2019-03

Characterization of mutations in drug resistant tuberculosis and diagnostic challenges in referral health facilities, Tanzania

Mnyambwa, Nicholaus Peter

NM-AIST

https://doi.org/10.58694/20.500.12479/297

CHARACTERIZATION OF MUTATIONS IN DRUG RESISTANT TUBERCULOSIS AND DIAGNOSTIC CHALLENGES IN REFERRAL HEALTH FACILITIES, TANZANIA

Nicholaus Peter Mnyambwa

A Dissertation Submitted in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Life Sciences of the Nelson Mandela African Institution of Science and Technology

Arusha, Tanzania

March, 2019

ABSTRACT

Tuberculosis remains one of the world's deadliest infectious diseases in resource-limited settings, including Tanzania. Diagnostic challenges and minimal information on resistant tuberculosis complicate building effective management strategies. The current study employed whole genome shotgun sequencing and genotyping methods to characterize genetics of drug resistant tuberculosis strains and diagnostic impedes of tuberculosis in Tanzania. A total of 134 positive sputa from collected at Central Tuberculosis Reference Laboratory from different parts of the country. Forty patients were regarded as multi-drug resistant tuberculosis (MDR-TB), of which 18 (45%) were classified as relapse cases. The remaining 94 were smear-positive culture-negative samples and treated as susceptible tuberculosis. Sequence analysis of 40 MDR-TB isolates identified a set of genetic markers (including additional variants) in the following known drug-resistant genes: *katG*, *inhA*, *embCAB*, *ethA*, *inhA*, *rpoB*, *rpoC*, *rpsL*, *gyrA*, *eis*, and *pncA*. Additionally, there was evidence of positive selection in other three novel genomic regions namely: *ndhC*, *ndhI* and *ndhK*. Sequence analysis also identified one isolate of *M. yongonense*, the first case to be described in Tanzania, suggesting that the patient was misdiagnosed with multi-drug resistant tuberculosis. Out of 94 smear-positive but culture negative sputa, 25 (26.60%) were GeneXpert® mycobacteria TB positive. Repeat-culture identified 11/94 (11.70) as culture positive, of which 5 were Capilia TB-Neo positive and confirmed by GenoType MTBC to be *Mycobacterium tuberculosis*/*Mycobacterium canettii*. The remaining 6 Capilia TB-Neo negative samples were typed by GenoType® CM/AS and identified 3 (3.19%) nontuberculous mycobacteria, 2 Gram positive bacteria, and 1 isolate tested negative, together, making a total of 6/94 (6.38%) confirmed false smear-positives. Overall, 28/94 (29.79%) isolates were confirmed TB cases while 60 (63.83%) remained unconfirmed tuberculosis cases. These findings on misdiagnosis and the suggestive of novel resistance-associated mutations in resistant tuberculosis emphasize the need for accurate molecular diagnostic tests for delineating the tuberculosis cases and their drug susceptibility profiles in clinical settings.

Keywords: Tuberculosis; *Mycobacterium tuberculosis* complex (MTBC); mutation; multi-drug resistant tuberculosis (MDR-TB); Tuberculosis diagnosis.

DECLARATION

I, Nicholaus Peter Mnyambwa do hereby declare to the Senate of the Nelson Mandela African Institution of Science and Technology that this dissertation is my own original work and that it has neither been submitted nor being concurrently submitted for degree award elsewhere.

