2013). Besides, molecular epidemiology can facilitate the comparison of strains from various geographical regions and provide evolutionary insight, in addition to distinguishing reinfection from reactivation and detecting laboratory cross-contamination. This is achieved through the application of various molecular genotyping methods that utilize the polymorphic sites at the nucleotide or gene level in the tuberculosis genome. Polymorphic sites include SNPs, clustered regularly interspaced short palindromic repeats (CRISPR), variable number of tandem repeats (VNTR), and insertion sequences (IS).

Spoligotyping is one of the genotyping procedures. This is aimed at polymorphisms in the direct repeat (DR) region of *M. tuberculosis*. DR region is composed of multiple copies of

a conserved 36-base-pair sequence (DR) which are separated by 35-41 unique spacer sequences (Figure 5a). It involves PCR amplification of the DR region using specific primers and hybridizing on a membrane. The presence and absence of the 43 unique spacer DNA sequences provide the basis for distinguishing strains (Kamerbeek et al., 1997). It has less discriminatory power when compared to mycobacterial interspersed repetitive units-variable number tandem repeat (MIRU-VNTR).

MIRU-VNTR typing utilizes the allelic variations in the VNTR loci in *M. tuberculosis* genome. The genome has 41 MIRU loci consisting of 40-100 bp repetitive sequences (Supply et al. 2006) (Figure 5b). The loci that a polymorphic are considered as VNTR loci. It's also a PCR-based assay in which MIRU-VNTR loci are amplified, visualized by gel electrophoresis, and molecular weights determined. The molecular weights are interpreted as the number of tandem repeats. The discriminatory power increases depending on the number of loci used, with the standard 24 MIRU-VNTR loci being the most discriminatory when compared to 15 and 12 MIRU-VNTR loci. When combined with spoligotyping the discriminatory power increases to almost that of IS6110 restricted fragment length polymorphism (RFLP).

IS6110 RFLP is the gold standard for studying MTBC strain relatedness by exploiting variations in restriction sites in the DNA of different strains. This variability is enough to yield the RFLP utilized in fingerprinting. A restriction enzyme (PvuII), digests the IS6110 element resulting in DNA fragments that are separated by gel electrophoresis and visualized by probing with an IS6110 probe. The IS6110 probe hybridizes to the right of the PvuII cleavage site. The blot containing HaeII-digested φX174DNA is then compared to the *M. tuberculosis* reference strain to generate unique fingerprints (Moström et al., 2002). IS6110 has 0–25 copies, and their positions on the *M. tuberculosis* chromosome differ from isolate to isolate (Soolingen et al. 1991). It has a higher discriminatory power when compared to MIRU-VNTR and spoligotyping, but lower than whole-genome sequencing (WGS). The advantages and disadvantages of each particular genotyping method are listed in the Table 1.

WGS is an ultimate high-resolution technique for determining strain relatedness, transmission drug resistance patterns, genomic evolution, and virulence factors. Other methods target polymorphic sites in only less than 1% of the TB genome. Hence, they have limited discriminatory power in differentiating closely related strains and are prone to homoplasy. Briefly, the procedure involves library preparation in which there's fragmentation of DNA and ligation of adapters. These are then attached to the flow cell and amplified to increase the signal emission. The signal is generated by the incorporation of a nucleotide/oligonucleotide during sequencing. The output raw sequences are then compared to the reference for drug-resistance

profile determination, transmission, evolution analysis, and lineage identification (Oudghiri et al., 2018).

MTBC has seven major human-adapted lineages, which are *M. tuberculosis sensu strictu* [Lineages 1 (L1) to 4], Lineage 7 (L7), and *M. africanum* (Lineages 5 to 6). In humans, these lineages are able to maintain full infection, disease, and transmission cycles (Ernst, 2012). Lineages 2 to 4 are the modern, while L1, L5 to L7 are termed the ancient lineages based on the genomic region TbD1 deletion. The TbD1 region is deleted in modern lineages but intact in ancient lineages (Gagneux, 2017). These lineages have been shown to have specific geographical distributions, with L2 and L4 being global, L1 around the Indian Ocean, and L3 in East Africa, central and South-Asia (Figure 3) whereas L5 (*M. africanum* West Africa 1) and L6 (*M. africanum* West Africa 2) are restricted to certain geographical areas, mainly West Africa. Lineages 1, 2, 3, and 4 are also known as Indo-Oceanic, East-Asian, Central Asian Strain (CAS), and Euro-American lineages, respectively. Other lineages, such as L7, have been restricted to the horn of Africa. The geographical preferences of MTBC lineages are a result of the higher fitness of the pathogen in a particular host population it has co-evolved with (Gagneux, 2012). However, some strains are now present in areas where they were originally in existence due to the migration of people and the evolution of tuberculosis over time. Furthermore, the lineages differ from each other as a result of varied compositions of single nucleotide polymorphisms (SNPs), insertion and deletion (indels), mobile and repetitive elements, large genomic deletions, and large duplications. The most diverse lineage is L1, with an average of 730 SNPs between any two L1 strains. While L7 is the least diverse, with an average of 230 SNPs between any two L7 strains.

The information on the circulating lineages in a specific region is important as different MTBC strains have different immunogenicity, virulence, and transmissibility (Coscolla and Gagneux 2014). It is also a factor that determines host response and the emergence of drug resistance, control, and clinical outcome of pulmonary and extra-pulmonary tuberculosis (Napier et al., 2020).

In Zambia, tuberculosis is among the major causes of morbidity and top ten causes of mortality (Ministry of Health, 2017). Zambia recorded a total TB incidence of 319/100,000 population in 2020 (WHO, 2021c). The trends for TB incidence, new and relapse TB cases notified, and HIV-positive/TB incidence have been relatively declining over the past decade, mainly because of a relatively well-established TB control program and the involvement of other stakeholders (Figure 6). Additionally, the treatment success for new and relapsed cases

registered in 2019 was 89%. However, Zambia is among the global high-burden countries for TB, HIV-associated TB, and MDR/RR-TB according to WHO (WHO, 2021a) (Figure 1).

Despite the global MDR/RR-TB burden remaining relatively stable, Zambia had a 2.8 times increase in the estimated percentage of TB cases with MDR/RR-TB among the new pulmonary TB cases for the period 2015 to 2018 (WHO, 2016, 2019a) (Figure 7). This was suggestive of possible transmission of primary MDR/RR-TB. In addition, previous studies and routine surveillance in Zambia revealed an increase in MDR-TB (Kapata et al., 2013; Mulenga et al., 2010; Solo et al., 2021). Similar to the global picture, MDR-TB in Zambia has been associated with a lower treatment success rate and high mortality when compared to pan susceptible tuberculosis. Nevertheless, MDR case surveillance to understand the circulating strains and their transmission patterns has been inadequate, leading to limited information on the subject. Therefore, to address this gap, this dissertation focused on the molecular epidemiological analysis of multi-drug resistant *Mycobacterium tuberculosis* in Lusaka, Zambia.

There are two chapters in this dissertation. I determined the genetic diversity and transmission patterns of multidrug-resistant *M. tuberculosis* in Lusaka, Zambia, in the first chapter. This was in response to an increase in routine surveillance reports of MDR-TB, but insufficient information on its transmission and diversity. In this chapter, it was revealed that to a larger degree, the MDR-TB spread was due to recent transmission rather than the independent acquisition of MDR. Further, clonal expansion of spoligotype international types (SIT) 21 CAS1-Kili (SIT21/CAS1-Kili) and SIT20 Latin American-Mediterranean one (LAM1) strains were implicated for this spread. Therefore, TB control programs involving genotyping coupled with conventional epidemiological methods to guide measures for stopping the spread of MDR-TB were recommended.

In the second chapter, I evaluated recent transmission among the lineage 3 (CAS1-Kili) strains that were linked to and responsible for MDR *M. tuberculosis* spread in the first chapter. A high-resolution technology, WGS was used to distinguish these strains. The CAS1-Kili strains were closely related, implying recent transmission as they exhibited low SNP variabilities. Furthermore, the phylogenetic analysis showed that Zambian strains were related to those from Tanzania, Malawi, and the United Kingdom. As a result, the chapter recommends using WGS in addition to standard epidemiological methodologies to better understand MDR-TB transmission.

Table 1. Advantages and disadvantages of molecular genotyping methods

Advantages	**Disadvantages**
Spoligotyping	
▪ Rapid, repeatable, robust and cost-effective ▪ Simple to compare findings from different laboratories ▪ Viable culture is not needed ▪ Small amount of DNA needed ▪ No need for DNA purification ▪ It is able to differentiate between re-infection and reactivation in patients who relapse during or after chemotherapy	▪ Poor discriminatory power hence overestimates clustering when compared to MIRU-VNTR and IS6110 ▪ Is unable to distinguish between several common strain families, particularly Beijing
Mycobacterial Interspersed Repetitive Unit-Variable Number of Tandem Repeats (MIRU-VNTR)	
▪ Viable culture is not needed ▪ Small amount of DNA needed ▪ Rapid, reproducible and less expensive when compared to RFLP ▪ It has a higher discriminatory power than spoligotyping ▪ Easier to compare findings from different laboratories	▪ In outbreak settings, it has limited ability to differentiate between relapse and re-infection ▪ Limited ability to differentiate some lineages eg Beijing ▪ Sometimes it's difficult to obtain a full MIRU-VNTR pattern ▪ Unclear stability or mutation rate of the loci

▪ IS6110 restriction fragment length polymorphism typing (RFLP)	
▪ High discriminatory power in differentiating epidemiologically unrelated strains especially in strains with multiple IS6110 insertion sites ▪ Thoroughly tested technique ▪ The polymorphic sites have high stability due to slow insertion rate	▪ Time-consuming and labor-intensive ▪ Requires a viable culture and high DNA quantity and digestion with restriction enzymes ▪ High level of technical expertise is needed ▪ Difficulties in reproducibility and comparability ▪ Poor discriminative power for strains with less than five copy number of IS6110
▪ Whole genome sequencing (WGS)	▪
▪ Ultimate highest discriminatory power in delineating closely related strains ▪ Can infer the likely direction of transmission without prior recourse to epidemiological data ▪ Provides high levels of evolutionary insight	▪ It's expensive and has a low turn-around time ▪ Programs for standardization and quality assurance have yet to be established. ▪ High level of technical expertise is needed ▪ Complicated bioinformatics and specialist software for storage and analysis ▪ standardization as well as quality assurance programs are yet to be established

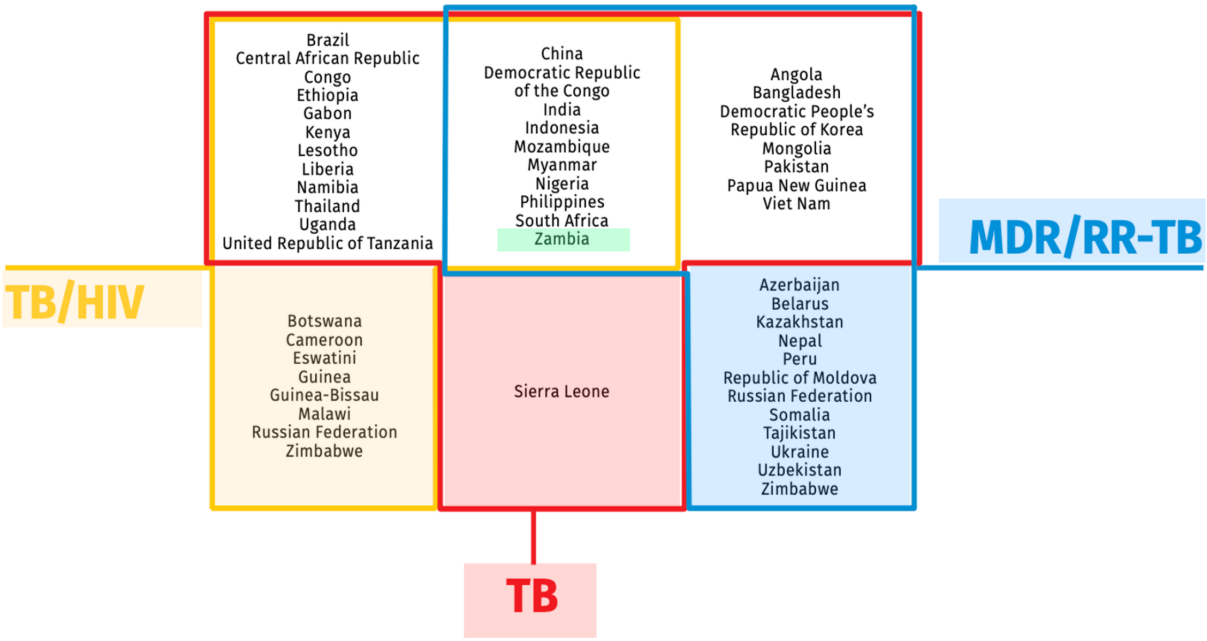


Figure 1. Figure showing Zambia among the three global lists for high-burden countries for TB, HIV-associated TB and MDR/RR-TB adapted from WHO 2021b. Each of the three lists is represented by a specific color; blue for MDR/RR-TB, red for TB and orange for TB/HIV. Some countries including Zambia, are in more than one list as shown by the overlapping areas.

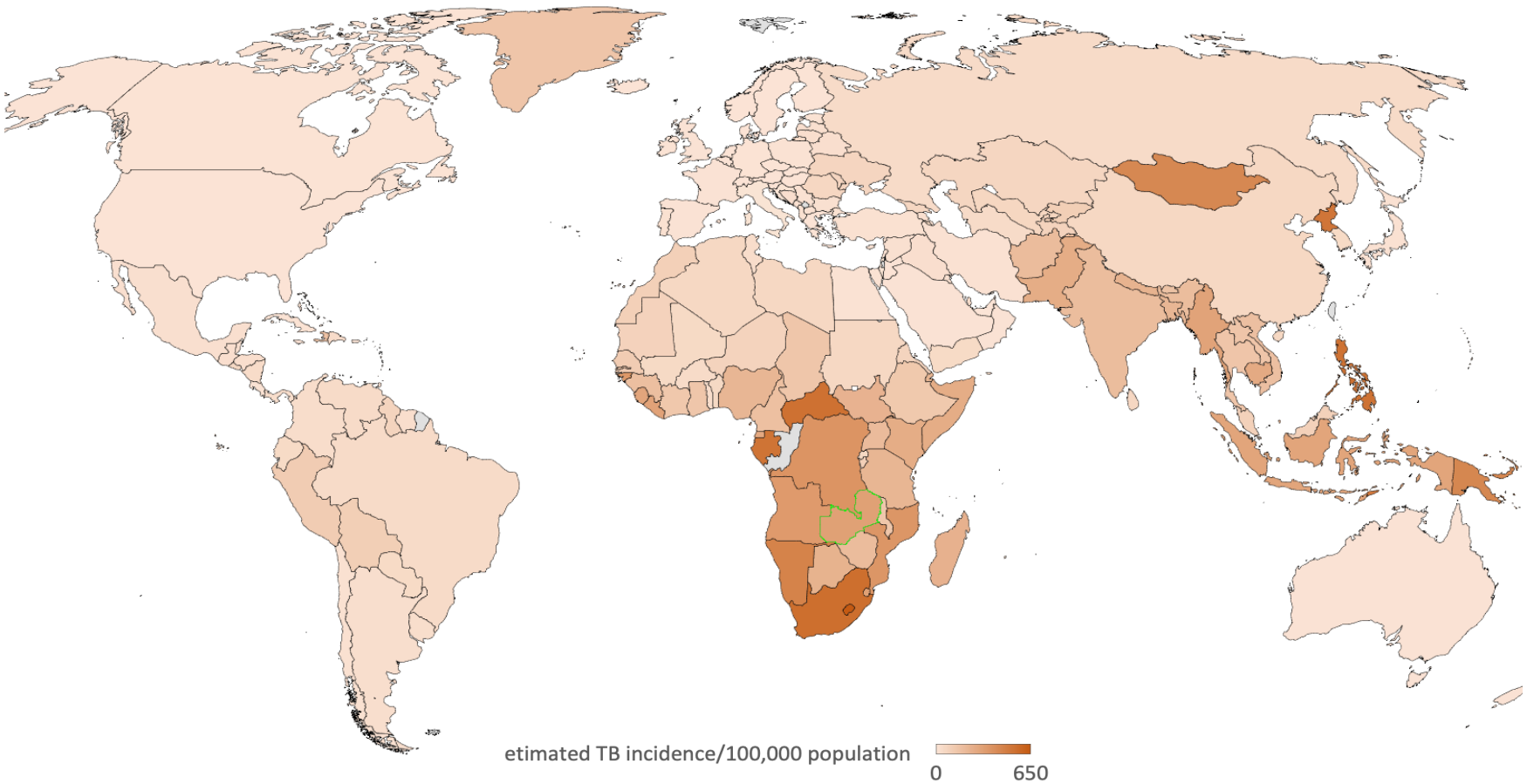


Figure 2. WHO TB incidence rate estimates for the year 2020 adapted from WHO 2021b. The map was shaded according to the incidence per 100, 000 population per year as indicated in the legend.

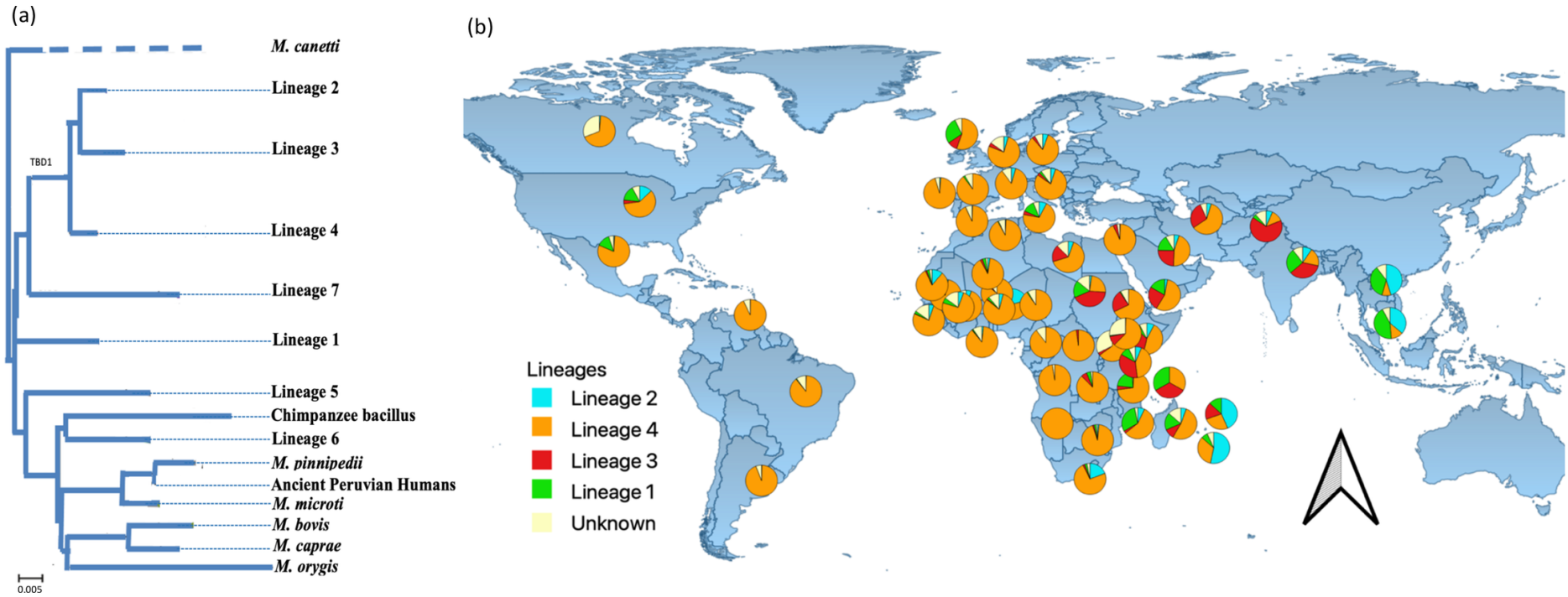
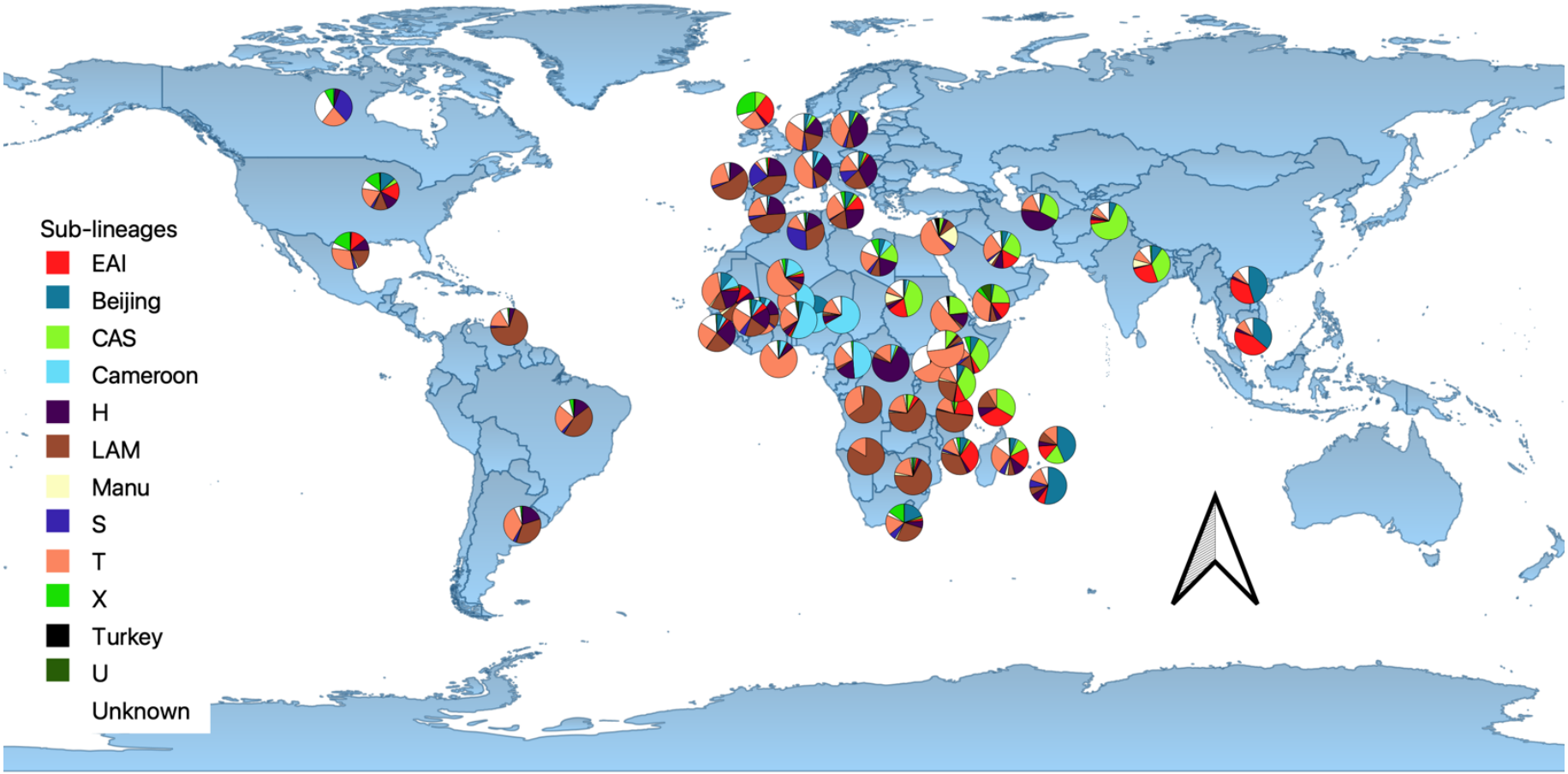


Figure 3. Phylogenetic tree showing the seven human and animal adapted lineages rooted to *M. canetti* adapted from Coscolla 2017. (b) Global map showing the distribution of four of the seven lineages (b) and (c) sub-lineages. (Data adapted from Chihota et al. 2018 and SITVIT2 WEB Database available on http://www.pasteur-guadeloupe.fr:8081/SITVIT2/index.jsp). Some lineages are widely distributed while others are restricted to certain geographic regions.