10/03/2019

Nicholaus Peter Mnyambwa **Date**

Name and signature of a candidate

The above declaration is confirmed

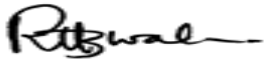
10/03/2019

Prof. Dong-Jin Kim **Date**

Name and signature of supervisor

Prof. Dr. Sayoki G Mfinanga 10/03/2019

Name and signature of supervisor **Date**



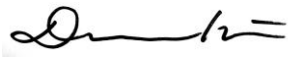
Prof. Rudovick Kazwala 10/03/2019

Name and signature of supervisor **Date**

COPYRIGHT

CERTIFICATION

The undersigned certifies that they have read and hereby recommend for final submission in an acceptable form a dissertation entitled **"Characterization of mutations in resistant drug tuberculosis and diagnostic challenges in referral health facilities, Tanzania"** in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Life Sciences (LiSe) of the Nelson Mandela African Institution of Science and Technology



10/03/2019

Prof. Dong-Jin Kim **Date**

Name and signature of supervisor

Prof. Dr. Sayoki G Mfinanga 10/03/2019

Name and signature of supervisor **Date**

Prof. Rudovick Kazwala 10/03/2019

Name and signature of supervisor **Date**

ACKNOWLEDGEMENT

The successful completion of this dissertation would be impossible without the assistance and guidance of several scholars in the thematic area. I would like to convey my gratefulness to every scholar who has contributed in either way to the success of this dissertation, most notably Prof. Dong-Jin Kim (NM-AIST), Prof. Dr. Sayoki Mfinanga (NIMR-Muhimbili Medical Research Centre) and Prof. Rudovick Kazwala (Sokoine University of Agriculture), for their endless support and invaluable critique kept me focused along with my thinking and writing this thesis. I am highly indebted to Prof. Pammla Petrucka (University of Saskatchewan, Saskatoon, Canada), Dr. Esther Ngadaya (NIMR, Muhimbili Medical Research Centre) and Prof. Addo Kwasi (University of Ghana) for their cooperation, assistance and comments that that kept me going. I thank Dr. Gonzalo Garcia (Earlham Institute, United Kingdom), Dr. Jean-Baka Domelevo and Dr. Dedan Githae (BeCA-ILRI) for their mentorship during sequence analysis. I would like to extend my sincere thanks to Prof Midori Kato-Maeda (University of California, San Francisco for technical advice on sequencing. I am immensely grateful to the Tanzania Government through the NM-AIST, Afrique-ONE Aspire and the National Institute for Medical Research (NIMR), Muhimbili Centre for funding my studies and research work.

Lastly, but not least, I convey my sincere heartfelt appreciations to my parents, siblings and friends for their love, assistance, courage and support in diverse ways to see me through this academic pursuit.

DEDICATION

This dissertation is lovingly dedicated to my beloved parents, Mr and Mrs Peter Mnyambwa and all family members for they have always honoured me with endless love, prayers, support and courage that sustained me throughout this academic pursuit.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AFB : Acid-Fast Bacilli

ART : Antiretroviral

AMK : Amikacin

AMK : Amikacin

CAP : Capreomycin

CIP : Ciprofloxacin

CTAB : Cetyltrimethyl Ammonium Bromide

CTRL : Central TB Reference Laboratory

CYS : Cycloserine.

DNA : Deoxyribonucleic acid

DM : Diabetic mellitus

DOT : Direct Observed Treatment

DST : Drug Susceptibility Testing

EAPHLN : East Africa Public Health Laboratory Network

EMB : Ethambutol

ETH : Ethionamide

EPTB : Extra-pulmonary tuberculosis

FQ : Fluoroquinolone

HIV : Human Immunodeficiency Virus

IGRs : Intergenic regions

INH : Isoniazid

KAN : Kanamycin

KCRI : Kilimanjaro Clinical Research Institute

KIDH : Kibong'oto Infectious Diseases Hospital

LEVO : Levofloxacin

LMIC : Low and Middle Income Countries

LSPs : Large sequence polymorphisms

MAC : *Mycobaterium avium* complex

MDR : Multi-drug resistant

MLST : Multilocus sequence typing

MTBC : *Mycobacterium tuberculosis* complex

NIMR : National Institute for Medical Research

NM-AIST : Nelson Mandela African Institution of Science and Technology

NTM : Nontuberculous mycobacteria

NTLP : National Tuberculosis and Leprosy Programme

OFX : Ofloxacin

PCR : Polymerase Chain Reaction

PZA : Pyrazinamide

RIF : Rifampicin

RR	:	Rifampicin-resistant
SM	:	Streptomycin
SNPs	:	Single nucleotide polymorphisms
TB	:	Tuberculosis
UV	:	Ultraviolet
VIM	:	Viomycin
WHO	:	World Health Organization
XDR	:	Extensively drug resistance