(c)


Sub-lineages
EAI
Beijing
CAS
Cameroon
H
LAM
Manu
S
T
X
Turkey
U
Unknown

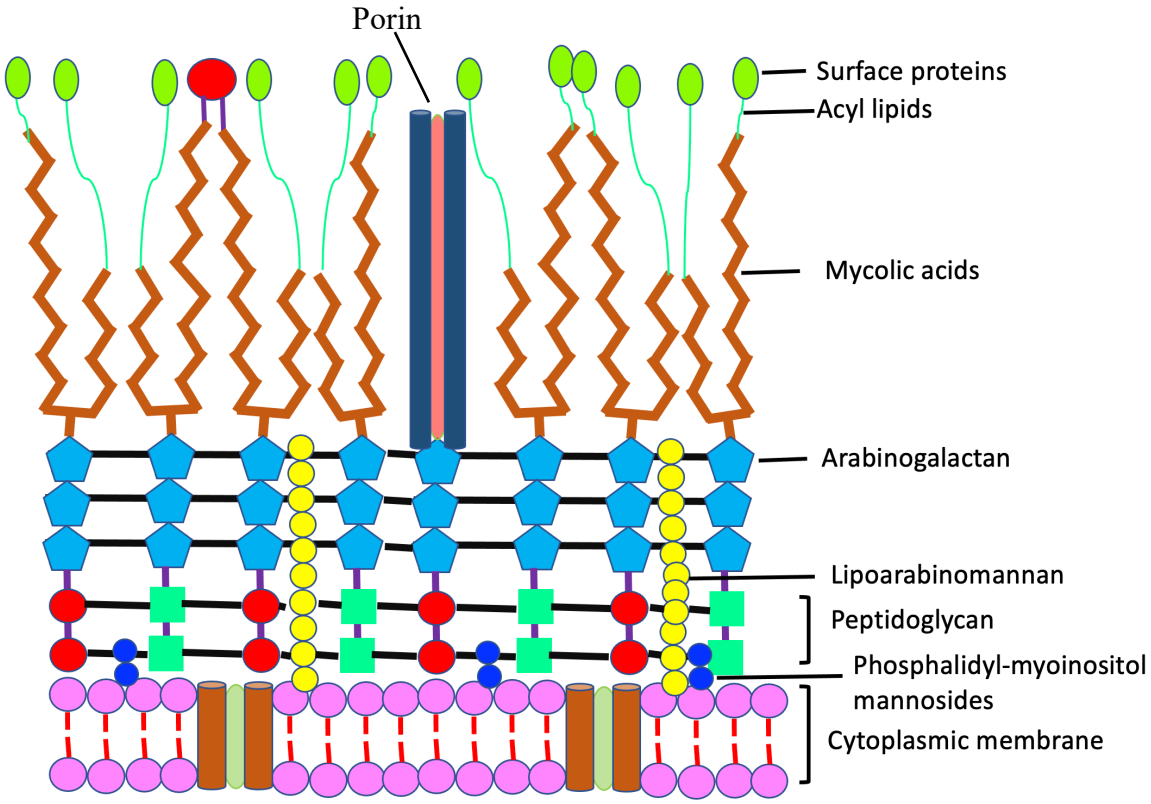


Figure 4. Mycobacterium cell wall structure adapted from Kaiser 2019.

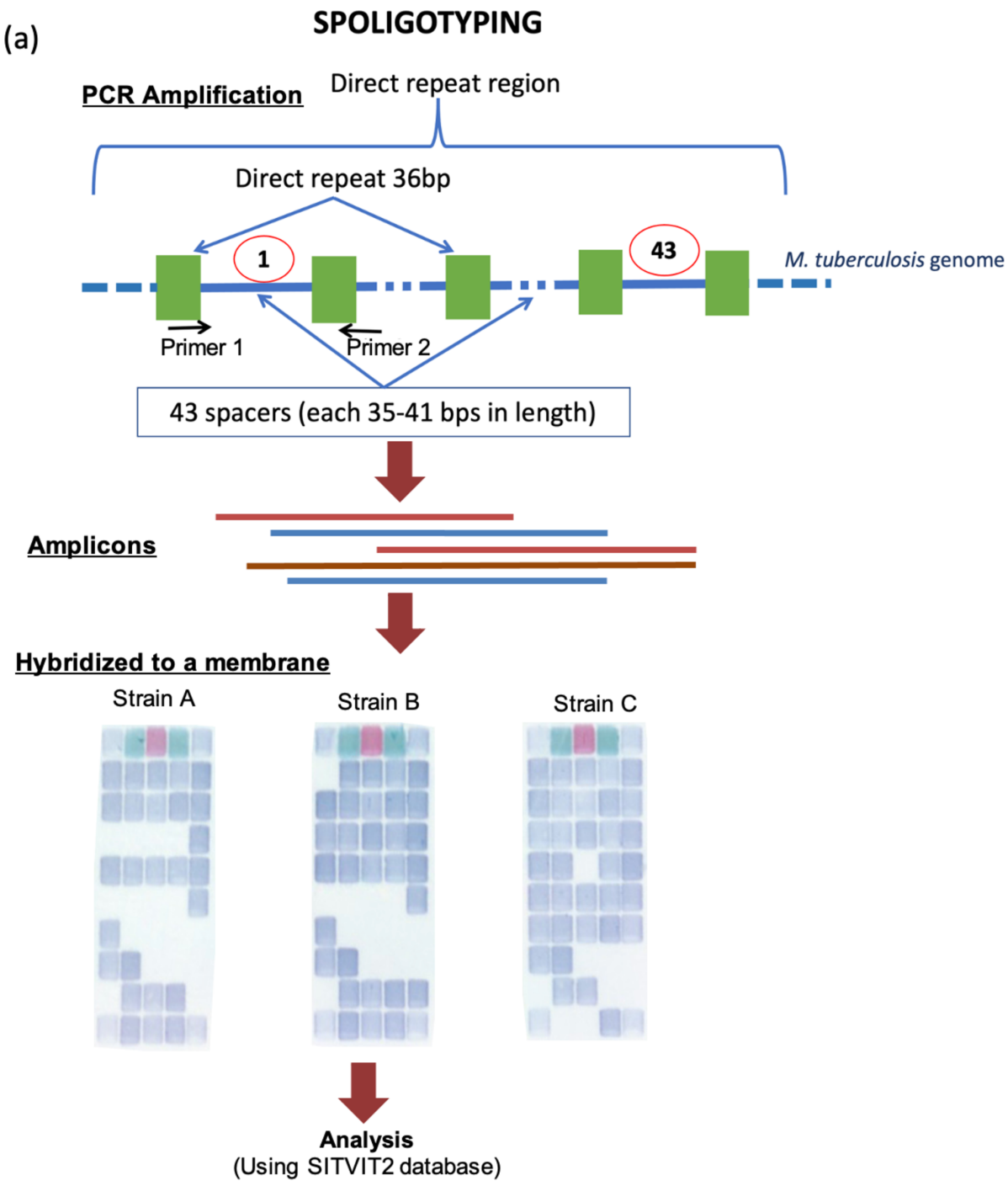
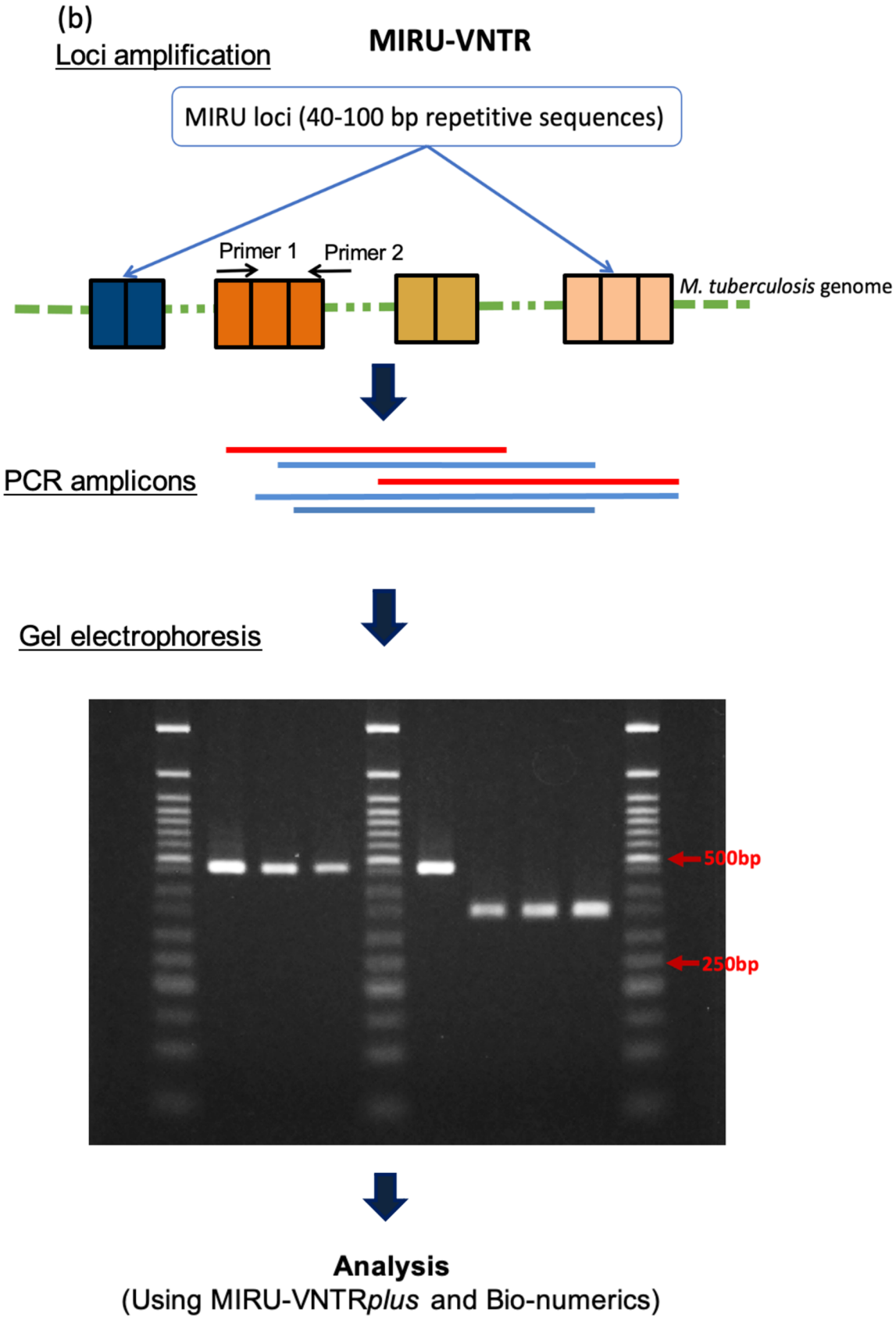
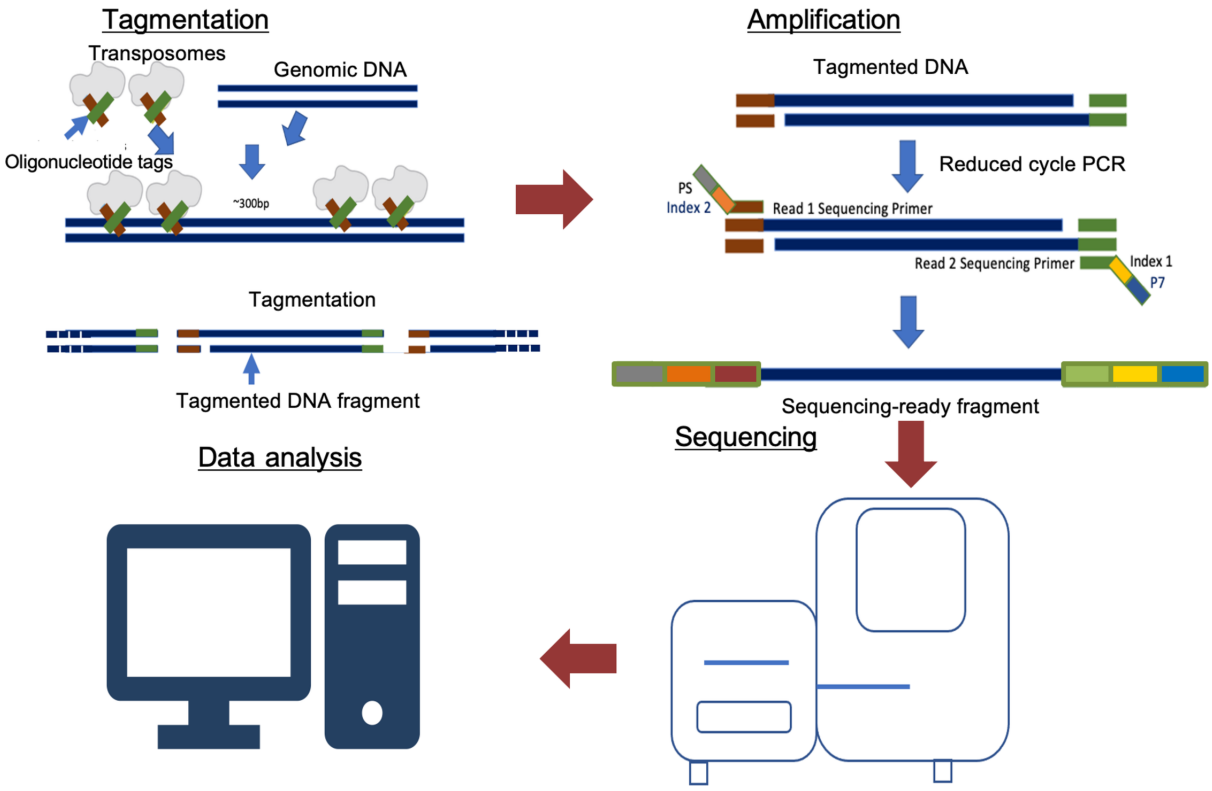


Figure 5. Diagram illustration the genotyping methods utilized in this dissertation. (a). Spoligotyping (Kamerbeek et al., 1997) (b). MIRU-VNTR (Supply et al., 2006) (c). Whole genome sequencing adapted from Disease Profiling Project 2018.


(b)
Loci amplification
MIRU-VNTR
MIRU loci (40-100 bp repetitive sequences)
Primer 1
Primer 2
M. tuberculosis genome
PCR amplicons
Gel electrophoresis
500bp
250bp
Analysis
(Using MIRU-VNTRplus and Bio-numerics)

(c) **Whole genome sequencing**



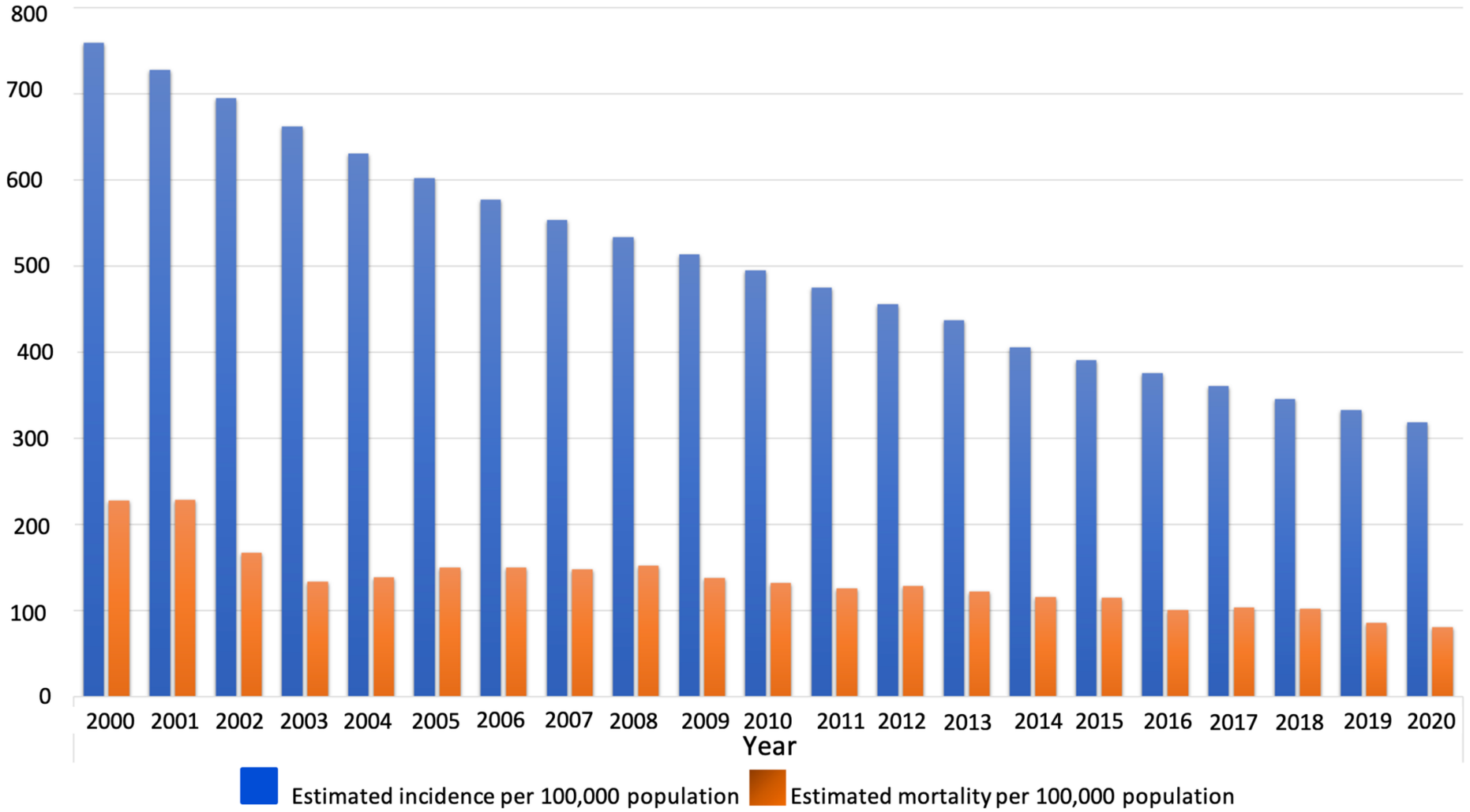


Figure 6. Zambia tuberculosis incidence and mortality estimates per 100,000 population for the period 2000 to 2020 (WHO, 2021a).

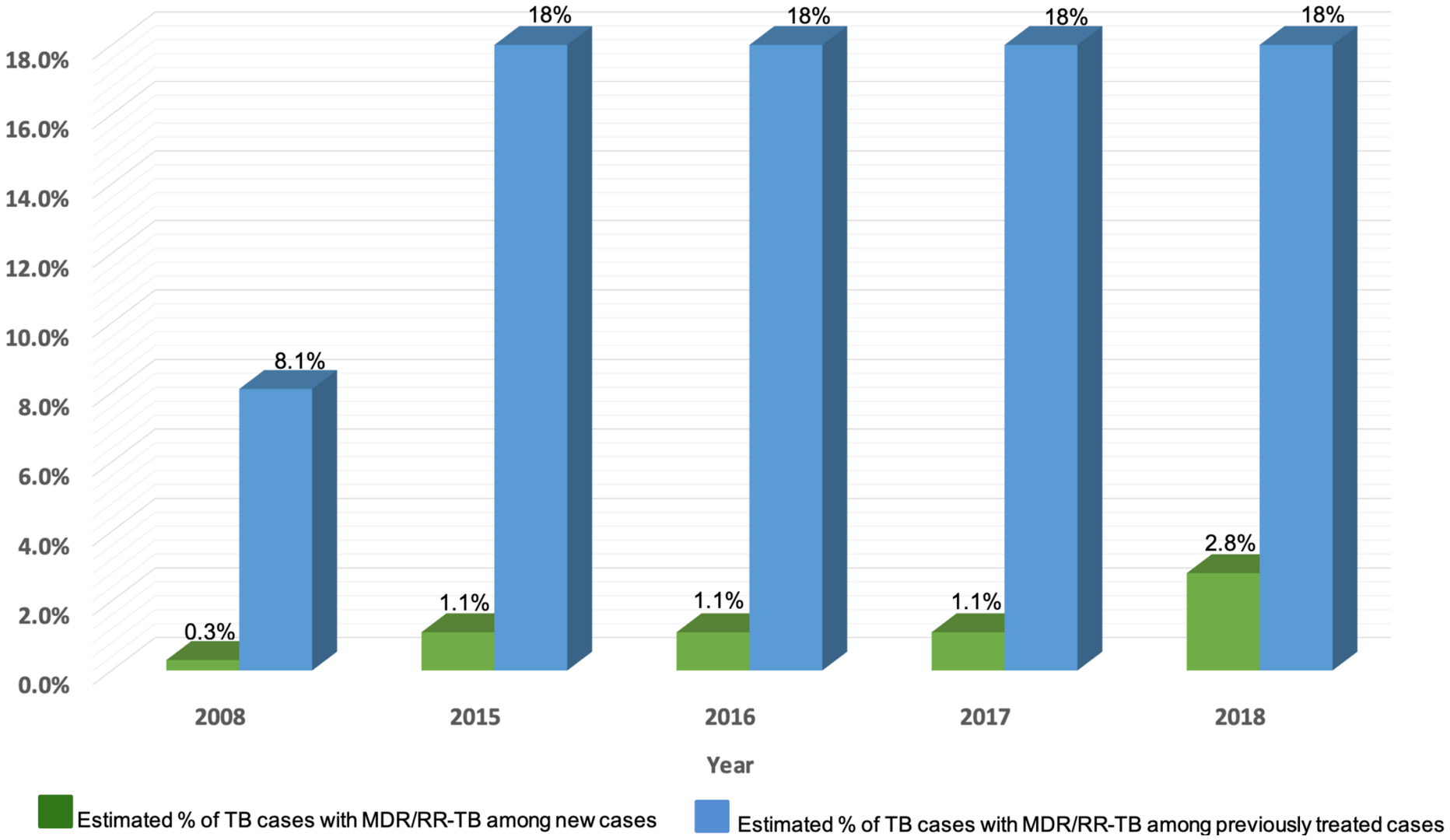


Figure 7. Estimated TB cases with MDR/RR-TB among the new and previously treated cases (WHO, 2019a, 2018, 2017, 2016).

CHAPTER 1

Genetic Diversity and Transmission of Multidrug Resistant Mycobacterium tuberculosis strains in Lusaka, Zambia

Introduction

Tuberculosis exacerbated by the emergence of MDR strains remains a major public health challenge in many countries worldwide. MDR-TB is caused by *M. tuberculosis* that is resistant to isoniazid and rifampicin, the most effective first-line drugs. Its treatment involves the use of second-line drugs which are expensive and have more side effects. MDR-TB is associated with a low treatment success rate and high mortality (WHO, 2019b). Moreover, MDR-TB has paused a threat to the achievement of End TB strategy by 2035 (2019 WHO, 2019). Tuberculosis with an incidence of 333/100,000 is among the top ten causes of mortality in Zambia (Ministry of Health, 2017; WHO, 2021d). Reports indicated the low occurrence of MDR-TB in the country (Ministry of Health, 2017; Mulenga et al., 2010; WHO, 2019b) though an incremental trend of notified cases per year was observed (Kapata et al., 2013). Drivers of MDR-TB in Zambia may be speculated to be late diagnosis and initiation of treatment due to fewer diagnostic centers, poor compliance among patients, and cross-border movements of people from high burden TB countries for economic activities and refuge.

Previous studies have indicated the predominance of Euro-American and East Asian lineages, the lineage 4 (L4) and 2 (L2), respectively, circulating in the southern African region (Chihota et al., 2018; Mbugi et al., 2016). Molecular genotyping methods such as spoligotyping and MIRU-VNTR are utilized in the understanding of genetic diversity and spread of *M. tuberculosis* (Said et al., 2012). Though differing in their ability to differentiate between unrelated strains, their combination provides a higher diversity index.