CHAPTER ONE

General introduction

1.1 Background information

This dissertation is comprised of seven chapters. Chapter one provides the background information on tuberculosis (TB), rationale, research questions, research objectives, and the significance of the present study. The next five chapters are paper-based chapters while Chapter 7 summarizes the findings of this dissertation.

Tuberculosis is a chronic contagious infectious disease pathologically characterized by the formation of granulomas (Silva *et al*., 2012), caused by different species and sub-species of tubercle bacillus that are collectively known as *Mycobacterium tuberculosis* complex (MTBC) (Köser *et al*., 2012; Zink *et al*., 2003). These include human adapted species (*M. tuberculosis, M. africanum* and *M. canetti)* and animal adapted species *M. bovis* (bovine), and *M. microti* (rodents), which displays similar clinical features in human (Homolka *et al*., 2010; Riojas *et al*., 2018). Others include *M. caprae* (goats), *M. pinnipedii* (seals) (Riojas *et al*., 2018), and *M. mungi* that infect banded mongooses (Alexander *et al*., 2010). The most widespread causative of human TB is *M. tuberculosis*. Transmission of the disease to another person occurs either through inhalation of droplet nuclei (airborne particles about 1-5 microns) which are aerosolized by person with active TB when speaking, laughing, coughing, singing or sneezing (Knechel, 2009) or it can be by drinking or eating raw animal products such as milk and meat from infected animal. The droplets can remain in airborne for minutes to hours after expectoration, and transmissibility of the pathogen depends on its concentration and virulence, exposure to UV light, occasions and the degree of ventilation (Klein and Yang, 2014; Knechel, 2009). The viability and infectivity of the pathogen can be up to months depending on conditions such as exposure to UV light (Klein and Yang, 2014).

Tuberculosis can be classified into various forms based on four factors which are: site of the disease, bacteriologic results (including drug resistance profile), history of previous TB treatment, and HIV status. While TB can occur in almost any part of the body, the site of TB disease is generally classified as either pulmonary TB or extra-pulmonary TB. Pulmonary TB is the most common form of TB that infect the lung; the disease can spread and affect other organs such as lymph nodes bones and meninges, which is generally classified as extra-

pulmonary TB (EPTB) (Behr and Waters, 2014). Extra-pulmonary TB which forms about 10-15 percent of all reported TB cases globally, cannot be transmitted from person to person, thus, drive the pathogen into an evolutionary dead-end (Behr and Waters, 2014). A patient with both pulmonary TB and EPTB constitutes a case of pulmonary TB. It should be noted that defining the site of TB disease is primarily for epidemiologic purposes and for identifying infectious patients rather than for determining treatment regimen or approach to management. Bacteriologic result includes Acid-Fast Bacilli (AFB) smear, culture and other methods for detection of MTBC. Bacteriologic monitoring for treatment response is most practical in smear-positive patients. A classification of TB as new or previously treated TB cases is also important for identification of patients at increased risk of acquired drug resistance and epidemiological monitoring of the TB epidemic and program performance at regional and country levels. All previously treated cases should be further classified by outcome of most recent course of treatment (TB relapse, treatment failure or default). Lastly, determining and recording HIV status for TB cases is critical making treatment decisions, monitoring trends and for assessment of the program performance.