A few studies in Zambia have described *M. tuberculosis* diversity using 15 loci MIRU-VNTR in Ndola and Namwala districts (Malama et al., 2014; Mulenga et al., 2010). Further, an association of drug resistance with specific genotypes in Lusaka based on spoligotyping was previously described (Solo et al., 2021). Nevertheless, information on the transmission of MDR *M. tuberculosis* strains in Lusaka, Zambia is still inadequate. Hence, the current study explores this gap to guide the implementation of TB control interventions.

Materials and methods

Sample collection, processing, DNA extraction

The study utilized 85 MDR-TB clinical isolates collected and banked at the University Teaching Hospital (UTH) TB laboratory in Lusaka, Zambia, from 2013 to 2017. UTH TB laboratory is one of the three TB culturing facilities in the country covering over one-third of the country's population (Figure 8). Sample decontamination, sub-culturing, and drug susceptibility testing to the first-line drugs (rifampicin, isoniazid, ethambutol, and streptomycin) using BACTEC™ 960 MGIT™ (Mycobacteria Growth Indicator Tube) system (Becton Dickson & Co., Franklin Lakes, NJ, USA) were done following the manufacturer's instructions. The recommended critical concentrations by the kit manufacturer (Becton, Dickinson, and Company) of 0.1, 1.0, 1.0, and 5.0 μg/mL for isoniazid, rifampicin, streptomycin and ethambutol respectively were followed. DNA was extracted as described previously (Solo et al. 2020). The demographic information such as residence, age, and year of collection for the strains was captured after a review of laboratory registers.

Drug-resistance associated gene sequencing

Briefly, sequencing for *rpo*B, *kat*G genes, and *inh*A promoter region was conducted using the specific primers to detect mutations in respective genes as previously described (Solo et al. 2020) and resolute that isolates were genotypically multidrug-resistant.

Spoligotyping

Sub-lineages were identified using spoligotyping as described elsewhere (Kamerbeek et al. 1997; Solo et al. 2021). Briefly, PCR amplified the DR region using primers to generate different-sized amplicons which were hybridized to a set of 43 complimentary immobilized oligonucleotides on a membrane to detect the presence or absence of spacers. The resulting binary spoligotyping patterns were evaluated using the SITVIT2 WEB Database available on *http://www.pasteur-guadeloupe.fr:8081/SITVIT2/index.jsp*.

MIRU-VNTR typing

Strains were further analyzed using 24 loci (MIRU02, MIRU04, MIRU10, MIRU16, MIRU20, MIRU23, MIRU24, MIRU26, MIRU27, MIRU31, MIRU39, MIRU40, Mtu04, Mtub21, Mtub30, Mtub39, exact tandem repeats (ETR)-A, ETR-B, ETR-C, ETR-F, Queen's University Belfast (QUB) 11a, QUB11b, QUB26, and QUB4156) to determine their genetic relatedness. To improve the discriminatory power, two of the 24 loci (Mtub29 and Mtub34) reported by Supply et al. 2006 were replaced with ETR-F and QUB11a (Shah et al., 2017). Briefly, each locus was amplified individually using the specific primers for the flanking regions as described Supply et al. 2006. The PCR products were identified by gel electrophoresis with a 50 bp DNA ladder stained with ethidium bromide to visualize the bands for molecular weights determination. The number of tandem repeats was determined by comparing the size of the amplicons to the corresponding base pairs on the conversion table.

Hunter-Gaston Discriminatory Index (HGDI) for the respective loci was determined using the formula;

$$\mathrm{HGDI} = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} x_j (x_j - 1)$$

where N is the total number of studied isolates, x_j is the number of isolates with j^{th} pattern and s is the number of genetic patterns (Hunter, 1990; Hunter et al., 1988). The discriminatory index was defined as high when HGDI>0.6, moderate 0.3<HGDI<0.6, and HGDI <0.3 as poor. The cumulative HGDI was calculated by adding the MIRU-VNTR loci cumulatively starting with the highest and ending with the lowest discriminatory locus. The corresponding clustering rates, cluster size, and patterns identified according to each number of loci used were also calculated. This was done using BioNumerics and the resulting identified patterns were then used to calculate HGDI using the above formula. I further compared the cumulative HGDI for the selected 9 and 15 high discriminatory loci among the major SITs and all the strains in the study to propose a convenient set of loci that may be used for MDR-TB genotyping in this area.

Phylogenetic analysis

The resulting MIRU-VNTR profiles were analyzed using MIRU-VNTR*plus* (Allix-Beguec et al., 2008; Weniger et al., 2010) and visualized using ggtree package in R-studio (Team, 2020; Yu, 2020; Yu et al., 2018, 2017). To visualize the evolutionary relationships of the strains, a dendrogram was constructed by an unweighted pair group method with arithmetic averages (UPGMA) based on 24 loci MIRU-VNTR and spoligotyping patterns. Further, the minimum spanning tree (MST) was calculated using 24-loci MIRU-VNTR data with a single locus variant as the maximum locus difference within a clonal complex using BioNumerics version 7.6 (Applied Maths, Sint-Martens-Latem, Belgium). Overall, a cluster was defined as 2 or more strains with identical spoligotyping and/or MIRU-VNTR profiles. Clustering rate was calculated using the formulae n_c/N, where n_c is the number of clustered strains, and N is the total number of strains analyzed.

Comparison of CAS1-Kili and LAM1 strains

The SIT21/CAS1-Kili and SIT20/LAM1 strains from the SITVIT2 database which had standard 12 MIRU-VNTR loci and spoligotyping information, together with our strains were considered for phylogenetic analysis because of their association with MDR in Zambia. Other loci were excluded due to the non-availability of information from the database for many strains. Three and eleven more strains were obtained from Mbugi et al 2015 and Perdigão et al 2017 respectively. As a result, together with our strains, a total of 101 CAS1-Kili and LAM1 strains were analyzed.

Detection of RD-Rio sub-lineage strains

Further, lineage 4 strains were subjected to multiplex PCR for RD-Rio sub-lineage determination. Briefly, multiplex PCR was done using specific primers following David *et al* (2012). The presence of RD-Rio deletion or wild type was determined by the presence of 530 bp or 1175 bp bands, respectively after gel electrophoresis.

Ethical approval

The study was approved by the Zambia National Research Authority, ERES Converge Ethical committee and permission was granted by the University Teaching Hospital management.

Results

Demographic profile

The majority 65% (55/85) of strains were from males while 32% (27/85) from female patients. The age range was 9 to 75 years with a median of 33 years old. Most of the strains were from Lusaka 69% (59/85), and 22% (19/85) were from referred patients from outside Lusaka (from towns such as Mongu, Kafue, Chongwe, Kabwe, Sesheke, and Monze). While for 4% (3/85) patients' information was missing for sex and 8% (7/85) for age and residence, respectively.

Drug-resistance associated gene sequencing

All samples were confirmed to be multidrug resistant by the presence of resistance-associated mutations. Mutations in the *rpoB* gene among the strains included Ser 531 Leu (54%, 46/85) followed by His 526 Tyr (11%, 9/85), and others had mutations in other codons including double mutations and deletions. Whereas all of the strains had Ser 315 Thr mutation in *katG* and no mutation in the *inhA* promoter region as described previously by Solo et al. 2020 and in Figure 9.

Spoligotyping

Among the 85 samples, 5 sub-lineages; LAM 48% (41/85), CAS 29% (25/85), T 14% (12/85), X 6% (5/85) and Harlem 2% (2/85) were identified by spoligotyping (Table 2). The most predominant SIT was SIT21/CAS1-Kili 29% (25/85) followed by SIT59/LAM11-ZWE 19% (16/85). In addition, 11 distinct patterns were identified with a clustering rate of 99% (Table 2).

MIRU-VNTR genotyping and phylogeny

Two strains had multiple bands in more than two loci, suggestive of mixed infection and thus, were excluded from the analysis. Hence, 83 strains were considered for further analysis. From the 83 MDR-TB strains, 24-loci MIRU-VNTR identified 47 genotype patterns with the clustering rate of 63%. The number of strains per cluster ranged from 2 to 9. The clustering rates among the major spoligotypes SIT20/LAM1, SIT21/CAS1-Kili, and

SIT59/LAM11-ZWE based on 24-loci MIRU-VNTR were 75%, 76%, and 67%, respectively (Figure 10 and table 3).

The 24 loci varied in allelic diversity ranging from 0.0 to 0.8 among all our study strains as shown in Table 4. The 10 loci (MIRU10, MIRU16, QUB11b, ETR-A, MIRU23, MIRU26, MIRU31, QUB11a, QUB26, and QUB4156) and 4 loci (MIRU02, MIRU04, MIRU20, and MIRU24) exhibited high and low diversities among the strains in this study respectively. Besides, alias ETR-A showed no bands after gel electrophoresis for 96% of SIT 21/CAS1-Kili isolates (Table 4).

In addition, the first 15 high discriminatory loci varied slightly in the discriminatory index when compared to 24 MIRU-VNTR loci, (cumulative HGDI: 0.965 vs 0.973). Further, the combination of the two genotyping methods (24-MIRU-VNTR and spoligotyping) yielded the same clustering rate and cumulative HGDI as when 24-MIRU-VNTR was applied alone (Table 5). While, among the predominant SITs 59/LAM11-ZWE and 21/CAS1-Kili, the combined methods had cumulative HGDI of 0.914 and 0.833 respectively as shown in Table 3. The minimum spanning tree (MST) grouped the strains in 7 clonal complexes. The largest clonal complex consisted of clades, SIT21/CAS1-Kili with 24 isolates followed by SIT20/LAM1 with 10 isolates (Figure 10).

Comparison of CAS1-Kili and LAM1 strains

Zambian CAS1-Kili strains had identical 12-loci MIRU-VNTR and spoligotyping patterns with many strains from Somali, a strain from Ethiopia, Sudan, and South Africa as shown in Figure 12. While Zambian SIT20/LAM1 strains formed a monophyletic clade separate from other African strains.

Detection of RD-Rio sub-lineage strains

Among the 58 strains from lineage 4, 66% (38/58) were classified as RD-Rio sub-lineage strains by multiplex PCR. All the strains that had RD-Rio deletion belonged to the LAM clade. This included all LAM1 and LAM11-ZWE strains. Two strains from SIT42/LAM9 did not belong to RD-Rio sub-lineage. While two other strains from non-LAM (H1 clade) had no bands visualized on gel electrophoresis (Figure 13).

Discussion

Despite the emergence of MDR-TB and increasing cases in routine surveillance systems (Kapata et al., 2013), there is limited information regarding genetic diversity and transmission of MDR *M. tuberculosis* in Zambia. Moreover, drug susceptibility testing (DST) is not routinely done on all TB patients, thus, reported cases are based on selected DST done after clinicians' recommendations which is likely to underestimate the prevailing situation regarding MDR-TB. Therefore, this called for the study to elucidate the transmission phenomenon of MDR *M. tuberculosis* strains in Zambia.

Among the study strains, spoligotyping revealed a clustering rate of 99%. This was not surprising as this genotyping method may overestimate clustering due to its lower discriminatory power or was as a result of some selection bias in the strains. While 24 loci MIRU-VNTR delineated the 11 spoligotypes into 47 unique patterns, 16 clusters with a clustering rate of 63%. This is because MIRU-VNTR is superior to spoligotyping, and prevents homoplasy (Mokrousov et al., 2016; Supply et al., 2006).

Studies have described TB transmission based on clustering of strains using spoligotyping and MIRU-VNTR profiles (Easterbrook et al., 2004; Hamblion et al., 2016; Mulenga et al., 2010; Xu et al., 2018). Generally, strains with identical and unique genotypic profiles are as a result of recent transmission and independent acquisition of MDR, respectively. Hence, the clustering rate of 63% by 24 loci MIRU-VNTR and spoligotyping was suggestive of recent transmission playing a major role in the spread of MDR-TB. The limited number of reports applying 24-MIRU-VNTR on MDR-TB strains in Africa made it difficult to compare directly the clustering rates with other studies. Nevertheless, the clustering rate from this study was higher than other reports from Morocco (0%), Pakistan (15%), and China (33.1%) (Bakuła et al., 2019; Luo et al., 2019; Oudghiri et al., 2021). The recent transmission may be attributed more to the subfamilies SIT21/CAS1-Kili and SIT20/LAM1 as they formed the largest clonal complexes when compared to SIT59/LAM11-ZWE and others as shown in Figure 10. Moreover, all SIT20/LAM1 strains had KatG Ser315Thr and RpoB Ser 531 Leu mutations as shown in Figure 9. This combination of low-fitness-cost mutations has been described to increase transmission of MDR strains of the modern Beijing sublineage (Li et al., 2017).

Further, SIT21/CAS1-Kili and SIT20/LAM1 were previously reported to be associated with MDR-TB in Zambia (Solo et al. 2021). Therefore, having identical genotypic profiles and mutation combinations among some strains from SIT21/CAS1-Kili and SIT20/LAM1 is

indicative of clonal expansion of MDR-TB by these clades. Moreover, SIT20/LAM1 has been reported of having high proportions of strains among XDR-TB while, SIT21/CAS1-Kili has been associated with MDR-TB (Agonafir et al., 2010; David et al., 2012; Perdigão et al., 2019). Besides, the occurrence of MDR-TB in children in this study is a sign of active ongoing transmission in this area. Four percent (3/85) of patients were less than 15 years old and the majority (73%) were below the age of 40 years. The median age was 33 years similar to other studies in southern Africa (Sagonda et al., 2014; Said et al., 2012). On the other hand, some strains from patients living outside Lusaka, clustered with those from within the city (Figure 9). This may be attributed to the fact that patients referred from distant hospitals usually seek residence from relatives near the hospital to reduce on logistical costs which in turn facilitates the transmission of MDR-TB. Therefore, timely and effective contact tracing may reduce MDR-TB transmission through this mode. Further, two patients were supposedly family members as they had the same surnames and from within Lusaka, with samples collected in the space of 3 years but were clustered with the same mutations for rifampicin (RpoB Leu 511 Arg, Asp 516 Tyr) and isoniazid (KatG Ser315Thr) resistances as well as same spoligotyping and MIRU-VNTR patterns. However, a proportion of singletons illustrated that the conversion to MDR within each patient, such as in reactivation cases of latent tuberculosis, was also contributing to the MDR-TB population structure.

There were marginal differences regarding the discriminatory indices for 15 and 24 loci MIRU-VNTR (Table 3) in agreement with other studies (Perdigão et al., 2017; Silva et al., 2014). However, the discriminatory index of each locus differs according to geographical location and circulating strains in a region (Supply et al., 2006). Thus, the 24-loci were assessed to give a comprehensive analysis for the proposal of the loci set for MDR *M. tuberculosis* genotyping in this region. This was to suggest a set that is less labor-intensive and cheaper rather than 24 MIRU-VNTR loci. Hence, sets of the first 9 and 15-highly discriminatory MIRU-VNTR loci were suggested to be utilized for epidemiological and phylogenetic studies in this resource-limited region respectively (Table 3 and 5) (Hunter et al., 1988).

The three major clades SIT59/LAM11-ZWE, SIT21/CAS1-Kili, and SIT20/LAM1 varied markedly in their diversity indices with HGDI of 0.91, 0.83, and 0.67 respectively based on combined 24 MIRU-VNTR and spoligotyping (Table 3). The relatively lower HGDI for SIT21/CAS1-Kili and SIT20/LAM1, when compared to SIT59/LAM11_ZWE, emphasize the spread of these strains in this population. Nevertheless, more discriminatory methods such as WGS can elucidate more accurately the recent transmission of these clades.

The study showed lineage 4 and 3 as the major MDR *M. tuberculosis* strains in Zambia. The predominance of LAM clade was expected as this is the most common spreading strain for drug-susceptible tuberculosis in this region (Mulenga et al., 2010). Despite limited studies focusing on MDR-TB in Zambia and its neighboring countries, other studies in the southern African region had similar findings of LAM been predominant among MDR-TB strains (Namburete et al., 2020; Sagonda et al., 2014). Moreover, reports have indicated that lineage 4 predominates across Africa while lineage 3 occurs in larger proportions in the eastern than southern regions of Africa (Chihota et al., 2018). On the other hand, many studies on MDR *M. tuberculosis* from South Africa which doesn't share any borders with Zambia, had Beijing as the most common genotype with low occurrences of CAS clade below 5% (Figure 11) (Bhembe and Green, 2020; Chihota et al., 2012). Our current study revealed a relatively higher percentage (29%) of CAS sub-family than other south African countries among MDR *M. tuberculosis* strains (Figure 11). The CAS clade is more common in Tanzania which borders Zambia. The high percentage in this study is suggestive of the spread of the strain possibly due to migration of people because of trade with neighboring East African countries like Tanzania. Further, Zambia hosts and has incorporated thousands of refugees from Burundi and Rwanda (East African countries) into the local community (Commissioner and Grandi, 2019). Moreover, CAS1-Kili strains from this study had the same spoligotyping and standard 12-loci MIRU-VNTR patterns as those from Ethiopia, Sudan, Somali, and South Africa as shown in Figure 12. This may be suggestive of possible relations between these strains or due to poor discrimination of spoligotyping and standard 12-loci MIRU-VNTR genotyping methods.

In this study, 95% (38/40) of LAM strains had 1 tandem repeat in MIRU40, and 2 tandem repeats in MIRU2 loci (Figure 13), a VNTR signature for RD-Rio/RD 174 sub-lineages (Lazzarini et al., 2007; Mokrousov et al., 2016). All the 38 LAM strains were confirmed to belong to RD-Rio sub-lineage by multiplex PCR. The RD-Rio sub-lineage is common in Brazil and Portugal and has been associated with high transmissibility, MDR, and severe form of TB (David et al. 2012; Lazzarini et al. 2007; Mokrousov et al. 2016; De Almeida et al. 2019). It may be speculated that the ancestral strain might have come to Zambia through neighboring former Portuguese colonies such as Mozambique and Angola, which then gave rise to LAM11-ZWE that for some reason became endemic in this region (Brynildsrud et al., 2018; Mokrousov et al., 2016). Moreover, LAM1 and LAM9 spoligotypes are reported in high prevalence among LAM strains in Angola while LAM11_ZWE and LAM9 in Mozambique (Perdigão et al., 2017; Viegas et al., 2010). Mozambique and Angola had inadequate information on RD-Rio sub-lineage. Nevertheless, reports for RD174 that usually coexists with RD-Rio deletion were

found in these countries, (Mokrousov et al., 2016; Namburete et al., 2020; Perdigão et al., 2017) suggesting the likely presence of the RD-Rio sublineage.

Study limitations

Detailed epidemiological data and treatment history were inadequate. Hence, this affected a thorough evaluation of the spread and association with HIV/AIDS in the affected patients and communities. In addition, the smaller sample size hindered the generalization of the results. Hence, a large population study utilizing WGS, if possible, is needed to provide a comprehensive analysis of the circulating MDR *M. tuberculosis* strains.

Conclusion

The study revealed a high clustering rate that was suggestive of a large degree of MDR-TB spread attributed to recent transmission than independent acquisition of MDR. A 9-loci MIRU-VNTR was proposed to be utilized for screening of MDR *M. tuberculosis* lineages in this area to reduce on labor and cost associated with 24-MIRU-VNTR loci. Further, the study illustrated the clonal expansion of SIT20/LAM1 and SIT21/CAS1-Kili indicating their major role in the spread of MDR-TB in Lusaka Zambia. Therefore, this study recommends that TB control programs focusing on genotyping coupled with other measures such as early diagnosis, monitored treatment, and contact tracing of MDR-TB patients must be enhanced. This will facilitate the understanding of the transmission dynamics of MDR-TB and hence guide the implementation of appropriate control measures.

Summary

Zambia is among the 30 high tuberculosis burden countries in the world. Despite increasing reports of MDR-TB in routine surveillance, information on the transmission of MDR *Mycobacterium tuberculosis* strains is largely unknown. This study elucidated genetic diversity and transmission of MDR *M. tuberculosis* strains in Lusaka, Zambia. Eighty-five MDR *M. tuberculosis* samples collected from the year 2013 to 2017 at the University Teaching Hospital were used. Drug-resistance associated gene sequencing, spoligotyping, 24-loci MIRU-VNTR, and multiplex PCR for RD-Rio sub-lineage identification were applied.

Clades identified were LAM (48%), CAS (29%), T (14%), X (6%) and Harlem (2%). Strains belonging to SITs 21/CAS1-Kili and 20/LAM1 formed the largest clonal complexes. Combined spoligotyping and 24 loci-MIRU-VNTR revealed 47 genotypic patterns with clustering rate of 63%. Ninety five percent of LAM strains belonged to RD-Rio sub-lineage. The high clustering rate suggested that a large proportion of MDR-TB was due to recent transmission rather than independent acquisition of MDR. This spread was attributed to clonal expansion of SIT21/CAS1-Kili and SIT20/LAM1 strains. Therefore, TB control programs recommending genotyping coupled with conventional epidemiological methods can guide measures for stopping the spread of MDR-TB.

Table 2. Table showing spoligotype-based diversity of the 85 strains under the study

Lineage	Sub-family	SIT	Number of strains	SIT (%)	Clade cumulative (%)
L3	CAS1-Kili	21	25	29	29
L4	LAM1	20	12	14	
	LAM9	42	2	2	
		59	16	19	48
	LAM11_ZWE	815	9	11	
		Orphan	2	2	
	H1	Orphan	2	2	2
	T1	53	8	9	
	T2	52	3	4	14
	T2	317	1	1	
	X2	137	5	6	6

Table 3. MIRU-VNTR discriminatory powers for 9, 15, and 24 loci on predominant clades in our study

		Identified patterns					**Clustering rate**					**Cumulative HGDI**				
			MIRU-VNTR loci					MIRU-VNTR loci					MIRU-VNTR loci			
Clade		a	9	15	24	b	a	9	15	24	b	a	9	15	24	b
21/CAS1_Kili	25	1	8	9	10	10	100	80	76	76	76	0	0.727	0.743	0.833	0.833
59/LAM11_ZWE	15	1	9	9	9	9	100	67	67	67	67	0	0.914	0.914	0.914	0.914
20/LAM1	12	1	4	5	5	5	100	83	75	75	75	0	0.561	0.667	0.667	0.667

a spoligotyping

b 24 MIRU-VNTR loci + spoligotyping

Table 4. Allelic diversity for each of the 24 loci MIRU-VNTR for 87 MDR-TB isolates

Locus	Alias	Number of tandem repeats											HGDI	Designation
		0	1	2	3	4	5	6	7	8	9	10 ND[a]		
4052	QUB26			4	5	8	26	4	17	17		2	0.81	High
b	QUB11a			6	12	26		5			25	9	0.78	High
1644	MIRU16		27	13	20	20			3				0.76	High
960	MIRU10			1	30	21	3	4	24				0.73	High
2531	MIRU23			7			30	34	12				0.68	High
2996	MIRU26		25		2	4	40	12					0.66	High
2165	ETR-A			40	15	4						24[c]	0.66	High
3192	MIRU31			11	39	28	4					1	0.65	High
2163	QUB11b		2	44	13	21	3						0.64	High
4156	QUB4156			47	9	22						5	0.60	High
3007	MIRU27		2	36	40	5							0.58	Moderate
2401	Mtub30		40	36		7							0.58	Moderate
802	MIRU40		45		6	31	1						0.57	Moderate
3690	Mtub39			52	22		7	2					0.54	Moderate
424	Mtub04			15	9	57	2						0.49	Moderate
577	ETR-C			20	2	57	4						0.47	Moderate
b	ETR-F			29	54								0.46	Moderate
4348	MIRU39			56	27								0.44	Moderate
2461	ETR-B		26	57									0.44	Moderate
1955	Mtub21			17	66								0.33	Moderate
154	MIRU02		2	81									0.05	Poor
2687	MIRU24		83										0.0	Poor
580	MIRU04			83									0.0	Poor
2059	MIRU20			87									0.0	Poor

[a] ND-not defined: no band after gel electrophoresis
b -non-standard 24 Loci by Supply et al. 2006
[c] ETR-A failed to amplify 96 % isolates of SIT 21/CAS1-Kili

Table 5. Cumulative discriminatory index for 24-MIRU-VNTR loci and combined with spoligotyping.

[a]Number	Loci	Alias	Number of patterns	Number of clusters	Cluster size	No. of strains clustered	CR[b] (%)	HGDI	[c]cHGDI
1	*4052*	QUB26	8	8	4-26	83	100	0.81	0.810
2	**d**	QUB11a	13	10	2-25	80	96	0.78	0.856
3	*1644*	MIRU16	20	12	2-20	75	90	0.76	0.900
4	*960*	MIRU10	27	16	2-19	72	87	0.73	0.922
5	*2531*	MIRU 23	28	18	2-13	73	88	0.68	0.945
6	*2996*	MIRU26	30	17	2-13	70	84	0.66	0.945
7	*2165*	ETR-A	30	17	2-13	70	84	0.66	0.945
8	*3192*	MIRU31	35	15	2-12	63	76	0.65	0.952
[e]9	***2163***	**QUB11b**	**40**	**15**	**2-12**	**58**	**70**	**0.64**	**0.959**
10	*4156*	QUB4156	40	15	2-12	58	70	0.60	0.959
11	*3007*	MIRU 27	41	16	2-12	57	69	0.58	0.959
12	*2401*	Mtub30	41	16	2-12	57	69	0.58	0.959
13	*802*	MIRU40	42	16	2-12	56	67	0.57	0.961
14	*3690*	Mtub39	43	15	2-12	55	66	0.54	0.961
[e]15	***424***	**Mtu04**	**46**	**15**	**2-12**	**52**	**63**	**0.49**	**0.965**
16	*577*	ETR-C	47	16	2-9	52	63	0.47	0.973
17	**d**	ETR-F	47	16	2-9	52	63	0.46	0.973
18	*4348*	MIRU 39	47	16	2-9	52	63	0.44	0.973
19	*2461*	ETR-B	47	16	2-9	52	63	0.44	0.973
20	*1955*	Mtub21	47	16	2-9	52	63	0.33	0.973
21	*154*	MIRU 02	47	16	2-9	52	63	0.05	0.973
22	*2687*	MIRU 24	47	16	2-9	52	63	0.0	0.973
23	*580*	MIRU04	47	16	2-9	52	63	0.0	0.973
[e]24	***2059***	**MIRU 20**	**47**	**16**	**2-9**	**52**	**63**	**0.0**	**0.973**
[e]24 loci *plus* spoligotyping			**47**	**16**	**2-9**	**52**	**63**	**0.974**	**0.973**

[a]The columns represent the following: Number is the cumulative number of loci used. Loci/Alias represents names of the individual MIRU-VNTR locus. Number of patterns, clusters and cHGDI columns show the increase of these parameters with the addition of each MIRU-VNTR locus. Cluster size, number of clustered strains and clustering rate columns show the reduction of these parameters with the addition of each MIRU-VNTR locus. While HGDI column represents the individual HGDI of each locus starting with the locus with the highest discriminatory power.

[b] Clustering rate

[c] cHGDI-cumulative Hunter-Gaston discriminatory index

d -non-standard 24 Loci by Supply et al., 2006

[e] Bold indicate results for cumulative 9, 15, 24, and 24 MIRU-VNTR loci plus spoligotyping

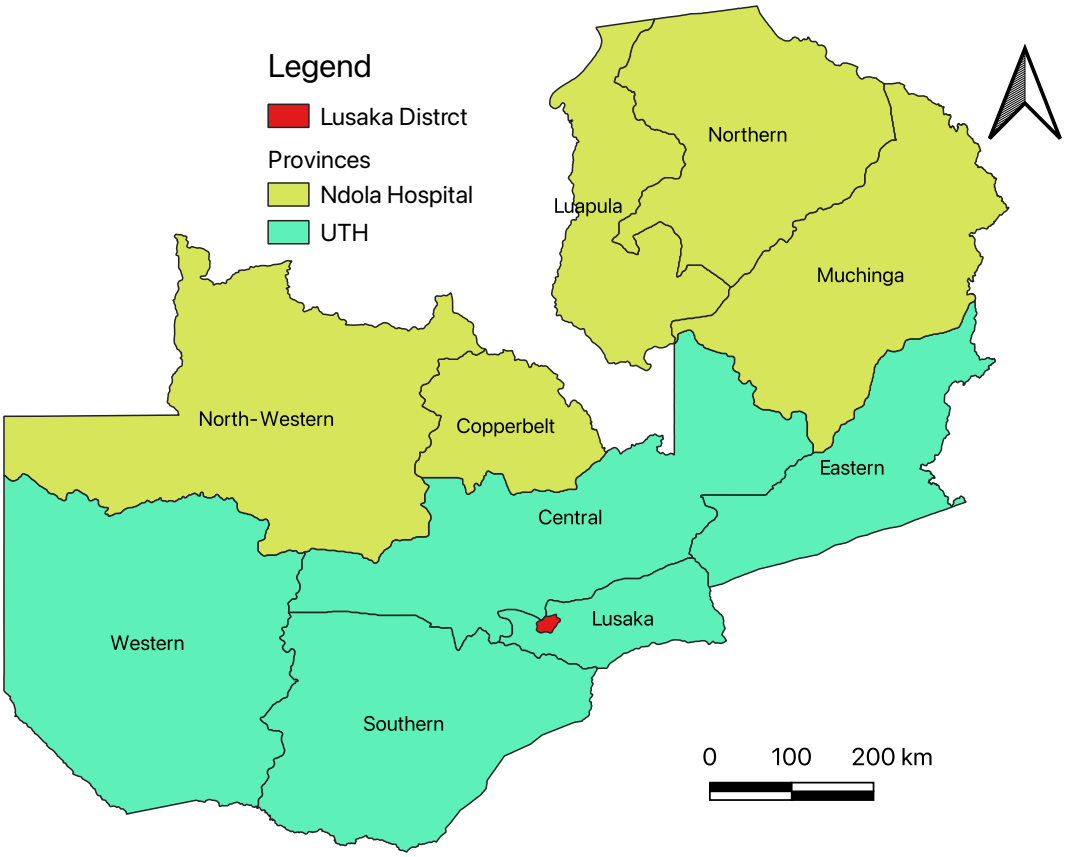


Figure 8. Map showing regions (in green) referring patients and samples to the University Teaching Hospital in Lusaka, Zambia. The regions (yellow) refer patients to Ndola Central Hospital.

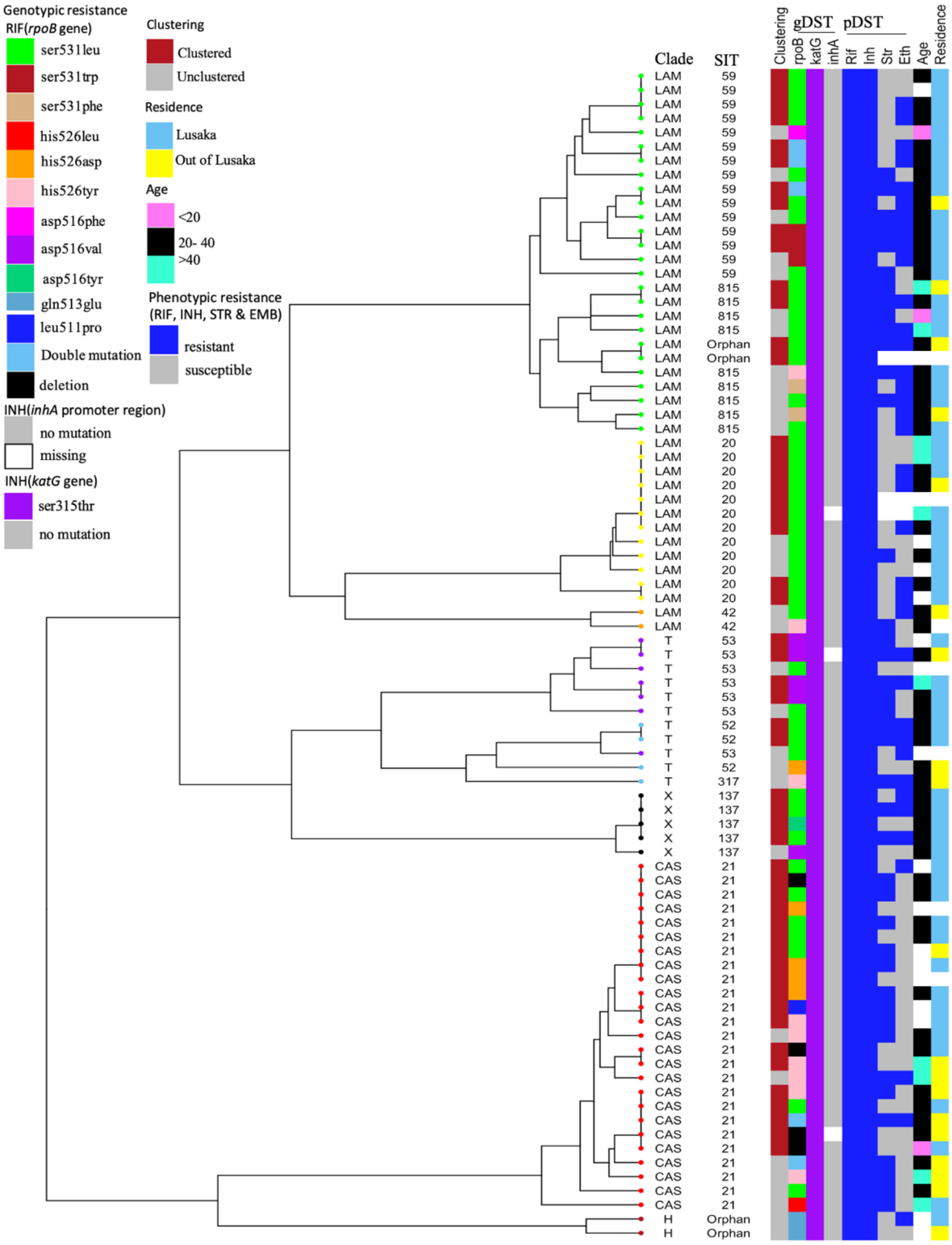


Figure 9. Dendrogram based on 24 loci MIRU-VNTR and spoligotyping patterns constructed by (UPGMA). pDST and gDST = phenotypic and genotypic drug susceptibility.

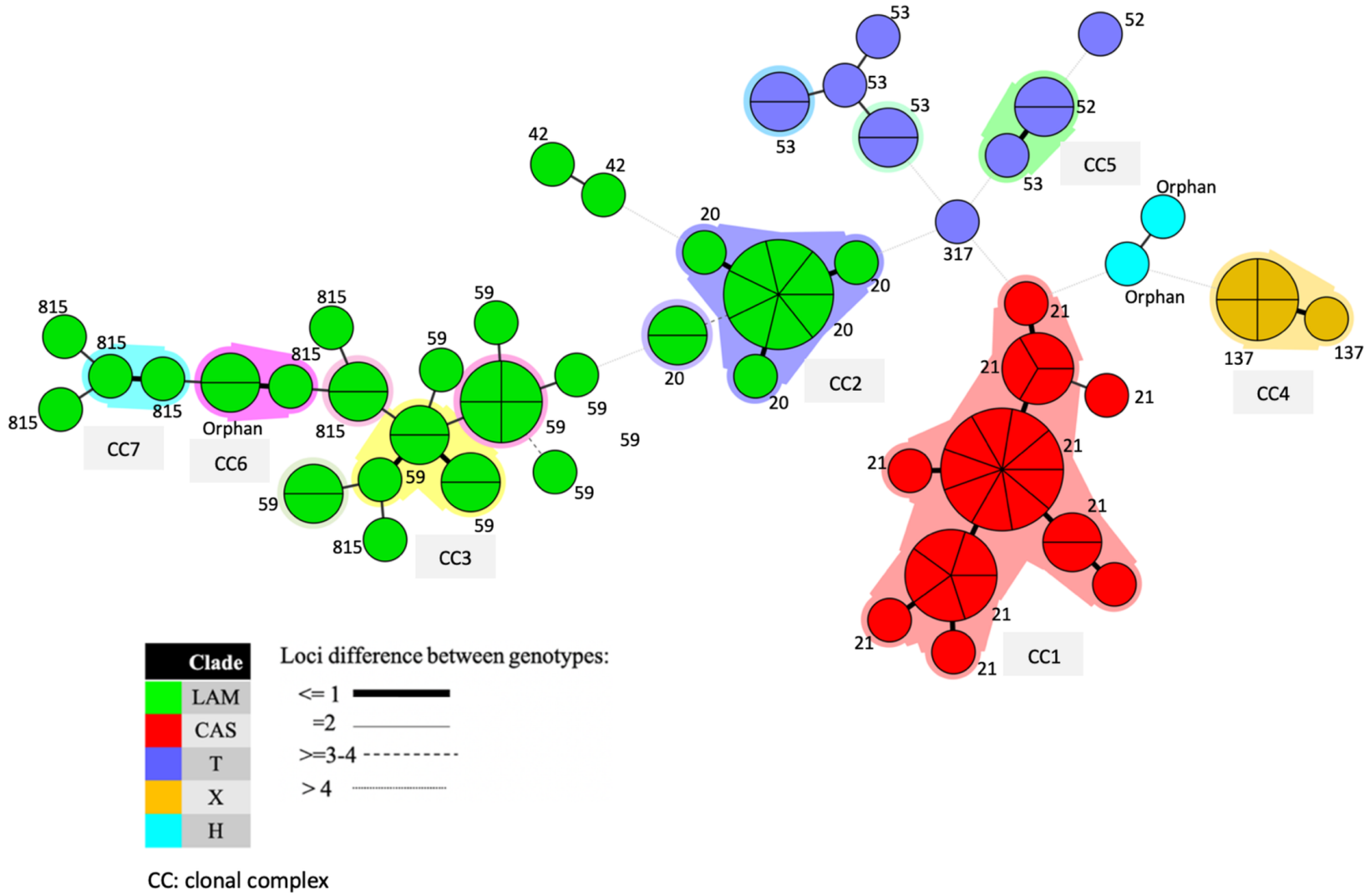
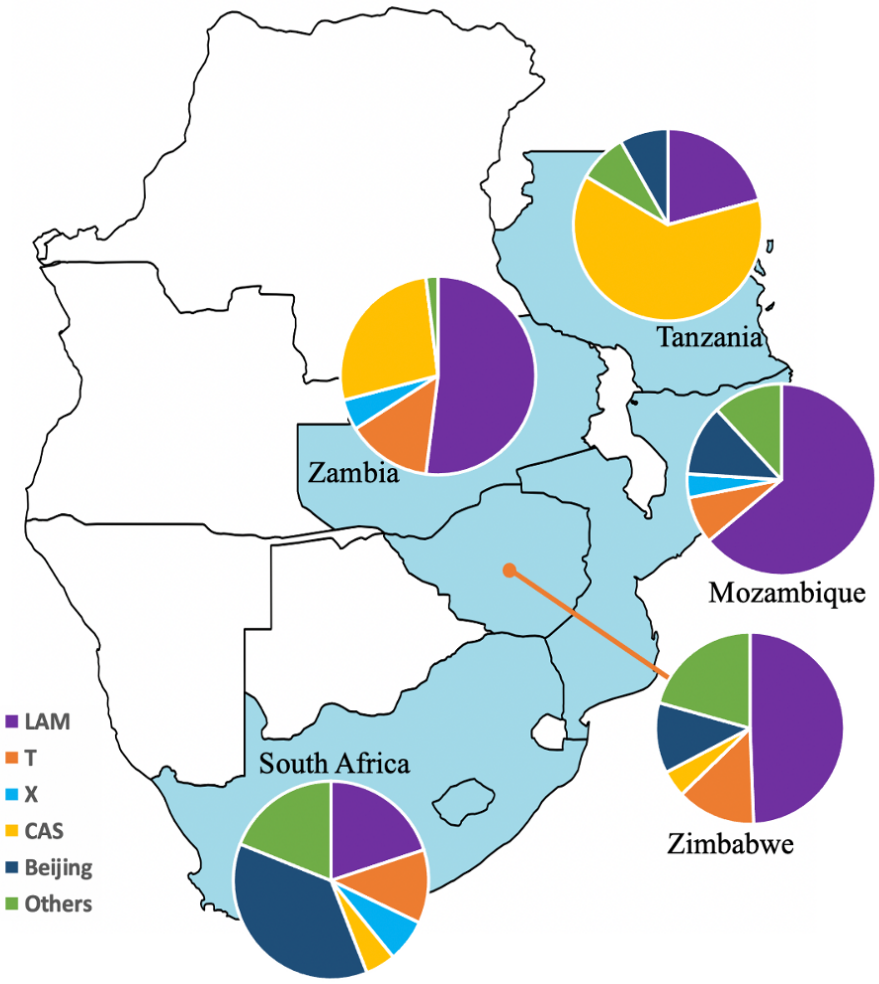


Figure 10. Minimum spanning tree constructed using 24-loci MIRU-VNTR profile. Strains with single-locus differences were grouped in 7 clonal complexes. Line thickness indicates loci differences. The variations in size and partitions in each node correspond to the number of strains. Figures on the nodes are SIT numbers.



Clade	Range (%)
LAM	20-64
T	0-14
X	0-8
CAS	0-63
Beijing	0-37
Others	1-21

(Chihota et al., 2012; Katale et al., 2020; Namburete et al., 2020; Sagonda et al., 2014)

Figure 11. Figure showing the distribution of MDR M. tuberculosis strains according to sub-lineages in Zambia and its neighboring countries. The countries in white implies that the data was not captured or was not adequate.

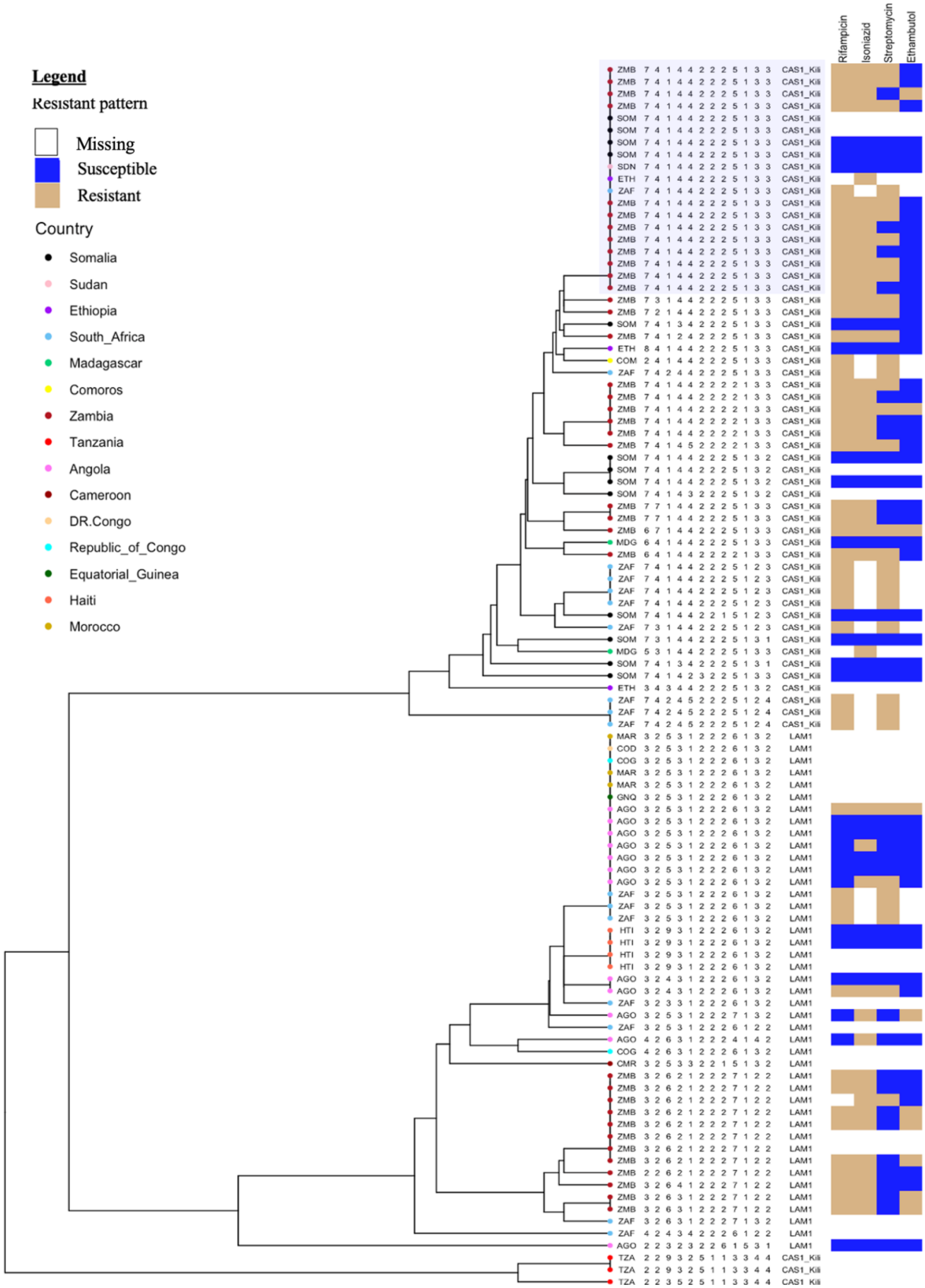


Figure 12. Dendrogram showing the relationship of the Zambian strains (SIT20/LAM1 and SIT21/CAS1-Kili) with African strains from SITVIT 2 database, Mbugi et al. 2015 and Perdigão et al. 2017 based on standard 12 loci MIRU-VNTR. The strains had same spoligotyping octal patterns; 703377400001771 for SIT21/CAS1-Kili and 677777607760771 for SIT20/LAM1.