1.2 Clinical manifestations and progression

Some specific clinical symptoms of TB depend on the specific site of infection. The most clinical presentation of pulmonary TB is cough (dry or productive) that lasts longer than two weeks and one may experience chest pain and coughing up blood or phlegm. Non-specific symptoms for all forms of TB include tiredness or fatigue, lack of appetite, weight loss, chills, fever, and night sweats (Knechel, 2009). In the first two to six weeks after primary infection, the cell–mediated immunity develops (Dannenberg and Rook, 1994). Granulomas and caseation are formed as a results of the immunological reactions at the site of infection (Silva *et al*., 2012). The infection can heal spontaneously, or more often, and be contained by the immune system in a latent stage (latent TB) which may last for years. Reactivation of latent TB can occur later in life if the immune system is compromised or failures to stop the pathogen from growing. Studies have shown that an estimate 10 percent of infected persons with normal immunity will develop TB at some point in life, of which 5 percent occurs in the first 1–2 years post-infection and another 5 percent later in life. Studies conducted during pre-chemotherapy era on untreated smear-positive estimated 50 percent case fatality, 30 percent self-cured and 20 percent remained with chronic disease (Grzybowski, 1991). TB reactivation and progression to active TB can be aggravated by a number of conditions including HIV infections, cancer, malnutrition, alcoholism and transplantation. Furthermore, risk varies

across populations, racial and ethnic and it can also be influenced by lifestyle (Knechel, 2009). Individuals with latent TB develops no signs and/or symptoms of the disease and cannot transmit the disease.

1.3 Global burden of TB

Tuberculosis is a global public health agenda with an immense burden in Low and Middle-Income Countries (LMICs) (WHO, 2016). Despite the availability of anti-TB antibiotics and BCG vaccine, TB continues to among the leading cause of illness and death, especially in the developing countries where co-infections with HIV constitutes a significant risk (Comas *et al*., 2013; WHO, 2016). The global emergence and spread of multi-drug-resistant tuberculosis (MDR-TB) which are at increasing pace, making diagnosis, treatment and control of the disease more difficult (Mnyambwa *et al*., 2017a; WHO, 2015). Details of drug-resistant TB are provided in the next sections.

Despite concerted global efforts aiming to eliminate TB, the decline has been far less than predicted by epidemiological modeling. The current overriding goal of the WHO's End TB Strategy as endorsed in 2014, is to reduce TB incidence rate and deaths due to TB by 90% (<10/100 000) and 95% (<5/100 000), respectively, and to ensure that no family is burdened with catastrophic costs due to TB by 2035 (Uplekar *et al*., 2015; WHO, 2014). Concurrently, the Sustainable Development Goals target reduction in TB incidence and death rates due to TB by 80% (<20/100 000) and 90% (<10/100 000), respectively, in 2030 compared with 2015 percentage. The strategy is comprised of three pillars: (i) Integrated patient-centered care and prevention (ii) Bold policies and supportive health systems and (iii) Intensified research and innovation. The strategy comes to life in the era of MDR-TB and extensively drug-resistant TB (XDR-TB); high relapses rates; and the double epidemic of HIV and non-infectious diseases particularly diabetes mellitus (DM), which altogether, presents major programmatic, diagnostic and management challenges of the disease, especially in sub-Saharan Africa, including Tanzania. Furthermore, strain epidemiology and transmission dynamics of TB remain partially characterized, thereby hampering eradication efforts. The emergence and global spread of MDR-TB have become a priority public health issue: at least 480 000 cases of MDR-TB are globally reported each year, and these figures are argued to be underestimated (WHO, 2016). Nearly 1 in 10 of the MDR-TB cases are categorized XDR-TB: most recently, totally drug-resistant TB has been notified of the XDR-TB cases in various part of the world (WHO, 2016). Multi-drug resistance TB do not respond to the

conventional standard treatment regimen with first-line anti-TB drugs for 6-9 months, instead requires long-term treatment regimens with toxic second-line anti-TB drugs and is associated with non-adherence and high mortality.