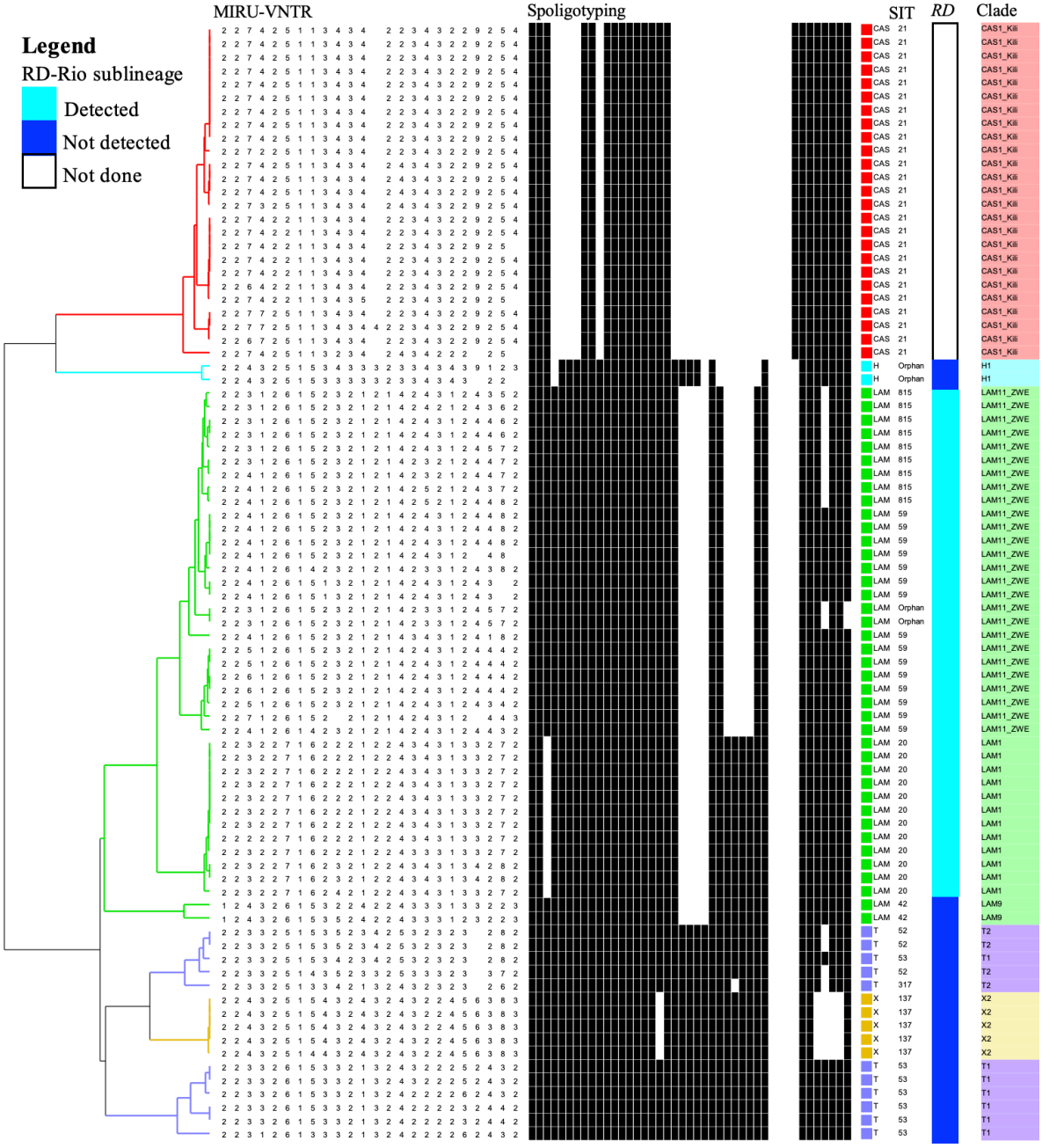


Figure 13. Dendrogram showing 24-MIRU-VNTR and spoligotyping patterns. The order of 24-MIRU-VNTR from left to right: MIRU02, MIRU04, MIRU10, MIRU16, MIRU20, MIRU23, MIRU24, MIRU26, MIRU27, MIRU31, MIRU39, MIRU40, ETR-A, ETR-B, ETR-C, ETR-F, Mtub04, Mtub21, Mtub30, Mtub39, QUB11a, QUB11b, QUB26, and QUB4156. *RD* represents RD-Rio sublineage.

CHAPTER 2

Whole-genome sequencing reveals recent transmission of multidrug-resistant *Mycobacterium tuberculosis* Lineage 3 CAS1-Kili strains in Lusaka, Zambia

Introduction

Worldwide, close to 4000 lives in a day are lost due to tuberculosis (TB). In Zambia, like many other developing countries, TB is the leading cause of death especially among people living with HIV/AIDS (UNAIDS, 2018).

Mycobacterium tuberculosis, the cause of TB has seven major human-adapted lineages. One of these is lineage 3 (L3) which is predominant in East Africa, the Middle East, and South Asia (Chihota et al., 2018; O'Neill et al., 2019). Central Asian Strain 1- Kili (CAS1-Kili) forms part of lineage 3 sub-clades and is more prevalent in the eastern part of Africa (Chihota et al., 2018; Gagneux et al., 2006) (Figure 14).

Molecular epidemiological methods such as MIRU-VNTR, spoligotyping, and IS6110-RFLP have been applied to understand *M. tuberculosis* transmission based on genotypic clustering (Easterbrook et al., 2004; Hamblion et al., 2016; Xu et al., 2018). Nevertheless, these genotypic methods though targeting polymorphic sites, only interrogate one percent of the *M. tuberculosis* genome and have limited discriminatory power (Coll I Cerezo, 2015; Luo et al., 2014; Roetzer et al., 2013). Therefore, WGS provides an ultimate method for high resolution of strain relatedness and investigation of recent transmission (Walker et al., 2013). Further, it explores more drug-resistant mutations that occur outside the targeted regions by PCR-based assays (Feliciano et al., 2018). Recently, WGS has become affordable, though it is still a very big challenge for developing nations. Based on WGS results, though not yet standardized, many reports have described strains belonging to a cluster and been involved in a recent transmission event when they differ by twelve or fewer SNPs and five or fewer SNPs respectively in their core genomes (Lalor et al., 2018; Walker et al., 2013).

Previously we reported a relatively higher percentage of lineage 3 Spoligotype International Types (SIT) 21/CAS1-Kili strains among multi-drug resistant (MDR) *M. tuberculosis* strains in Zambia when compared to other southern African countries (Chizimu et al., 2021). Despite its low proportions in the region when compared to predominant lineage 4, lineage 3 strains (SIT21/CAS1-Kili) were associated with MDR TB and implicated for its

spread in Lusaka, Zambia (Solo et al. 2021; Chizimu et al. 2021). Among the studied MDR *M. tuberculosis* strains, 24-loci MIRU-VNTR and spoligotyping showed clonal expansion of the SIT21/CAS1-Kili strains (Chizimu et al., 2021). The findings in those studies supported the possibility of recent transmission though this could not be concluded due to the limitations of the used genotyping methods in discriminating between closely related strains. Thus, this study aims to investigate recent transmission among the MDR *M. tuberculosis* SIT21/CAS1-Kili strains in Lusaka, Zambia, by WGS.

Materials and Methods

Study samples

I used strains from the previous study conducted in Lusaka from 2013 to 2017 (Chizimu et al., 2021). In the previous study, 87 MDR-TB strains were typed by spoligotyping and 24-loci MIRU-VNTR, of which 25 CAS1-Kili strains formed the largest clonal cluster and were suspected to represent a recent transmission event (Chizimu et al., 2021). All CAS1-Kili strains were considered for WGS though only 13 were successfully sequenced. The other 12 strains had poor DNA quality, hence excluded from the current study.

Culturing and drug susceptibility testing

Briefly, sub-culturing of the isolates was done in BACTEC™ 960 MGIT™ (Mycobacteria Growth Indicator Tube) system (Becton Dickson & Co., Franklin Lakes, NJ, USA) following the manufacturer's instructions. The critical concentrations for isoniazid, rifampicin, streptomycin, and ethambutol, as prescribed by the kit manufacturer (Becton, Dickinson, and Company), of 0.1, 1.0, 1.0, and 5.0 g/mL were used, respectively.

DNA extraction and genotyping

DNA was extracted by heating method as described previously (Chizimu et al., 2021; Solo et al., 2021). The extracted DNA was transported to the Hokkaido University International Institute for Zoonosis Control, Japan for analysis. Further, 24-loci MIRU-VNTR (Supply et al., 2006) and spoligotyping (Kamerbeek et al., 1997) were performed as previously described (Chizimu et al. 2021; Solo et al. 2021).

Whole-genome sequencing analysis

Library preparations were performed following Nextera XT DNA Library Preparation Kit (Illumina Inc., San Diego, CA, USA) manufacturer's instructions, and Illumina MiSeq 2500 platform was used for sequencing. WGS data analysis was carried out as previously described (Batut et al., 2018; Heusden et al., 2021). The raw paired reads were checked for quality using fastqc, results compiled using multiQC (Ewels et al., 2016). Trimmomatic was

used to trim adapters, low-quality bases and filtering for a minimum read length of 20 (SLIDINGWINDOW:4:20 MINLEN:20) (Bolger et al., 2014). Variants were called using Snippy pipeline (Seemann. T, 2020); briefly, reads were aligned to the reference strain H37Rv (NC_000962.3) by Burrows Wheel Aligner (BWA), manipulated with SAMtools, and variant calling with freebayes, while variant annotation was performed with SnpEff. Further, Snippy-core generated full and core genome alignments. Gubbins (Croucher et al., 2014) was further used to generate filtered polymorphic sites.

Therefore, filtered SNPs in variable and invariable sites were used to calculate a pairwise distance matrix and maximum likelihood phylogenetic tree using RAxMLv8.2.11 in Geneious v 10.2.6, Biomatters, Ltd., Auckland, New Zealand https://www.geneious.com (accessed on Dec 16, 2021). The mean coverages of the analyzed reads for 12 CAS1-Kili strains ranged from 35.73 to 139.32 (Table 6). One strain (strain 02) was excluded due to poor coverage. Resistance patterns and sub-lineages were determined by TB Profiler (Coll et al., 2015; Phelan et al., 2019) and PhyResSE (Feuerriegel et al., 2015). We further manually confirmed the variants using CLC Genomics workbench 10 (QIAGEN, Hilden, Germany) (CLC Genomics Workbench, 2013). To understand the likely pattern of drug-resistant mutations acquisition by the strains, the drug-resistant mutations were conveniently plotted on the phylogenetic tree.

Phylogenetic assessment of global sub-lineage CAS1-Kili (L3.1.1)

Sub-lineage 3.1.1 strain's accession numbers were obtained from TB-Profiler, projects PRJEB29435 (Katale et al., 2020) and PRJEB33896 (Shuaib et al., 2020). A total of 527 pairs of raw reads were downloaded from the ENA browser and 480 of these were successfully analyzed. All raw reads were processed as described above. Additionally, to avoid false hits in the repetitive regions, SNPs in Proline-glutamate (PE) or proline-proline-glutamate (PPE) gene families were filtered by *M. tuberculosis* BED mask file included in Snippy-package. Further, Gubbins generated filtered polymorphic sites and a maximum likelihood phylogenetic tree rooted to *M. canetti*. The phylogenetic tree was trimmed of *M. canetti* but topology maintained, for better visualization using ggtree (Team, 2020; Yu, 2020; Yu et al., 2018, 2017). To confirm the sub-lineage and drug resistance patterns for some strains, TB-Profiler was used. Further, SNP distances using Ape package (Paradis E, Claude J, 2004) and Wilcoxon signed-rank test were calculated in R-studio (Team, 2020). Only countries with more than 10 strains were considered for SNP distance calculation.

Results

Cluster analysis

Among our Zambian strains, 11/12 (92%) differed by not more than 12 SNPs, while 6 (50%) strains differed with at least one other strain by 5 or fewer SNPs. Strain 7 was closely related to strains 4, 5, and 11. Additionally, strain 7 had the lowest average SNP difference of 9 SNPs to other strains, as shown in Table 7. On the contrary, strain 6 had the highest average SNP difference of 23 SNPs to the other strains.

Resistance patterns and Phylogeny

All strains had the same mutations for isoniazid resistance (KatG, Ser 315 Thr). With the exception of one strain with EmbB Met 306 Leu (Strain 6, Figure 15), 92% (11/12) had the same mutation associated with ethambutol resistance (Met 306 Ile) though only 25% (3/12) had the corresponding phenotypic resistance, as shown in Figure 15. Among 11 EmbB Met 306 Ile mutants, 36% (4/11) strains had additional mutations in *embA,* and three of these were phenotypically resistant to ethambutol. Additionally, all the 12 strains had resistance-associated mutations to streptomycin. Of these twelve strains, four and one had mutations in *rpsL* and *rrs* genes, respectively, and all were resistant to streptomycin, while 7/12 had mutations in *gid* and only two out of those seven were resistant to streptomycin. Overall, strains had different mutations towards rifampicin resistance while all RpoB Ser 450 Leu mutants had compensatory mutations in RpoC Val 483 (Figure 16). Among four strains having RpoB Asp 435 deletion, three had a mutation in RpoB Thr 1047 Ile and one had that in RpoC Trp 105 Arg. The strains seemed to have accumulated drug-resistance-associated mutations in a stepwise pattern (Figure 15). One strain (strain 3) had both -16 C > T and -16 C > G nucleotide variants in the *embA* gene. Strains having the same mutations for drug resistance to several drugs had a correspondingly small number of SNP differences between them (Table 7, Figures 15 and 16).

Phylogenetic assessment of global sub-lineage CAS1-Kili (L3.1.1)

Most of the lineage 3.1.1 (SIT21/CAS1-Kili) strains were from Europe, 52% (250/480), followed by Africa, 48% (228/480) (Table 8). Zambian strains formed a monophyletic clade,

closer to Malawian strains but both descending from the Tanzanian strains (Figure 17a and b). The clade having the Zambian strains also had strains from the United Kingdom (Figure 17b). The majority of the global strains were pan-susceptible (78%, 375/480), while 13% (60/480) were MDR and 2% (9/480) were resistant to other anti-tuberculosis drugs (Figures 17a and 19). Moreover, Zambian strains had the lowest median SNP distance when compared to strains from other countries ($p < 0.000$) (Figure 18).

Discussion

The SIT21/CAS1-Kili *M. tuberculosis* strains were previously shown to be associated with the clonal spread of MDR-TB in the Lusaka District (Chizimu et al., 2021; Solo et al., 2021). In this study, we applied WGS to investigate recent transmission events among SIT21/CAS1-Kili strains that was suggested by the combination of 24-MIRU-VNTR loci and spoligotyping. Of the 12 evaluated SIT21/CAS1-Kili strains, 11 belonged to a cluster, as they differed by 12 or fewer SNPs in their core genomes, while 6 strains were involved in recent transmission events as they differed by ≤ 5 SNPs to at least one other strain. Therefore, WGS revealed that the SIT21/CAS1-Kili strains were closely related as they exhibited comparatively low variabilities in SNPs. In addition, several strains shared the same combinations of drug-resistance-associated mutations to isoniazid, ethambutol, rifampicin, and other anti-TB drugs. This provided more evidence on the clonal spread of MDR SIT21/CAS1-Kili strains in this region, calling for interventions to stop this possible outbreak.

Despite some reports indicating the reduced fitness as a result of drug resistance evolution (Andersson, 2006; Asare et al., 2020), others have indicated emanating high fitness strains with preserved ability to spread (Leung et al., 2013; Metcalfe et al., 2010). Thus, the latter might explain the scenario in our current study. In addition, 75% (9/12) of strains had additional mutations besides the major resistance-associated mutations to rifampicin in *rpoB*, of which 33% (4/12) of strains with RpoB Ser 450 Leu mutation also had compensatory mutations in RpoC, Val 483 Ala (2/4) and Val 483 Gly (2/4). These compensatory mutations have been implicated in mitigating the fitness defects caused by RpoB Ser450Leu substitution in *M. tuberculosis* (Comas et al., 2011) which in turn contribute to the successful spread of MDR-TB. Interestingly, other strains had novel mutations in *rpoB* (3/12, Thr 1074 Ile) and *rpoC* (1/12, Trp 105 Arg) which were suspected to be compensatory as they occurred in association with a known mutation, Asp 435 deletion (4/12), in the *rpoB* RRDR. However, this can only be confirmed by allelic exchange experiments.

Similarly, 33% (4/12) of strains had mutations in more than a single resistance-conferring gene to ethambutol. The phenomenon has been reported to be associated with high resistance to ethambutol (Bakuła et al., 2013; Li et al., 2020; Qing et al., 2021; Safi et al., 2008). In fact, all three strains which were phenotypically resistant to ethambutol in this study had multiple mutations towards ethambutol resistance. In addition, one of the three EMB-resistant strains (strain 8) had a novel mutation in the *embA* gene regulatory region (-35 A > deletion), whilst one other strain (strain 3) had -16 C > T and -16 C > G nucleotide variants in *embA*,

which was suggestive of ongoing evolution within the strain. All the 12 MDR strains had mutations in codon *embB* 306 which has been associated with isoniazid resistance (Parsons et al., 2005), MDR, and more likely to evolve to XDR-TB (Bakuła et al., 2013; Hazbón et al., 2005; Safi et al., 2008).

Further, codon *embB* 306 mutations have been reported to reduce susceptibility of *M. tuberculosis* to other drugs used i. Therefore, based on this understanding, it can be speculated that possibly the strains first acquired drug resistance to isoniazid (KatG, Ser 315 Thr) and ethambutol (*EmbB* 306) then independently became resistant to other drugs, rifampicin, streptomycin, and pyrazinamide, while propagating to the next patients (Figure 15). As a result, rapid and low-cost diagnostic techniques for detecting *embB* 306 mutations in drug-resistant strains, particularly in high-burden drug-resistant TB areas, are encouraged to guide patient management and control of drug-resistant TB (Safi et al., 2008; Shen et al., 2007; Shuaib et al., 2020).

Even though some strains were resistant to all the first-line anti-TB drugs and streptomycin, none had resistance mutations to other second-line drugs. This at least left some treatment options for these patients, though on the other hand, calls for improved resistance detection and individualized treatment. The observed discrepancy in phenotypic and genotypic drug susceptibility patterns for ethambutol can be attributed to lack of consistency of phenotypic DST to ethambutol resistance determination as indicated by several reports (Hazbón et al., 2005; Leung et al., 2013).

Mutations in *embB* 306, which is associated with the majority of drug-resistant strains to ethambutol, have also been detected in susceptible strains (Shen et al., 2007), whilst the same discrepancy for streptomycin might be due to low-level resistance-conferring mutations in *gid* to the drug, which may result in susceptible phenotype (Wong et al., 2011). Mutations in *rrs* and *rpsL* genes have been associated with intermediate to high (Meier et al., 1996) and high levels of resistance to streptomycin (Bwalya et al., 2021; Meier et al., 1996; Sun et al., 2016; Wong et al., 2011), respectively. In addition, drug susceptibility testing is not routinely performed on all patients with TB, hence the treatment regimen in some cases may not be appropriate and this facilitates resistance amplification and further transmission of primary drug-resistant TB (Shuaib et al., 2020).

Phylogeny showed that some strains involved in recent transmission events were from patients residing in different districts (Monze, Chikankata) and Lusaka City (Figures 16 and 20). This conceivably suggests that there could be many unknown related intermediary cases

not captured by this study hence posing a threat for more outbreaks in the future. Therefore, this calls for more improved contact tracing strategies to curtail this transmission.

Further, some previously identical strains by MIRU-VNTR and spoligotyping were delineated as unique strains by WGS, indicating its supremacy to traditional genotyping methods (Alaridah et al., 2019; Xu et al., 2019). Though the application of WGS on TB epidemiology is still challenging in developing nations like Zambia, its combination where possible, with traditional genotyping methods, can make it cost-effective and be utilized as a supporting tool, especially for large-size population studies.

Generally, the global SIT21/CAS1-Kili (L3.1.1) strains were pan-susceptible to anti-TB drugs (Figure 17a and 19). Therefore, the outbreak of drug resistance among Zambian strains might be due to local factors such as late diagnosis, poor compliance, incomplete contact investigations, or other unknown reasons in the TB control system. Interestingly, the relationship of our Zambian strains to those from Malawi and Tanzania was suggestive of the possible origin of this strain to be Tanzania, in agreement with Chihota et al., 2018 (Chihota et al., 2018). Then, it spread to Malawi before Zambia. This finding also illustrated that the TB structure in a country is likely to be influenced by TB events in the neighboring countries in addition to local factors. Therefore, regional coordination in TB control is cardinal as movements of people for trade, migration, and refuge facilitate the spread of TB (Alyamani et al., 2019; Casper et al., 1996). SIT21/CAS1-Kili (L3.1.1) strains related to the Zambian and other African countries are also causing TB disease in European countries such as the United Kingdom, as suggested by the existence of these strains in the same clade. This spread can be attributed to the movements of people from the eastern part of Africa to Europe for economic activities and refuge (Alaridah et al., 2019; Walker et al., 2018).

The study limitations included small sample size due to poor quality DNA for some samples and inadequate patient history to precisely infer the direction for person-to-person transmission. In addition, a larger sample size with a longer collection period of the strains would have facilitated the determination of the emergency of MDR-TB and its hotspots in Lusaka. Additionally, the sample size also affected the generalization of these results. Therefore, we recommend MDR-TB surveillance on the large scale, whenever possible, to provide more information on the spread of this strain and other genotypes in the community, which can inform policy.

Conclusions

The study revealed ongoing transmission of MDR SIT21/CAS1-Kili strains in Lusaka, Zambia, which is a public health concern and needs more evaluation. It also showed the high resolution of WGS in delineating closely related strains and determining antibiotic resistance. Further, the study supports the complimentary use of WGS with traditional methods. Intensified case finding, improved drug resistance detection and adherence to treatment can help interrupt the transmission chains.

Summary

Globally, tuberculosis (TB) is a major cause of death due to antimicrobial resistance. *Mycobacterium tuberculosis* CAS1-Kili strains that belong to lineage 3 (Central Asian Strain, CAS) were previously implicated in the spread of multidrug-resistant (MDR)-TB in Lusaka, Zambia. Thus, we investigated recent transmission of those strains by whole-genome sequencing (WGS) with Illumina MiSeq platform. Twelve MDR CAS1-Kili isolates clustered by traditional methods (MIRU-VNTR and spoligotyping) were used. A total of 92% (11/12) of isolates belonged to a cluster (≤12 SNPs) while 50% (6/12) were involved in recent transmission events, as they differed by ≤5 SNPs. All the isolates had KatG Ser 315 Thr (isoniazid resistance), EmbB Met306 substitutions (ethambutol resistance) and several kinds of *rpoB* mutations (rifampicin resistance). WGS also revealed compensatory mutations including a novel deletion in *embA* regulatory region (−35A > del). Several strains shared the same combinations of drug-resistance-associated mutations indicating transmission of MDR strains. Zambian strains belonged to the same clade as Tanzanian, Malawian and European strains, although most of those were pan-drug-susceptible. Hence, complimentary use of WGS to traditional epidemiological methods provides an in-depth insight on transmission and drug resistance patterns which can guide targeted control measures to stop the spread of MDR-TB

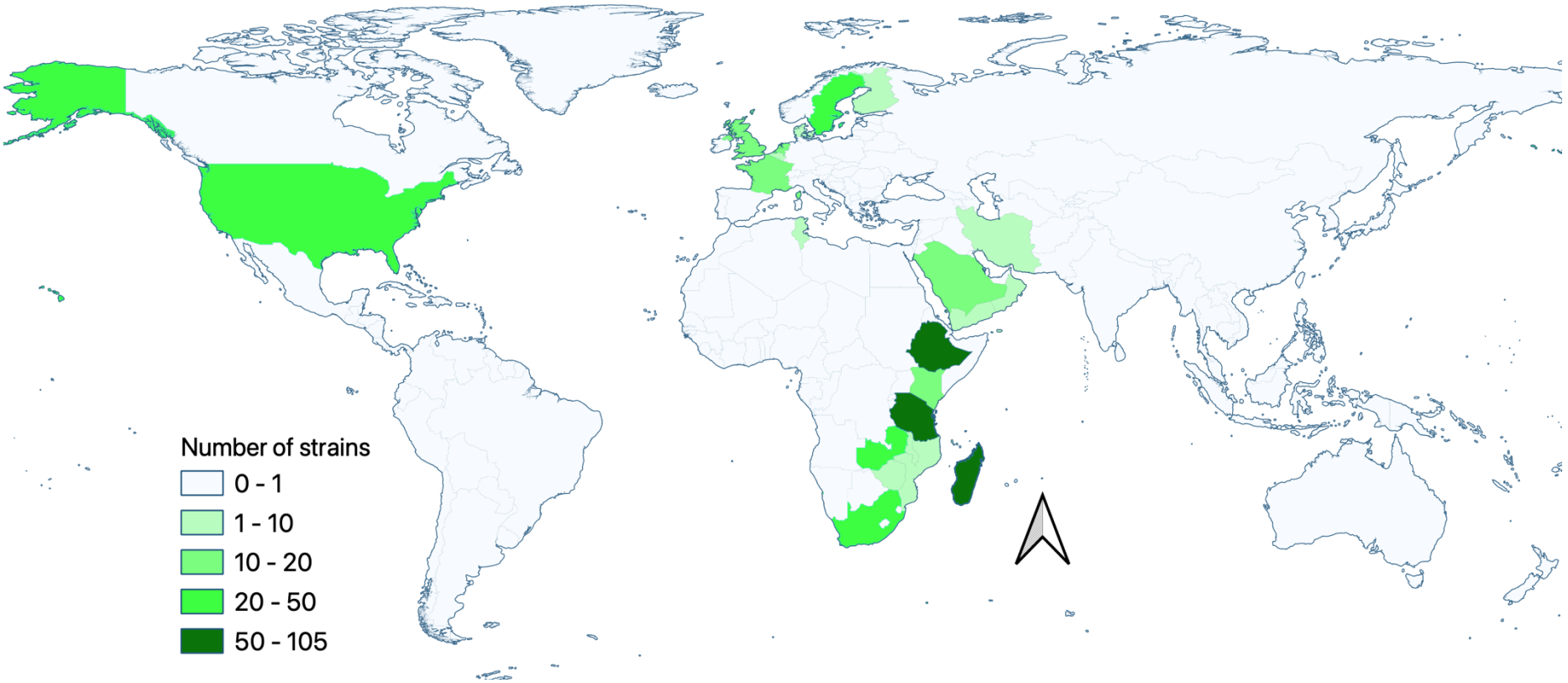


Figure 14. Distribution of SIT21/CAS1-Kili strains by country. Data obtained from SITVIT2 MTBC Database (Demay et al., 2012).

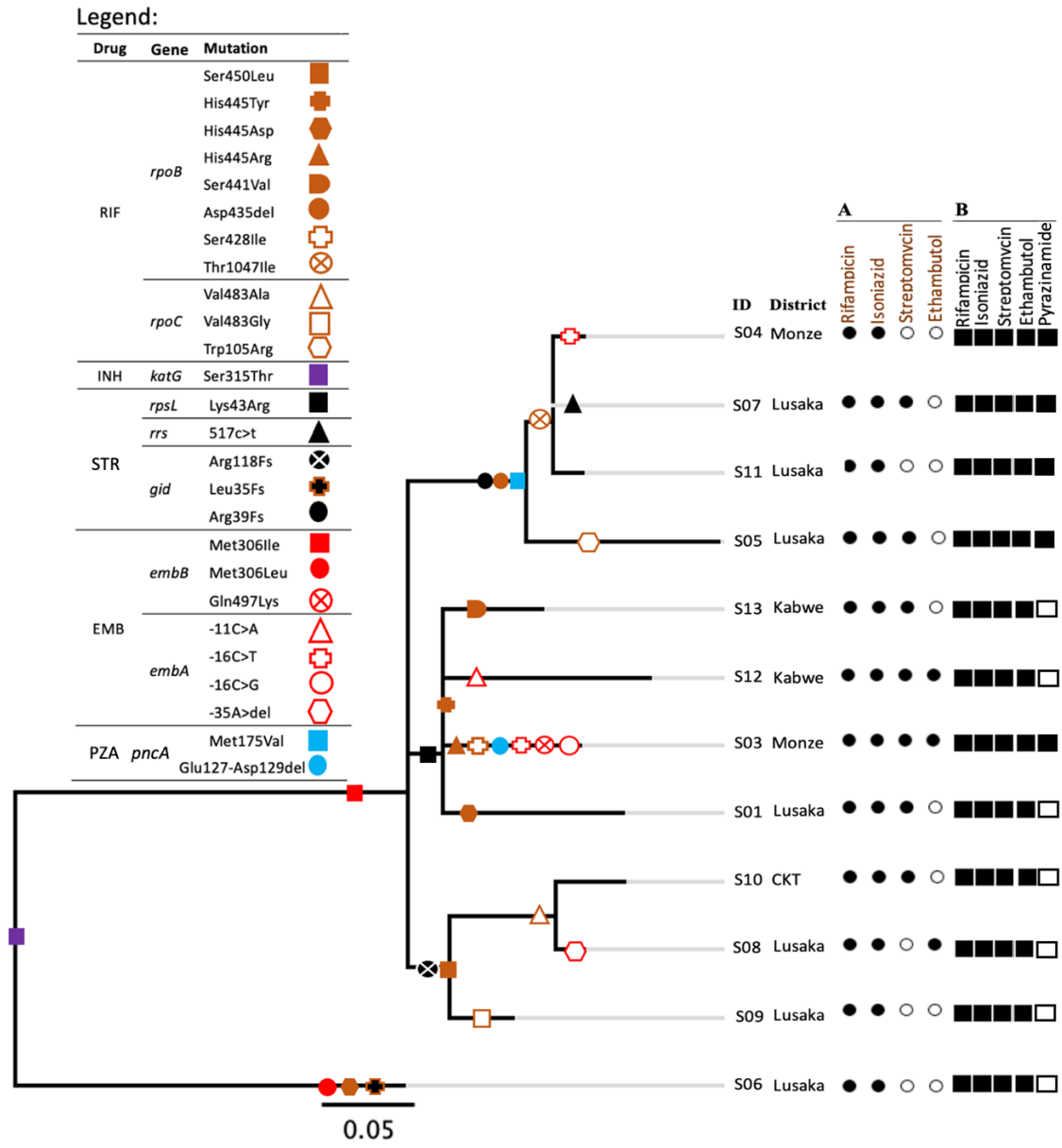
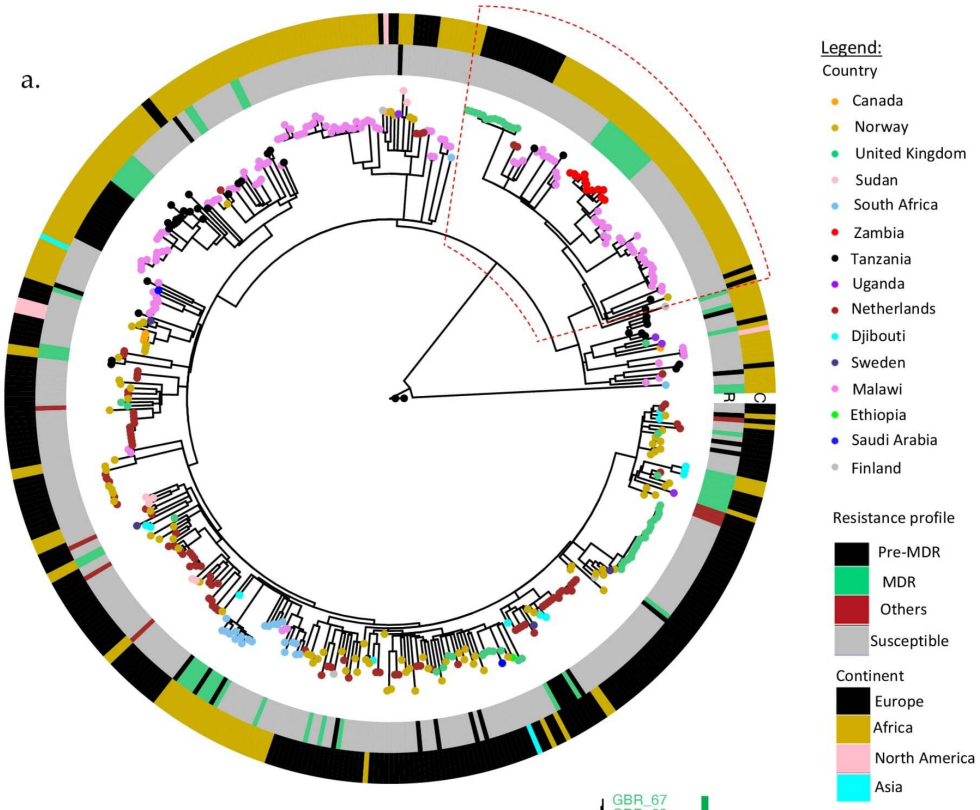


Figure 15. Phylogenetic tree illustrating the relations of the studied SIT21/CAS1-Kili strains. Drug-resistance-associated mutations of each strain are represented by different shapes on the branches and highlighted as indicated by the legend. Fs in the legend stand for frameshift. A, is for phenotypic drug susceptibility patterns. Black and white circles represent results of drug-resistant and susceptible phenotypes, respectively. B, Represents drug-resistant-associated mutations. Black and white squares indicate the presence and absence of drug-resistance-associated mutations to a particular drug, respectively. CTK in the district column stands for a district Chikankata. An 'S' letter before each ID number stands for strain. In the legend, RIF, INH, STR, EMB and PZA stand for rifampicin, isoniazid, streptomycin, ethambutol and pyrazinamide, respectively.

Drug	Gene	Mutation	Nucleotide change	Number of strains	Strain ID											
					S04	S07	S11	S05	S13	S12	S03	S01	S10	S08	S09	S06
Rifampicin	*rpoB*	Ser450Leu	TCG > TTG	4									■	■	■	
		His445Tyr	CAC > TAC	2					■	■						
		His445Asp	CAC > GAC	2								■				■
		His445Arg	CAC > CGC	1							■					
		Ser441Val	TCG > GTG	1					■							
		Ser428Ile	AGC > ATC	1							■					
		Asp435del	GAC > ---	4	■	■	■	■								
		Thr1047Ile	ACC > ATC	3	■	■	■									
	rpoC	Val483Ala	GTG > GCG	2									■	■		
		Val483Gly	GTG > GGG	2											■	
		Trp105Arg	TGG > CGG	1				■								
Isoniazid	*katG*	Ser315Thr	AGC > ACC	13	■	■	■	■	■	■	■	■	■	■	■	■
Streptomycin	*rpsL*	Lys43Arg	AAG > AGG	4					■	■	■	■				
	rrs	517c>t	C > T	1		■										
	gid	Arg118Fs	351del	3									■	■	■	
		Leu35Fs	102del	1												■
		Arg39Fs	115del	4	■	■	■	■								
Ethambutol	*embB*	Met306Ile	ATG > ATA	11	■	■	■	■	■	■	■	■	■	■	■	
		Met306Leu	ATG > TTG	1												■
		Gln497Lys	CAG > AAG	1							■					
	embA	-11C>A	C > A	1						■						
		-35A>del	A > del	1										■		
		-16C>T	C > T	2	■						■					
		-16C>G	C > G	1							■					
Pyrazinamide	*pncA*	Met175Val	ATG > GTG	4	■	■	■	■								
		Glu127-Asp129del	380 - 388del	1							■					

Figure 16. Nucleotide changes in drug resistance associated genes and number of strains having the same nucleotide change among the 13 SIT21/CAS1-Kili strains. The strain IDs are arranged in relation to Figure 15 (Starting from the top). Fs in the figure represents frameshift.

a.


Legend:
Country
Canada
Norway
United Kingdom
Sudan
South Africa
Zambia
Tanzania
Uganda
Netherlands
Djibouti
Sweden
Malawi
Ethiopia
Saudi Arabia
Finland
Resistance profile
Pre-MDR
MDR
Others
Susceptible
Continent
Europe
Africa
North America
Asia

b.

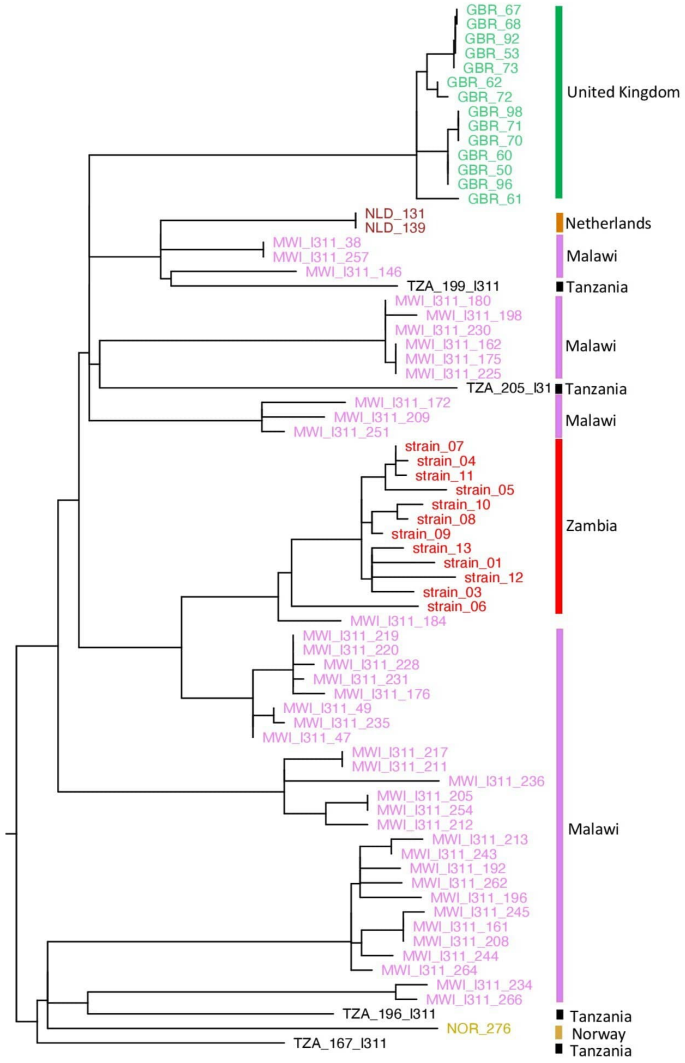

GBR_67
GBR_68
GBR_92
GBR_53
GBR_73
GBR_62
GBR_72
GBR_98
GBR_71
GBR_70
GBR_60
GBR_50
GBR_96
GBR_61
United Kingdom
NLD_131
NLD_139
Netherlands
MWI_I311_38
MWI_I311_257
MWI_I311_146
Malawi
TZA_199_I311
Tanzania
MWI_I311_180
MWI_I311_198
MWI_I311_230
MWI_I311_162
MWI_I311_175
MWI_I311_225
Malawi
TZA_205_I31
Tanzania
MWI_I311_172
MWI_I311_209
MWI_I311_251
Malawi
strain_07
strain_04
strain_11
strain_05
strain_10
strain_08
strain_09
strain_13
strain_01
strain_12
strain_03
strain_06
Zambia
MWI_I311_184
MWI_I311_219
MWI_I311_220
MWI_I311_228
MWI_I311_231
MWI_I311_176
MWI_I311_49
MWI_I311_235
MWI_I311_47
MWI_I311_217
MWI_I311_211
MWI_I311_236
MWI_I311_205
MWI_I311_254
MWI_I311_212
MWI_I311_213
MWI_I311_243
MWI_I311_192
MWI_I311_262
MWI_I311_196
MWI_I311_245
MWI_I311_161
MWI_I311_208
MWI_I311_244
MWI_I311_264
MWI_I311_234
MWI_I311_266
Malawi
TZA_196_I311
Tanzania
NOR_276
Norway
TZA_167_I311
Tanzania

Figure 17. (a). Phylogenetic tree for global L3.1.1 strains. The countries are represented by small circles on the tips of the tree and colored as shown in the legend. The inner (R) and outer circles (C) were colored according to the drug resistance patterns and regions of strain isolation (as continents), respectively. (b). Enlargement of the clade containing the Zambian strains.

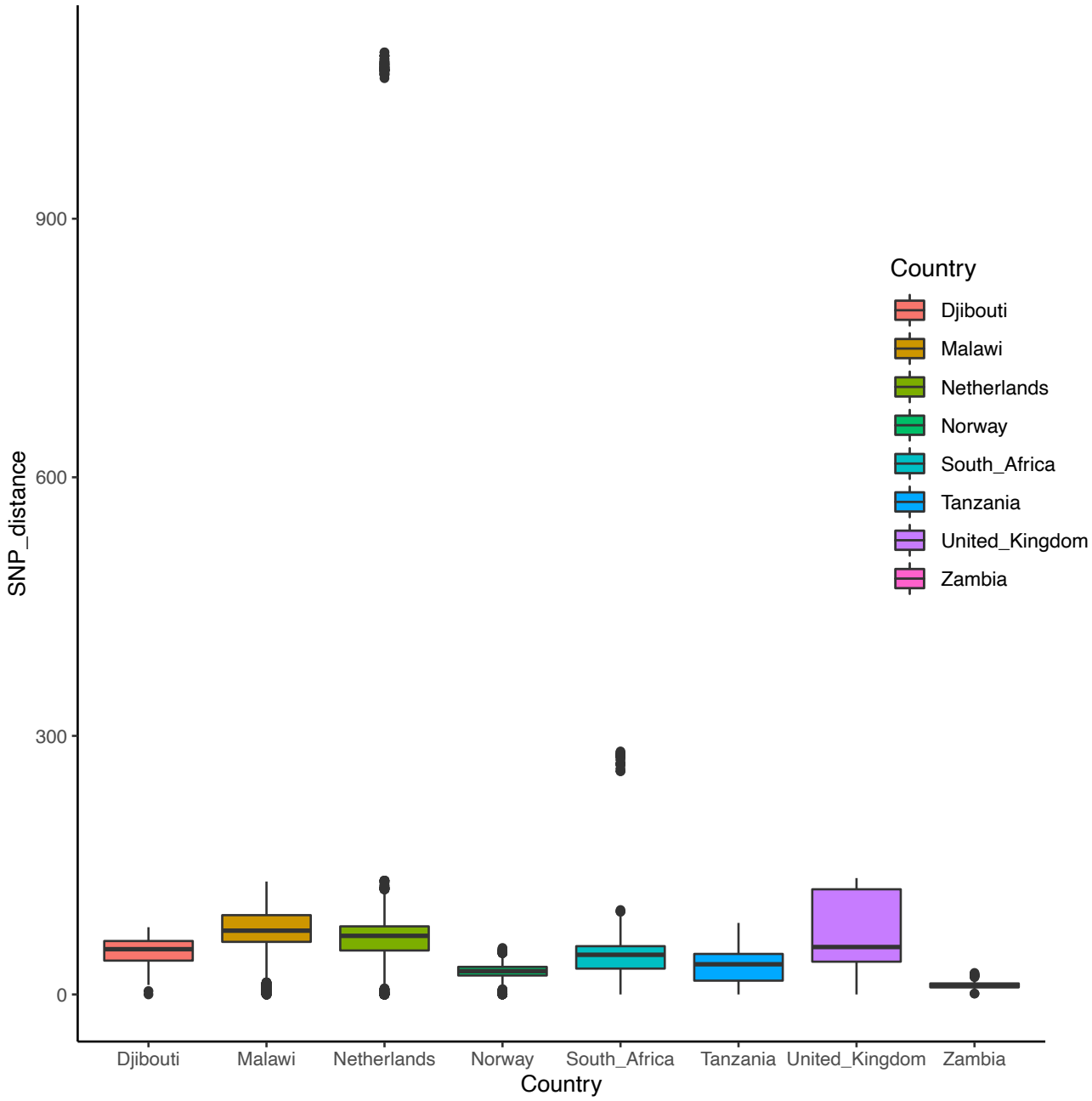


Figure 18. Comparison of median SNP distances of global representative L3.1.1 strains by country to Zambian strains. The differences in median SNP distance were significant ($p <$ 0.000). Only countries with more than 10 strains were considered.

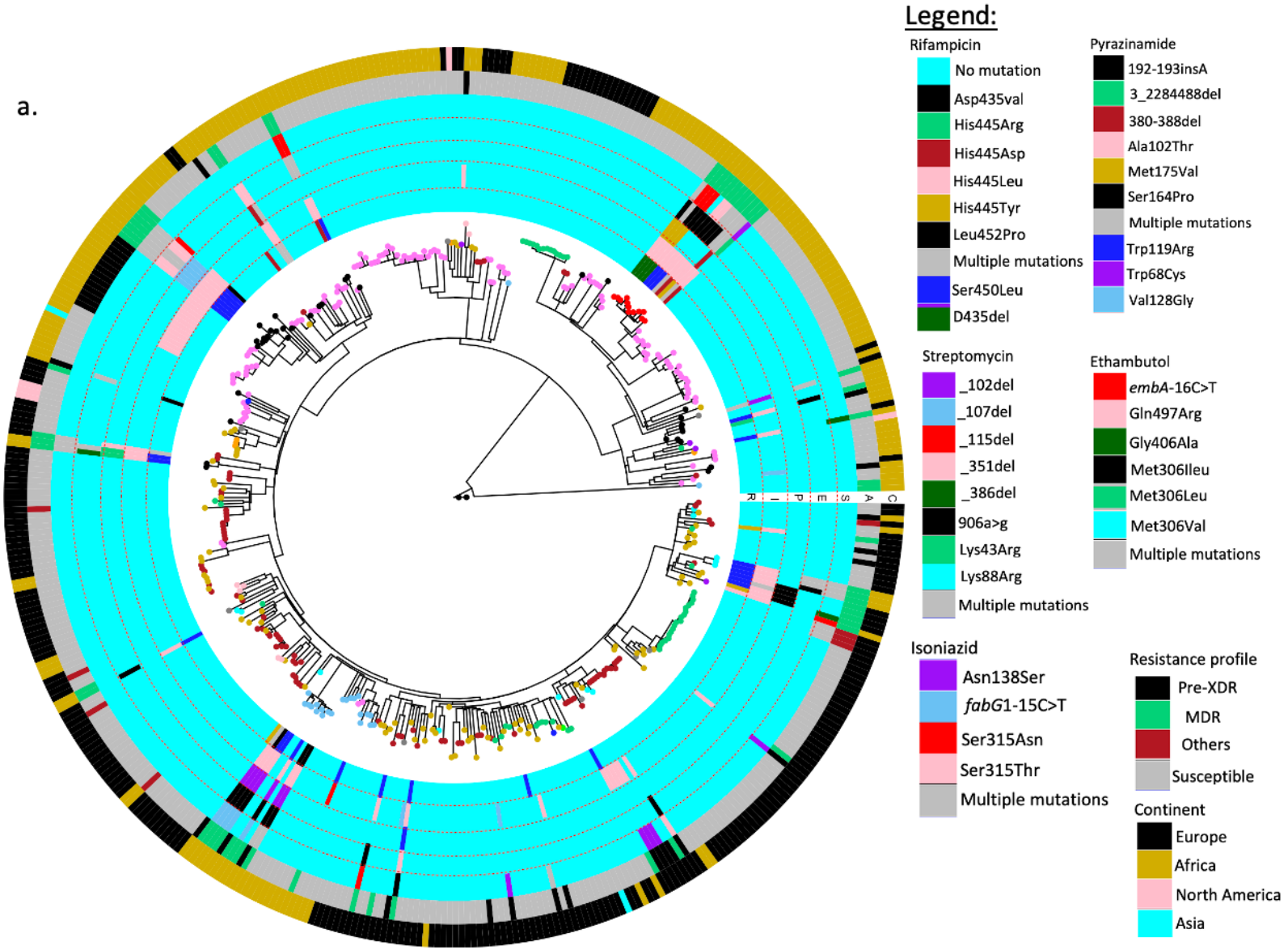


Figure 19. Phylogenetic tree with detailed drug resistant mutations for each global L3.1.1 strain. R = rifampicin, I = isoniazid, P = pyrazinamide, E = ethambutol, S = streptomycin, A = drug resistance pattern, C = continent

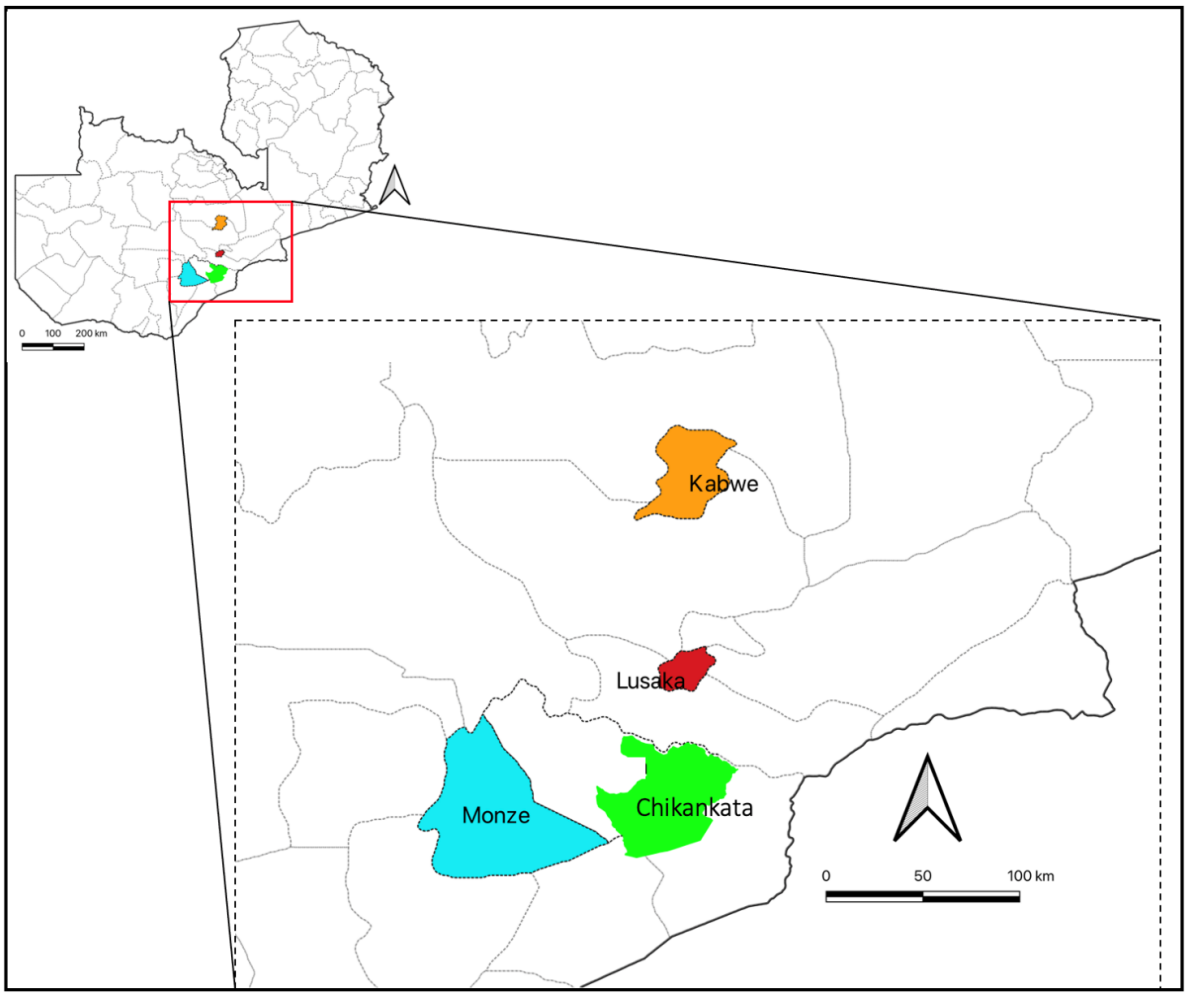


Figure 20. Map showing the residence districts of the 12 patients highlighted in different colors. Seven patients were from Lusaka, and one from Chikankata. Monze and Kabwe had two patients each.