Diabetic mellitus is one of the most common chronic diseases in nearly all countries (Shaw *et al*., 2010; WHO, 2016). In 2014, a total of 422 million adults were living with DM, of which three-quarters were living in LMIC (WHO, 2016). Epidemiologic models anticipate that between 2010 and 2030, there will be a 69% increase in numbers of adults with DM in developing countries and 20% increase in developed countries (Shaw *et al*., 2010). Urbanization and lifestyle changes have also led to an emerging epidemic of type 2 DM (WHO, 2016). DM is one of the important modifiable risk factors for TB acquisition that often have synergistic and detrimental health consequences among TB patients, similar to that of HIV infections (Harries *et al*., 2015), while stress response to infection due to TB leads hyperglycemia which is reported to disappears after TB treatment (Boillat-Blanco *et al*., 2016). DM is present in about 15% of TB cases worldwide (Lönnroth *et al*., 2014). The steadily growing epidemic of DM cases will translate into higher TB incidence rates particularly in TB high burden countries. Patients with TB-DM have higher mycobacterial burden than non-diabetics; have increased risk of TB relapse; require a longer time to culture conversion; have higher risk of treatment failure and death (Baker *et al*., 2011; Dooley *et al*., 2009; Harries *et al*., 2011). Life style may have conflicting implication during treatment of TB-DM that require specific clinical guidance (Harries *et al*., 2015). Diabetic person with poor glycaemic control have even worse outcome (Lo *et al*., 2016), thus, enhanced case management of DM reduces risks and improves treatment outcome of TB. Integrated interventions such as systematic bidirectional screening in TB high-burden countries and co-management TB-DM as for TB-HIV are highly recommended. However, some critical questions remain unanswered: (i) when and how often to screen, (ii) what are the most appropriate tools for screening (iii) diagnostic implication of hyperglycaemia in TB patients and (iv) how best to provide dual case management of the diseases are critical research questions need to be answerd (Harries *et al*., 2015).

Designing new effective country-specific interventions is central to the WHO's "End-TB Strategy", in addition to the development of improved affordable vaccines, diagnostics, and drugs (Uplekar *et al*., 2015; WHO, 2016). Developing new effective TB vaccines or improving the available BCG that fails to effectively protect adolescents and adults would

significantly reduce infection and deaths due to TB. The same groups adolescent and adults at aged 15-49 are more likely to be exposed to DM risk factors which include: unhealthy diet, smoking, excessive alcohol consumption and physical inactivity. However, all these require prioritizing TB research agenda and collaboration for effective and evidence-based control interventions. This review reflects current TB situation in Tanzania and highlights possible opportunities for enhancing TB control programs in the country in response to the global strategy to end TB epidemics by 2035.

1.4 Diagnosis

Effective TB control measures require early identification of the TB infected persons followed by timely tailored therapies. Diagnosis of active TB in LMIC is mainly based on the detection of Mycobacteria as achieved by sputum smear examination, mycobacterial culture (McNerney and Daley, 2011) and more recently the GeneXpert MTB/RIF tests (Trébucq *et al*., 2011). However, most TB patients do not produce sputum after the second month of treatment due to the cessation or dramatically cough reduction, making expectorating difficult. EPTB, HIV co-infected individuals, and children, often do not present with positive sputum (McNerney and Daley, 2011). Smear microscopy has limited sensitivity and does not distinguish between live and dead organisms; non-tuberculosis mycobacteria and MTBC (McNerney and Daley, 2011). However, dual TB/DM present with high bacillary density therefore smear microscopy can easily detect. The sensitivity of GeneXpert is also reported low in settings where rifampicin resistance is rare (Weyer *et al*., 2013). The culture method is more sensitive and highly specific than microscopy and GeneXpert and hence serves as the gold standard (Agrawal *et al*., 2016). Unfortunately, most routine laboratories do not perform culture due to lack of Biosafety set-up as MTBC poses biohazards risk and the slow generation time (3-8 weeks) in the Lowenstein Jensen media or 2-6 on liquid media (Ani *et al*., 2009). Test for TB infection can be performed by the use tuberculin skin test or the TB blood test, however, the tests do not tell whether the person has latent TB or has progressed to active TB. Also false positive TB skin test is common in people previously received vaccination with BCG. Thus, additional tests are required to discriminate the infection. Chest radiography is used to detect chest lesions (abnormalities) that may develop infections, but cannot be used to definitively diagnose TB. Evolving novel and/or unfamiliar mutations are revolutionizing the pathways of antibiotic resistance of clinical tuberculosis, further challenging both diagnostic tools and treatment (Mnyambwa *et al*., 2017a). The WHO defined a definite case of TB as a patient with MTBC identified from a clinical specimen by