Table 6. Information on the 12 CAS1-Kili Zambian strains.

Sample Number	Mean Coverage	Standard deviation	Mean read length	Total sequences	Number of reads
S1	63.75	± 17.34	249.97	579110	1224370
S3	139.32	± 39.43	209.31	1522426	3195791
S4	35.73	± 21.40	260.14	318802	750634
S5	118.74	± 36.57	178.23	1534056	3292956
S6	46.17	± 15.43	250.57	426310	900458
S7	111.68	± 27.18	167.49	1550554	3334987
S8	43.29	± 13.69	254.62	387347	799903
S9	44.75	16.08	250.83	415133	868824
S10	73.28	± 20.76	278.31	604304	1252391
S11	75.05	± 21.14	266.89	641298	1345606
S12	76.3	± 21.52	271.42	644606	1347590
S13	99.68	± 25.76	212.6	1066338	2159212

Table 7. Similarity matrix showing core-SNP differences of the 12 SIT21/CAS1-Kili strains. SNPs differences between strains ranged from 1 to 26 SNPs.

	Strain ID												SNP Difference[1]		
	S01	S03	S13	S08	S10	S09	S12	S04	S07	S11	S05	S06	**≤5**	**≤12**	**≥13**
S01		9	8	11	12	9	11	11	10	11	14	24	0	9	2
S03	9		7	10	11	8	10	10	9	10	13	23	0	9	2
S13	8	7		9	10	7	9	9	8	9	12	22	0	10	1
S08	11	10	9		3	5	12	10	9	10	13	23	2	9	2
S10	12	11	10	3		6	13	11	10	11	14	24	1	8	3
S09	9	8	7	5	6		10	8	7	8	11	21	1	10	1
S12	11	10	9	12	13	10		12	11	12	15	25	0	8	3
S04	11	10	9	10	11	8	12		1	2	7	23	2	8	1
S07	10	9	8	9	10	7	11	1		1	6	22	2	10	1
S11	11	10	9	10	11	8	12	2	1		7	23	2	10	1
S05	14	13	12	13	14	11	15	7	6	7		26	0	5	6
S06	24	23	22	23	24	21	25	23	22	23	26		0	0	11

Legend: ≤ 5 SNPs; ≤ 12 SNPs; ≥ 13 SNPs. [1] Number of strains differing by ≤5, ≤12, and ≥13 SNPs to an individual strain. The letter S before the ID stands for strain. Boxes are highlighted according to the SNP differences as shown in the legend.

Table 8. Number of L3.1.1 global strains from each country.

Country	Susceptible	Pre-MDR	MDR	Others	Total
Canada	5				5
United Kingdom	48	1	1	1	51
Finland	3		1		4
Norway	69	8	3	6	86
Netherlands	88	5	6		99
Sweden	1		4		5
Saudi Arabia	2				2
Djibouti	5	2	5		12
Uganda	1	1	2		4
Malawi	119	15			134
Tanzania	16	3	17		36
Zambia			12		12
South Africa	12	1	9	1	23
Sudan	6			1	7
Total	**375**	**36**	**60**	**9**	**480**

Table 9. List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler.

accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
SRR6397767	Sensitive	L3.1.1	North_America	Canada	-	-	-	-	-
SRR6152824	Sensitive	L3.1.1	North_America	Canada	-	-	-	-	-
SRR6153047	Sensitive	L3.1.1	North_America	Canada	-	-	-	-	-
SRR6397490	Sensitive	L3.1.1	North_America	Canada	-	-	-	-	-
SRR6397545	Sensitive	L3.1.1	North_America	Canada	-	-	-	-	-
SRR5818587	MDR	L3.1.1	Africa	Djibouti	Ser450Leu	Ser315Thr	-	Met306Ile	-
SRR5818617	MDR	L3.1.1	Africa	Djibouti	Ser450Leu	Ser315Thr	-	multiple	multiple
SRR5818618	Pre-MDR	L3.1.1	Africa	Djibouti	His445Tyr	-	-	-	-
SRR5818620	MDR	L3.1.1	Africa	Djibouti	Ser450Leu	Ser315Thr	-	Met306Ile	multiple
SRR5818636	MDR	L3.1.1	Africa	Djibouti	Ser450Leu	Ser315Thr	-	multiple	multiple
SRR5818637	MDR	L3.1.1	Africa	Djibouti	-	-	-	-	-
SRR5818704	Pre-MDR	L3.1.1	Africa	Djibouti	-	-	-	-	-
SRR5818578	Sensitive	L3.1.1	Africa	Djibouti	-	-	-	-	-
SRR5818585	Sensitive	L3.1.1	Africa	Djibouti	-	-	-	-	-
SRR5818607	Sensitive	L3.1.1	Africa	Djibouti	-	-	-	-	-
SRR5818683	Sensitive	L3.1.1	Africa	Djibouti	-	-	-	-	-
SRR5818689	Sensitive	L3.1.1	Africa	Djibouti	-	-	-	-	-
ERR1200630	Sensitive	L3.1.1	Africa	Ethiopia	-	-	-	-	-
ERR2503423	MDR	L3.1.1	Europe	Finland	-	-	-	-	-
ERR2503494	Sensitive	L3.1.1	Europe	Finland	-	-	-	-	-
ERR2503469	Sensitive	L3.1.1	Europe	Finland	-	-	-	-	-
ERR2503524	Sensitive	L3.1.1	Europe	Finland	-	-	-	-	-
ERR046968	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046969	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046970	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046982	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046983	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR400465	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR2516860	Pre-MDR	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR2516908	MDR	L3.1.1	Europe	United_Kingdom	Ser450Leu	multiple	192_193insA	Met306Val	Lys88Arg
ERR025834	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038264	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038265	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038266	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038269	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038270	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038271	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038272	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038273	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038298	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038299	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR038300	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR039323	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040086	Other	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040087	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040088	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040090	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040091	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040093	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040106	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040107	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR040108	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046729	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046769	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046770	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046771	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046772	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046775	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046845	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046847	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046854	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046857	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046858	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046860	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046861	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046866	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046867	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046914	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046921	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046936	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046940	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR046942	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-

Table 9: List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler continues…..

Accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
ERR046965	Sensitive	L3.1.1	Europe	United_Kingdom	-	-	-	-	-
ERR190347	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR212146	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR245757	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR245777	Pre-MDR	L3.1.1	Africa	Malawi	-	fabG1_c.-1	-	-	-
ERR245792	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR245805	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR036193	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR036210	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR036216	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245832	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR037502	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161040	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245838	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR161064	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161065	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245646	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245668	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245671	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245676	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245684	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245687	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245688	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245689	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245694	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245702	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245704	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245713	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245715	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245726	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245732	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245766	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245768	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245789	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245796	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245803	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245827	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245839	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR245840	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR036218	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161083	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161099	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161107	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161167	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161198	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161199	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161200	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR163939	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR163951	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR163966	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR163973	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR163978	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176455	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176466	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181692	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR176471	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176479	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176481	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176497	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176504	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176528	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176577	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176600	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176601	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176603	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176648	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176657	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176662	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176668	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176670	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176674	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176704	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-

Table 9: List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler continues…..

Accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
ERR176731	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR176791	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181683	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181707	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181713	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181715	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181717	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181718	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181736	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181763	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR181770	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181790	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181800	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181802	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181804	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181812	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181817	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181830	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181842	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181846	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181851	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181854	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181856	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181863	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181883	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181892	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181906	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181953	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181970	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181994	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR182044	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR190341	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR190357	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR190372	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR190404	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR190407	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR190408	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR211997	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212007	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212026	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212027	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212029	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212031	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212142	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR212147	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR216937	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR216981	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR221576	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR221587	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR221591	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR036243	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR037492	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181763	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR161079	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161082	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161169	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR161192	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR163974	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181808	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR164011	Sensitive	L3.1.1	Africa	Malawi	-	-	-	-	-
ERR181815	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR181852	Pre-MDR	L3.1.1	Africa	Malawi	His445Asp	-	-	-	-
ERR212017	Pre-MDR	L3.1.1	Africa	Malawi	-	Ser315Thr	-	-	-
ERR2446176	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446296	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446308	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446364	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446372	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446390	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446430	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446437	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446467	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446511	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-

Table 9: List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler continues…..

Accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
ERR2446524	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446565	Pre-MDR	L3.1.1	Europe	Netherlands	-	Ser315Thr	-	-	-
ERR2446580	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275191	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275214	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275217	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275218	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	_102del
ERR3275229	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275314	Pre-MDR	L3.1.1	Europe	Netherlands	-	Ser315Thr	-	-	-
ERR3275320	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275321	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275402	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275405	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275443	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275456	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275470	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275481	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275520	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275574	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275614	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275616	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275643	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275686	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275704	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275708	Pre-MDR	L3.1.1	Europe	Netherlands	-	Ser315Thr	-	-	-
ERR3275711	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275718	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275737	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275744	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275761	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275764	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275766	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275770	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275822	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275825	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275841	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275842	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275870	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275877	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275904	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275908	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275932	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275943	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3275962	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276012	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276018	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276049	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276059	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276070	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276125	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276143	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276159	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276183	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276197	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276211	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276230	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276270	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR3276291	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446105	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2446116	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	multiple
ERR2446234	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2516466	Pre-MDR	L3.1.1	Europe	Netherlands	Ser450Leu	-	-	-	-
ERR2516567	MDR	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2516594	MDR	L3.1.1	Europe	Netherlands	His445Tyr	Ser315Thr	-	-	Lys88Arg
ERR2516609	MDR	L3.1.1	Europe	Netherlands	Ser450Leu	Ser315Thr	-	-	-
ERR2516634	MDR	L3.1.1	Europe	Netherlands	Ser450Leu	Ser315Thr	192_193insA	Met306Val	-
ERR2517142	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517148	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517152	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	_115del
ERR2517158	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	_102del
ERR2517188	MDR	L3.1.1	Europe	Netherlands	Ser450Leu	Ser315Thr	-	Met306Ile	-
ERR2517235	Pre-MDR	L3.1.1	Europe	Netherlands	-	Ser315Thr	-	-	-
ERR2517240	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-

Table 9: List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler continues…..

Accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
ERR2517249	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517251	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517278	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517331	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517348	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	906a>g
ERR2517360	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517371	MDR	L3.1.1	Europe	Netherlands	Ser450Leu	Ser315Thr	Trp119Arg	Gln497Arg	-
ERR2517404	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517407	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517419	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517441	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517581	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	-
ERR2517597	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	_386del
ERR2517626	Sensitive	L3.1.1	Europe	Netherlands	-	-	-	-	multiple
ERR2245301	MDR	L3.1.1	Europe	Norway	His445Tyr	Ser315Thr	Ser164Pro	Met306Val	Lys88Arg
ERR2245391	MDR	L3.1.1	Europe	Norway	multiple	Ser315Thr	Ala102Thr	Met306Val	Lys43Arg
ERR2245421	MDR	L3.1.1	Europe	Norway	Ser450Leu	Ser315Thr	192_193insA	Met306Val	Lys43Arg
ERR2245337	Other	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245338	Other	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245340	Other	L3.1.1	Europe	Norway	-	-	-	-	Lys43Arg
ERR2245375	Other	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245378	Other	L3.1.1	Europe	Norway	-	-	-	-	_351del
ERR2245396	Other	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245296	Pre-MDR	L3.1.1	Europe	Norway	Ser450Leu	-	-	-	-
ERR2245305	Pre-MDR	L3.1.1	Europe	Norway	-	fabG1_c.-1	-	-	-
ERR2245342	Pre-MDR	L3.1.1	Europe	Norway	-	-	-	-	_102del
ERR2245346	Pre-MDR	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245363	Pre-MDR	L3.1.1	Europe	Norway	-	fabG1_c.-1	-	-	_351del
ERR2245365	Pre-MDR	L3.1.1	Europe	Norway	-	Ser315Thr	-	-	_102del
ERR2245409	Pre-MDR	L3.1.1	Europe	Norway	-	Ser315Thr	-	Met306Ile	-
ERR2245416	Pre-MDR	L3.1.1	Europe	Norway	-	Ser315Thr	-	-	-
ERR2245279	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245284	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245285	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245287	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245288	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245289	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245290	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245292	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245294	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245295	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245297	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245298	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245299	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245300	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245303	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245306	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245333	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245335	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245336	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245341	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245353	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245356	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245364	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245366	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245367	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245372	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245373	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245374	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245376	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245377	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245379	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245380	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245381	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245382	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245383	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245384	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245385	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245386	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245388	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245389	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245390	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245392	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-

Table 9: List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler continues…..

Accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
ERR2245393	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245395	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245397	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245398	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245399	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245401	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245402	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245403	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245406	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245407	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245408	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245410	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245411	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245412	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245413	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245414	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245415	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245417	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245418	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245419	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245420	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245422	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245423	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245424	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2245427	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2512376	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR2512378	Sensitive	L3.1.1	Europe	Norway	-	-	-	-	-
ERR1035167	Sensitive	L3.1.1	Middle_East	Saudi_Arabia	-	-	-	-	-
ERR1035172	Sensitive	L3.1.1	Middle_East	Saudi_Arabia	-	-	-	-	-
ERR3470579	Sensitive	L3.1.1	Africa	Sudan	-	-	-	-	-
ERR3470580	Sensitive	L3.1.1	Africa	Sudan	-	-	-	-	-
ERR3470573	Sensitive	L3.1.1	Africa	Sudan	-	-	-	-	-
ERR3470620	Sensitive	L3.1.1	Africa	Sudan	-	-	-	-	-
ERR3470613	Other	L3.1.1	Africa	Sudan	-	-	-	-	-
ERR3470556	Sensitive	L3.1.1	Africa	Sudan	-	-	-	-	Lys43Arg
ERR3470623	Sensitive	L3.1.1	Africa	Sudan	-	-	-	-	Lys43Arg
strain_01	MDR	L3.1.1	Africa	Zambia	His445Asp	Ser315Thr	-	Met306Ile	_115del
strain_03	MDR	L3.1.1	Africa	Zambia	multiple	Ser315Thr	380_388del	multiple	_115del
strain_04	MDR	L3.1.1	Africa	Zambia	-	Ser315Thr	Met175Val	multiple	_102del
strain_05	MDR	L3.1.1	Africa	Zambia	-	Ser315Thr	Met175Val	Met306Ile	multiple
strain_06	MDR	L3.1.1	Africa	Zambia	His445Asp	Ser315Thr	-	Met306Leu	_351del
strain_07	MDR	L3.1.1	Africa	Zambia	-	Ser315Thr	Met175Val	Met306Ile	_351del
strain_08	MDR	L3.1.1	Africa	Zambia	Ser450Leu	Ser315Thr	-	Met306Ile	-
strain_09	MDR	L3.1.1	Africa	Zambia	Ser450Leu	Ser315Thr	-	Met306Ile	_115del
strain_10	MDR	L3.1.1	Africa	Zambia	Ser450Leu	Ser315Thr	-	Met306Ile	Lys43Arg
strain_11	MDR	L3.1.1	Africa	Zambia	-	Ser315Thr	Met175Val	Met306Ile	Lys43Arg
strain_12	MDR	L3.1.1	Africa	Zambia	His445Tyr	Ser315Thr	-	multiple	-
strain_13	MDR	L3.1.1	Africa	Zambia	multiple	Ser315Thr	-	Met306Ile	-
SRR1162689	Sensitive	L3.1.1	Europe	Sweden	-	-	-	-	-
SRR1144826	MDR	L3.1.1	Europe	Sweden	-	-	-	-	-
SRR1144734	MDR	L3.1.1	Europe	Sweden	-	-	-	-	-
SRR1144803	MDR	L3.1.1	Europe	Sweden	-	-	-	-	multiple
SRR1162878	MDR	L3.1.1	Europe	Sweden	-	-	-	-	-
ERR3077930	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	Val128Gly	Gln497Arg	-
ERR3077934	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	_115del
ERR3077936	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077939	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	-	multiple	-
ERR3077940	Pre-MDR	L3.1.1	Africa	Tanzania	His445Arg	-	-	-	-
ERR3077948	MDR	L3.1.1	Africa	Tanzania	-	-	-	-	_115del
ERR3077949	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077950	MDR	L3.1.1	Africa	Tanzania	His445Asp	Ser315Thr	-	multiple	-
ERR3077952	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	multiple
ERR3077955	MDR	L3.1.1	Africa	Tanzania	multiple	Ser315Thr	multiple	Gln497Arg	-
ERR3077956	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	Val128Gly	multiple	-
ERR3077959	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077962	Pre-MDR	L3.1.1	Africa	Tanzania	Asp435Val	-	-	multiple	multiple
ERR3077963	MDR	L3.1.1	Africa	Tanzania	-	-	-	-	multiple
ERR3077964	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	Val128Gly	multiple	-
ERR3077965	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	Val128Gly	Gln497Arg	multiple
ERR3077970	MDR	L3.1.1	Africa	Tanzania	Leu452Pro	Ser315Thr	-	embA_c.-16C>T	-
ERR3077971	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	3_2284488del	multiple	-
ERR3077972	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077974	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-

Table 9: List of global L3.1.1 (CAS1-Kili) strains and their mutations according to TB-profiler continues…..

Accession No.	Drug resistance	Sublineage	Continent	Country	Drug resistance mutations				
					RIF	INH	PRZ	EMB	STM
ERR3077976	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077979	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077984	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077985	MDR	L3.1.1	Africa	Tanzania	multiple	Asn138Ser	-	multiple	-
ERR3077987	Pre-MDR	L3.1.1	Africa	Tanzania	-	Ser315Thr	-	-	-
ERR3077988	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077991	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077992	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	-	Met306Ile	-
ERR3077994	MDR	L3.1.1	Africa	Tanzania	multiple	Ser315Thr	_388del	Gln497Arg	multiple
ERR3077995	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3077999	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	3_2284488del	Gly406Ala	-
ERR3078001	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	-
ERR3078002	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	multiple
ERR3078012	Sensitive	L3.1.1	Africa	Tanzania	-	-	-	-	multiple
ERR3078016	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	Val128Gly	Gln497Arg	_115del
ERR3078017	MDR	L3.1.1	Africa	Tanzania	Ser450Leu	Ser315Thr	Val128Gly	Gln497Arg	-
SRR1062878	MDR	L3.1.1	Africa	Uganda	His445Leu	Ser315Thr	-	Met306Val	_386del
SRR1062909	Pre-MDR	L3.1.1	Africa	Uganda	-	Ser315Thr	-	-	-
ERR040126	MDR	L3.1.1	Africa	Uganda	Ser450Leu	Ser315Thr	-	Met306Val	-
ERR040133	Sensitive	L3.1.1	Africa	Uganda	-	-	-	-	-
ERR2515846	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515691	MDR	L3.1.1	Africa	South_Africa	Ser450Leu	Ser315Asn	-	-	_107del
SRR833020	Pre-MDR	L3.1.1	Africa	South_Africa	-	-	-	-	_107del
SRR1011477	MDR	L3.1.1	Africa	South_Africa	Asp435Val	Ser315Thr	Trp68Cys	Met306Ile	_107del
SRR1165230	MDR	L3.1.1	Africa	South_Africa	multiple	Ser315Thr	Trp68Cys	Met306Ile	multiple
SRR1165270	MDR	L3.1.1	Africa	South_Africa	Ser450Leu	Ser315Thr	Trp68Cys	Met306Ile	_107del
SRR1165624	Other	L3.1.1	Africa	South_Africa	Asp435Val	Ser315Thr	Trp68Cys	Met306Ile	_107del
SRR1165649	MDR	L3.1.1	Africa	South_Africa	Ser450Leu	Ser315Thr	Trp68Cys	Met306Ile	multiple
SRR1165655	MDR	L3.1.1	Africa	South_Africa	Ser450Leu	Ser315Thr	Trp68Cys	Met306Ile	-
SRR1165657	MDR	L3.1.1	Africa	South_Africa	Asp435Val	Ser315Thr	Trp68Cys	Met306Ile	-
SRR1172744	MDR	L3.1.1	Africa	South_Africa	-	-	-	-	-
SRR1175029	MDR	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR171128	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR171129	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR1034944	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515140	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515200	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515366	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515404	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515564	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515696	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	-
ERR2515736	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	
ERR2515846	Sensitive	L3.1.1	Africa	South_Africa	-	-	-	-	

CONCLUSION

MDR-TB is of great concern globally. It is a threat to TB control in Zambia. As country, Zambia is listed among the 30 countries with high TB, MDR/RR and TB/HIV burden countries in the world. Despite increasing reports of MDR-TB in routine surveillance, and an estimated increase in the number of MDR-TB cases among the new pulmonary TB patients, the information on transmission of MDR Mycobacterium tuberculosis strains is largely unknown. Therefore, this study elucidated the genetic diversity and transmission of MDR *M. tuberculosis* strains in Lusaka, Zambia.

In chapter 1, eighty-five MDR *M. tuberculosis* samples were evaluated using drug-resistance associated gene sequencing, spoligotyping, and 24-loci MIRU-VNTR. Multiplex PCR was applied for RD-Rio sub-lineage identification. The identified clades were LAM (48%), CAS (29%), T (14%), X (6%) and Harlem (2%). Many strains belonged to SIT 21/CAS1-Kili (29%) and SIT59/LAM11-ZWE (16%). Strains belonging to SITs 21/CAS1-Kili and 20/LAM1 formed the largest clonal complexes. Combined spoligotyping and 24 loci-MIRU-VNTR revealed a clustering rate of 63%. Additionally, ninety-five percent of LAM strains belonged to the RD-Rio sub-lineage. The cumulative Hunter Gaston discriminatory indices for 9, 15 and 24 MIRU-VNTR loci sets were 0.959, 0.965 and 0.973 respectively. The high clustering rate revealed by the study was suggestive of MDR-TB spread to be due to recent transmission than independent acquisition of MDR. This was attributed more to the clonal expansion of SIT21/CAS1-Kili, and SIT20/LAM1. Besides, based on the performance of the assessed MIRU-VNTR loci sets, the study proposed the use 9 and 15-loci MIRU-VNTR for epidemiological and phylogenetic studies on MDR-TB strains in this region, respectively.

While in chapter 2, twelve of the MDR CAS1-Kili isolates clustered by the traditional genotyping methods (spoligotyping and MIRU-VNTR) were further investigated for recent transmission using a high-resolution technique, WGS. Among the 12 CAS1-Kili strains, 92% (11/12) belonged to a cluster (≤12 SNPs) while 50% (6/12) were involved in recent transmission events, as they differed by ≤5 SNPs. All the 12 strains had KatG Ser315Thr (isoniazid resistance), EmbB Met306 substitutions (ethambutol resistance) and several kinds of *rpoB* mutations (rifampicin resistance). The study also revealed compensatory mutations including a novel deletion in *embA* regulatory region (−35A > del). Several strains shared the same combinations of drug-resistance-associated mutations indicating transmission of MDR strains. Moreover, the 12 Zambian strains belonged to the same clade as Tanzanian, Malawian and European strains. Therefore, the study emphasized the ongoing transmission of MDR

SIT21/CAS1-Kili strain in Lusaka, which is of public health concern and needs more evaluation. It also revealed the sequential acquisition of drug-resistance associated mutations which has both public health and clinical implications. Besides, the study suggested the origin of the CAS1-Kili strain to be Tanzania, then spread to Malawi before Zambia. This showed the importance of regional coordination in TB control. Therefore, diagnostic methods and local treatment protocols to rapidly identify these strains and isolate patients affected can help terminate the transmission chains.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my supervisors, Professor Suzuki Yasuhiko and Professor Nakajima Chie for their unwavering support and guidance throughout my studies. I am also thankful to Professor Hideaki Higashi and Professor Norikazu Isoda for their opinions, recommendations, and guidance during the program. Many thanks to Professor Mukonka Victor for his timely advice and support he gave me during the course of the program. I would also like to extend my thanks to Yukari Fukushima and Mr. Mwale Kaemba for their technical support. Special thanks to Dr Solo Eddie for grating me permission to work on his samples. I also value the contributions of Assistant Professor Thapa Jeewan and all members of Bioresources. I am grateful to Japan International Cooperation Agency **(**JICA**)** for providing me with a scholarship to pursue this PhD program.

To my wife, Assah Namute, and my boys, Joseph Chizimu and Benjamin Chizimu, words can't express how much I am grateful to you, thank you. Thank you to my parents, brother, sisters and friends for their support and encouragements, which kept me going. My God, your grace has been so sufficient in my life. Who am I without you? Thank you for everything.

REFERENCES

Agonafir M, Lemma E, Wolde-Meskel D, Goshu S, Santhanam A, Girmachew F, et al. Phenotypic and genotypic analysis of multidrug-resistant tuberculosis in Ethiopia. Int J Tuberc Lung Dis 2010;**14**:1259–65.

Alaridah N, Hallbäck ET, Tångrot J, Winqvist N, Sturegård E, Florén-Johansson K, et al. Transmission dynamics study of tuberculosis isolates with whole genome sequencing in southern Sweden. Sci Rep 2019;**9**:1–9.

Allix-Beguec C, Harmsen D, Weniger T, Supply P, Niemann S, Allix-Béguec C, et al. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J Clin Microbiol 2008;**46**:2692–9.

De Almeida IN, Vasconcellos SEG, de Assis Figueredo LJ, Dantas NGT, Augusto CJ, Hadaad JPA, et al. Frequency of the *Mycobacterium tuberculosis* RDRio genotype and its association with multidrug-resistant tuberculosis. BMC Infect Dis 2019;**19**:1–9.

Alyamani EJ, Marcus SA, Ramirez-Busby SM, Hansen C, Rashid J, El-kholy A, et al. Genomic analysis of the emergence of drug-resistant strains of *Mycobacterium tuberculosis* in the Middle East. Sci Rep 2019;**9**:1–12.

Andersson DI. The biological cost of mutational antibiotic resistance: any practical conclusions? Curr Opin Microbiol 2006;**9**:461–5.

Asare P, Otchere ID, Bedeley E, Brites D, Loiseau C, Baddoo NA, et al. Whole Genome Sequencing and Spatial Analysis Identifies Recent Tuberculosis Transmission Hotspots in Ghana. Front Med 2020;**7**:161.

Bakuła Z, Javed H, Pleń M, Jamil N, Tahir Z, Jagielski T. Genetic diversity of multidrug-resistant *Mycobacterium tuberculosis* isolates in Punjab, Pakistan. Infect Genet Evol 2019;**72**:16–24.

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.

Balasubramanian V, Wiegeshaus EH, Taylor BT, Smith DW. Pathogenesis of tuberculosis: pathway to apical localization. Tuber Lung Dis 1994;**75**:168–78.

Batut B, Hiltemann S, Bagnacani A, Baker D, Bhardwaj V, Blank C, et al. Community-Driven Data Analysis Training for Biology. Cell Syst 2018;**6**:752--758.e1.

Bhembe NL, Green E. Molecular epidemiological study of multidrug-resistant tuberculosis isolated from sputum samples in Eastern Cape, South Africa. Infect Genet Evol 2020;**80**:104182.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;**30**:2114–20.

Boshoff HIM, Barry CE. Tuberculosis — metabolism and respiration in the absence of growth. Nat Rev Microbiol 2005 31 2005;**3**:70–80.

Brandis G, Wrande M, Liljas L, Hughes D. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. Mol Microbiol 2012;**85**:142–51.

Brynildsrud OB, Pepperell CS, Suffys P, Grandjean L, Monteserin J, Debech N, et al. Global expansion of *Mycobacterium tuberculosis* lineage 4 shaped by colonial migration and local adaptation. Sci Adv 2018;**4**:1–12.

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 (Basel) 2021;**10**:1169.

Casper C, Singh SP, Rave S, Daley CL, Schecter GS, Riley LW, et al. The transcontinental transmission of tuberculosis: A molecular epidemiological assessment. Am J Public Health 1996;**86**:551–3.

CDC. TB Risk Factors | Basic TB Facts | TB | CDC 2016.

Chihota VN, Müller B, Mlambo CK, Pillay M, Tait M, Streicher EM, et al. Population structure of multi- and extensively drug-resistant *Mycobacterium tuberculosis* strains in South Africa. J Clin Microbiol 2012;**50**:995–1002.

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**:e0200632.

Chizimu JY, Solo ES, Bwalya P, Kapalamula TF, Akapelwa ML, Lungu P, et al. Genetic Diversity and Transmission of Multidrug Resistant *Mycobacterium tuberculosis* strains in Lusaka, Zambia. Int J Infect Dis 2021;**114:** 142–150,.

CLC Genomics Workbench. CLC Genomics Workbench 20.0 (QIAGEN). Workbench 2013:1–70.

Coll F, McNerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill-Cawthorne G, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. Genome Med 2015;**7**:51.

Coll I Cerezo F. Bioinformatic analysis of *Mycobacterium tuberculosis* whole genome data. PhD thesis. London School of Hygiene & Tropical Medicine, London, UK, 2015.

Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nat Genet 2011;**44**:106–10.

Commissioner UNH, Grandi RF. UNHCR chief applauds Zambia's openness to refugees 2019.

Coscolla M. Biological and Epidemiological Consequences of MTBC Diversity. Adv Exp Med Biol 2017;**1019**:95–116.

Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res 2014;**43**:e15.

Daffé M, Draper P. The Envelope Layers of Mycobacteria with Reference to their Pathogenicity. Adv Microb Physiol 1997;**39**:131–203.

David S, Duarte EL, Leite CQF, Ribeiro J-NN, Maio J-NN, Paixão E, et al. Implication of the RDRio *Mycobacterium tuberculosis* sublineage in multidrug resistant tuberculosis in Portugal. Infect Genet Evol 2012;**12**:1362–7.

Demay C, Liens B, Burguière T, Hill V, Couvin D, Millet J, et al. SITVITWEB – A publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. Infect Genet Evol 2012;**12**:755–66.

Disease Profiling Project. DNA Sequencing: Illumina Workflow Part-2 (DNA Tagmentation, Amplification, and Cleanup) 2018.

Easterbrook PJ, Gibson A, Murad S, Lamprecht D, Ives N, Ferguson A, et al. High rates of clustering of strains causing tuberculosis in Harare, Zimbabwe: A molecular epidemiological study. J Clin Microbiol 2004;**42**:4536–44.

Ernst JD. The immunological life cycle of tuberculosis. Nat Rev Immunol 2012;**12**:581–91.

Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;**32**:3047–8.

Feliciano CS, Namburete EI, Rodrigues Plaça J, Peronni K, Dippenaar A, Warren RM, et al. Accuracy of whole genome sequencing versus phenotypic (MGIT) and commercial molecular tests for detection of drug-resistant *Mycobacterium tuberculosis* isolated from patients in Brazil and Mozambique. Tuberculosis 2018;**110**:59–67.

Feuerriegel S, Schleusener V, Beckert P, Kohl TA, Miotto P, Cirillo DM, et al. PhyResSE: a Web Tool Delineating *Mycobacterium tuberculosis* Antibiotic Resistance and Lineage from Whole-Genome Sequencing Data. J Clin Microbiol 2015;**53**:1908–14.

Gagneux S. Strain Variation in the *Mycobacterium tuberculosis* Complex : Its Role in Biology , Epidemiology and Control 2017; **1019**:319.

Gagneux S. Host-pathogen coevolution in human tuberculosis. Philos Trans R Soc London Ser B, Biol Sci 2012;**367**:850–9.

Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, et al. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 2006;**103**:2869–73.

Gillespie S. Tuberculosis: evolution in millennia and minutes. Biochem Soc Trans 2007;**35**:1317–20.

Gutierrez MC, Brisse S, Brosch R, Fabre M, Omaïs B, Marmiesse M, et al. Ancient Origin and Gene Mosaicism of the Progenitor of *Mycobacterium tuberculosis*. PLOS Pathog 2005;**1**:e5.

Hamblion EL, Le Menach A, Anderson LF, Lalor MK, Brown T, Abubakar I, et al. Recent TB transmission, clustering and predictors of large clusters in London, 2010-2012: Results from first 3-years of universal MIRU-VNTR strain typing. Thorax 2016;**71**:749–56.

Hazbón MH, Valle MB del, 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.

Heusden P van, Gladman S, Lose T. *M. tuberculosis* Variant Analysis (Galaxy Training Materials) 2021.

Hunter PR. Reproducibility and indices of discriminatory power of microbial typing methods. J Clin Microbiol 1990;**28**:1903–5.

Hunter PR, Gaston MA, Hunter PR, Gaston MA, Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: An application of Simpson's index of diversity. J Clin Microbiol 1988;**26**:2465–6.

Kaiser G. The Prokaryotic cell- Bacteria. Microbiology. Community College of Baltimore County (Cantonsville): LibreTexts libraries 2019;782.

Kalscheuer R, Palacios A, Anso I, Cifuente J, Anguita J, Jacobs Jr WR, et al. The *Mycobacterium tuberculosis* capsule: a cell structure with key implications in pathogenesis. Biochem J 2019;**476**:1995–2016.

Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 1997;**35**:907–14.

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.

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**:174.

Kochi A, Vareldzis B, Styblo K. Multidrug-resistant tuberculosis and its control. Res Microbiol 1993;**144**:104–10.

Kurbatova E V., Kaminski DA, Erokhin V V., Volchenkov G V., Andreevskaya SN, Chernousova LN, et al. Performance of Cepheid® Xpert MTB/RIF® and TB-Biochip® MDR in two regions of Russia with a high prevalence of drug-resistant tuberculosis. Eur J Clin Microbiol Infect Dis 2013;**6**:735–43.

Lacoma A, García-Sierra N, Prat C, Maldonado J, Ruiz-Manzano J, Haba L, et al. GenoType MTBDRsl for molecular detection of second-line-drug and ethambutol resistance in *Mycobacterium tuberculosis* strains and clinical samples. J Clin Microbiol 2012;**50**:30–6.

Lalor M, Casali N, Walker T, Anderson L, Davidson J, Ratna N, et al. The use of whole-genome sequencing in cluster investigation of an MDR-TB outbreak. Eur Respir J 2018;**51**:1702313.

Lazzarini LCO, Huard RC, Boechat NL, Gomes HM, Oelemann MC, Kurepina N, et al. Discovery of a novel *Mycobacterium tuberculosis* lineage that is a major cause of tuberculosis in Rio de Janeiro, Brazil. J Clin Microbiol 2007;**45**:3891–902.

Lee ASG, Othman SNK, Ho YM, Wong SY. Novel mutations within the embB gene in ethambutol-susceptible clinical isolates of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2004;**48**:4447–9.

Leung ECC, Leung CC, Kam KM, Yew WW, Chang KC, Leung WM, et al. Transmission of multidrug-resistant and extensively drug-resistant tuberculosis in a metropolitan city. Eur Respir J 2013;**41**:901–8.

Li M, Chen R, Lin S, Lu Y, Liu H, Li G, 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**:781.

Li Q jing, Jiao W wei, Yin Q qin, Li Y jia, Li J qiong, Xu F, et al. Positive epistasis of major low-cost drug resistance mutations rpoB531-TTG and katG315-ACC depends on the phylogenetic background of *Mycobacterium tuberculosis* strains. Int J Antimicrob Agents 2017;**49**:757–62.

Luo D, Chen Q, Xiong G, Peng Y, Liu T, Chen X, et al. Prevalence and molecular characterization of multidrug-resistant *M. tuberculosis* in Jiangxi province, China. Sci Rep 2019;**9**:7315.

Luo T, Yang C, Peng Y, Lu L, Sun G, Wu J, et al. Whole-genome sequencing to detect

recent transmission of *Mycobacterium tuberculosis* in settings with a high burden of tuberculosis. Tuberculosis (Edinb) 2014;**94**:434–40.

Malama S, Muma J, Munyeme M, Mbulo G, Muwonge A, Shamputa IC, et al. Isolation and Molecular Characterization of *Mycobacterium tuberculosis* from Humans and Cattle in Namwala District, Zambia. EcoHealth 2014;**11**:564-70.

Maruri F, Sterling TR, Kaiga AW, Blackman A, van der Heijden YF, Mayer C, et al. A systematic review of gyrase mutations associated with fluoroquinolone-resistant *Mycobacterium tuberculosis* and a proposed gyrase numbering system. J Antimicrob Chemother 2012;**67**:819–31.

Mbugi E V., Katale BZ, Siame KK, Keyyu JD, Kendall SL, Dockrell HM, et al. Genetic diversity of *Mycobacterium tuberculosis* isolated from tuberculosis patients in the Serengeti ecosystem in Tanzania. Tuberculosis 2015;**95**:170–8.

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**:e0154571.

Meier A, Sander P, Schaper KJ, Scholz M, Böttger EC. Correlation of molecular resistance mechanisms and phenotypic resistance levels in streptomycin-resistant *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 1996;**40**:2452-54.

Metcalfe JZ, Kim EY, Lin S-YG, Cattamanchi A, Oh P, Flood J, et al. Determinants of Multidrug-Resistant Tuberculosis Clusters, California, USA, 2004–2007. Emerg Infect Dis 2010;**16**:1403-9.

Ministry of Health. Zambia National TB Strategic Plan 2017-2021 2017.

Mokrousov I, Vyazovaya A, Iwamoto T, Skiba Y, Pole I, Zhdanova S, et al. Latin-American-Mediterranean lineage of *Mycobacterium tuberculosis*: Human traces across pathogen's phylogeography. Mol Phylogenet Evol 2016;**99**:133–43.

Moström P, Gordon M, Sola C, Ridell M, Rastogi N. Methods used in the molecular epidemiology of tuberculosis. Clin Microbiol Infect 2002;**8**:694–704.

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.

Namburete EI, Dippenaar A, Conceição EC, Feliciano C, Nascimento MMP do, Peronni KC, et al. Phylogenomic assessment of drug-resistant *Mycobacterium tuberculosis* strains from Beira, Mozambique. Tuberculosis 2020;**121**:101905.

Napier G, Campino S, Merid Y, Abebe M, Woldeamanuel Y, Aseffa A, et al. Robust barcoding and identification of *Mycobacterium tuberculosis* lineages for epidemiological and clinical studies. Genome Med 2020 121 2020;**12**:1–10.

O'Neill MB, Shockey A, Zarley A, Aylward W, Eldholm V, Kitchen A, et al. Lineage specific histories of *Mycobacterium tuberculosis* dispersal in Africa and Eurasia. Mol Ecol 2019;**28**:3241–56.

Oudghiri A, Chaoui I, Elmzibri M. Molecular Epidemiology of Tuberculosis: A Review of Tools and Applications. J Infect Dis Ther 2018;**6**:1–7.

Oudghiri A, Momen G, Aainouss A, Laglaoui A, El Messaoudi MD, El Mzibri M, et al. Genotypic diversity of multi- and pre-extremely drug-resistant *Mycobacterium tuberculosis* isolates from Morocco. PLoS One 2021;**16**:e0253826.

Palomino JC, Martin A. Drug resistance mechanisms in *Mycobacterium tuberculosis*. Antibiotics 2014;**3**:317–40.

Paradis E, Claude J SK. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 2004;**20**:289–90.

Parsons LM, Salfinger M, Clobridge A, Dormandy J, Mirabello L, Polletta VL, et al. Phenotypic and Molecular Characterization of *Mycobacterium tuberculosis* Isolates Resistant to both Isoniazid and Ethambutol. Antimicrob Agents Chemother 2005;**49**:2218.

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.

Perdigão J, Silva C, Diniz J, Pereira C, Machado D, Ramos J, et al. Clonal expansion across the seas as seen through CPLP-TB database: A joint effort in cataloguing *Mycobacterium tuberculosis* genetic diversity in Portuguese-speaking countries. Infect Genet Evol 2019;**72**:44–58.

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**:41.

Qing S, Tong-yang X, Hai-can L, Xiu-qin Z, Zhi-guang L, Ya-nan L, et al. Mutations within embCAB Are Associated with Variable Level of Ethambutol Resistance in *Mycobacterium tuberculosis* Isolates from China. Antimicrob Agents Chemother 2021;**62**:e01279–317.

Ramaswamy S, Musser JM. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber Lung Dis 1998;**79**:3–29.

Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. PLoS Med 2013;**10**:e1001387.

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.

Sagonda T, Mupfumi L, Manzou R, Makamure B, Tshabalala M, Gwanzura L, et al. Prevalence of Extensively Drug Resistant Tuberculosis among Archived Multidrug Resistant Tuberculosis Isolates in Zimbabwe. Tuberc Res Treat 2014;**2014**:1–8.

Said HM, Kock MM, Ismail NA, Mphahlele M, Baba K, Omar S V, et al. Molecular characterization and second-line antituberculosis drug resistance patterns of multidrug-resistant *Mycobacterium tuberculosis* isolates from the northern region of South Africa. J Clin Microbiol 2012;**50**:2857–62.

Scorpio A, Lindholm-Levy P, Heifets L, Gilman R, Siddiqi S, Cynamon M, et al. Characterization of *pncA* mutations in pyrazinamide-resistant *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 1997;**41**:540–3.

Seemann. T. snippy: Rapid haploid variant calling and core genome alignment 2020.

Shah Y, Maharjan B, Thapa J, Poudel A, Diab HM, Pandey BD, et al. High diversity of multidrug-resistant *Mycobacterium tuberculosis* Central Asian Strain isolates in Nepal. Int J Infect Dis 2017;**63**:13–20.

Shen X, Shen G, Wu J, Gui X, Li X, Mei J, et al. Association between *embB* Codon 306

Mutations and Drug Resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother 2007;**51**:2618–20.

Shuaib YA, Khalil EAG, Wieler LH, Schaible UE, Bakheit MA, Mohamed-Noor SE, et al. *Mycobacterium tuberculosis* Complex Lineage 3 as Causative Agent of Pulmonary Tuberculosis, Eastern Sudan. Emerg Infect Dis 2020;**26**:427–36.

Silva C, Perdigão J, Jordão L, Portugal I. Mycobacterial interspersed repetitive unit typing and mutational profile for multidrug-resistant and extensively drug-resistant tuberculosis surveillance in Portugal: a 3-year period overview. Int J Antimicrob Agents 2014;**44**:546–51.

Smith I. *Mycobacterium tuberculosis* Pathogenesis and Molecular Determinants of Virulence. Clin Microbiol Rev 2003;**16**:463–96.

Solo ES, Nakajima C, Kaile T, Bwalya P, Fukushima Y, Chila S, et al. Mutations of rpoB, katG and inhA genes in multidrug-resistant *Mycobacterium tuberculosis* isolates from Zambia. J Glob Antimicrob Resist 2020;**22:**302-7.

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.

Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, et al. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc Natl Acad Sci U S A 1997;**94**:9869–74.

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.

Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, et al. Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit – Variable-Number Tandem Repeat Typing of *Mycobacterium tuberculosis*. J Clin Microbiol 2006;**44**:4498–510.

Suzuki Y, Katsukawa C, Tamaru A, Abe C, Makino M, Mizuguchi Y, et al. Detection of kanamycin-resistant *Mycobacterium tuberculosis* by identifying mutations in the 16S rRNA gene. J Clin Microbiol 1998;**36**:1220–5.

Team Rs. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA 2020.

Telenti A, Philipp WJ, Sreevatsan S, Bernasconi C, Stockbauer KE, Wieles B, et al. The *emb* operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol. Nat Med 1997;**3**:567–70.

UNAIDS. Tuberculosis and Hiv Tb Is the Leading Cause of Death. vol. 54, p. 20. UAIDS: Geneva, Switzerland: 2018.

Viegas SO, MacHado A, Groenheit R, Ghebremichael S, Pennhag A, Gudo PS, et al. Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mozambique. BMC Microbiol 2010;**10**:195.

Walker TM, Ip CLC, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis 2013;**13**:137–46.

Walker TM, Merker M, Knoblauch AM, Helbling P, Schoch OD, van der Werf MJ, et al. A cluster of multidrug-resistant *Mycobacterium tuberculosis* among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. Lancet Infect Dis 2018;**18**:431–

40.

Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D. MIRU-VNTRplus: A web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. vol. 38. 2010.

WHO. Global Tuberculosis Report 2021 2021a.

WHO. Global Tuberculosis Programme 2021b.

WHO. TB country, regional and global profiles 2021c.

WHO. TB profile 2021d.

WHO. WHO consolidated guidelines on tuberculosis Module 4: Treatment Drug-resistant tuberculosis treatment 2020.

WHO. ANNEX 2 Country profiles for 30 high-TB burden countries 2019a.

WHO. Global Tuberculosis Report 2019. 2019b.

WHO. ANNEX 2 Country profiles for 30 high-TB burden countries 2018.

WHO. Annex 2: Country profiles for 30 high-TB burden countries. Glob Tuberc Rep 2017 2017:1–61.

WHO. ANNEX 2 Country profiles for 30 high-TB burden countries 2016.

WHO 2019. Executive summary. 2019.

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.

Xu G, Mao X, Wang J, Pan H. Clustering and recent transmission of *Mycobacterium tuberculosis* in a Chinese population. Infect Drug Resist 2018;**11**:323–30.

Xu Y, Cancino-Muñoz I, Torres-Puente M, Villamayor LM, Borrá R, Borrá s-Má ñez M, et al. High-resolution mapping of tuberculosis transmission: Whole genome sequencing and phylogenetic modelling of a cohort from Valencia Region, Spain. PLoS Med. 2019;**16**:e1002961.

Yu G. Using ggtree to Visualize Data on Tree-Like Structures. Curr Protoc Bioinforma 2020;**69**:e96.

Yu G, Lam TT-Y, Zhu H, Guan Y. Two Methods for Mapping and Visualizing Associated Data on Phylogeny Using Ggtree. Mol Biol Evol 2018;**35**:3041–3.

Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol 2017;**8**:28–36.