culture or newer method (e.g., nucleic acid amplification). However, TB case definitions are based on the level of certainty of the diagnosis, and whether or not laboratory confirmation is available.

In Tanzania, the diagnosis of TB is decentralized to TB diagnostic and treatment units, which are integrated in primary healthcare facilities, with 100% DOTS coverage throughout the country. Like in many other LMICs, diagnosis of TB in Tanzania remains a grim. Diagnosis of active TB is achieved by detection of bacilli through smear examination, mycobacterial culture (McNerney and Daley, 2011) and more recently the GeneXpert (Trébucq *et al*., 2011) as well as molecular MTBDRsl tests for DST (WHO, 2016). Smear microscopy is done in all diagnostic units while specimens for culture are sent to zonal laboratories. The zonal laboratories send culture positive isolates to CTRL for drug sensitivity testing. The CTRL monitors the quality of routine services of smear microscopy at peripheral levels. The EPTB is normally diagnosed on basis of clinical findings following guidelines. However, it is well known smear microscopy suffers from low sensitivity and does not distinguish between live vs nonviable organisms; NTM vs MTBC (McNerney and Daley, 2011). In recent years, there has been an increasing trends of NTM in Tanzania, which are easily misdiagnosed as TB (Hoza *et al*., 2016a; Hoza *et al*., 2016b; Kilale *et al*., 2016; Mfinanga *et al*., 2014; Mnyambwa *et al*., 2017b) and sometime as MDR-TB (Shahraki *et al*., 2015) which may result into unnecessary rendering of toxic second-line anti-TB drugs. Currently, mycobacteria culture activities are mainly centralized at CTRL and decentralization of laboratory for culture and DST with improving peripheral specimen collection is underway with the aim to absorb the net increases in case detection.

1.5 Species identification

Advances in genetic techniques have facilitated the better understanding of many epidemiological and bio-physiological attributes of the pathogen. The recent introduction of genotyping methods like IS6110 Restriction Fragment Length Polymorphism (Park *at al*., 2000), Spacer Oligonucleotide Typing (Spoligotyping) and mycobacterial interspersed repetitive units (Sola *et al*., 2003; Supply *et al*., 2001) has been useful in understanding molecular epidemiology of mycobacteria as they generate new insight regarding transmission pathways, relapse, and re-infection of TB (Alland *et al*., 1994; Small *et al*., 1994a). The general principle is that patients infected with clustered strains are epidemiologically linked

(Alland *et al*., 1994; Kamerbeek *et al*., 1997; Small *et al*., 1994a). Such data help to understand the modes of transmission and to put in place an adapted control strategy.

The penetration of high-throughput sequencing has fundamentally changed how scientists study attribute of these pathogens (Mnyambwa *et al*., 2017a). High-throughput whole-genome sequencing is an invaluable tool for detection of mutations associated with drug resistance and phylogenetic analysis (Farhat *et al*., 2013; Zhang *et al*., 2013). With the use of sequencing technologies, discernment of bacterial heterogeneity in an individual patient and the prediction of the pathways to drug-resistant mutations becomes plausible (Ford *et al*., 2012). However, sequencing is very sophisticated technologies that require careful planning and experimental designing. A decision on biological/technical replicas and DNA isolation method, choice of sequencing platform and sequence analysis procedures needs carefully planning because inappropriate or poor design might influence and affect the whole downstream analysis and may lead to a wrong biological conclusion. Procedure for a good sequencing experimental design summarized in Fig 1.

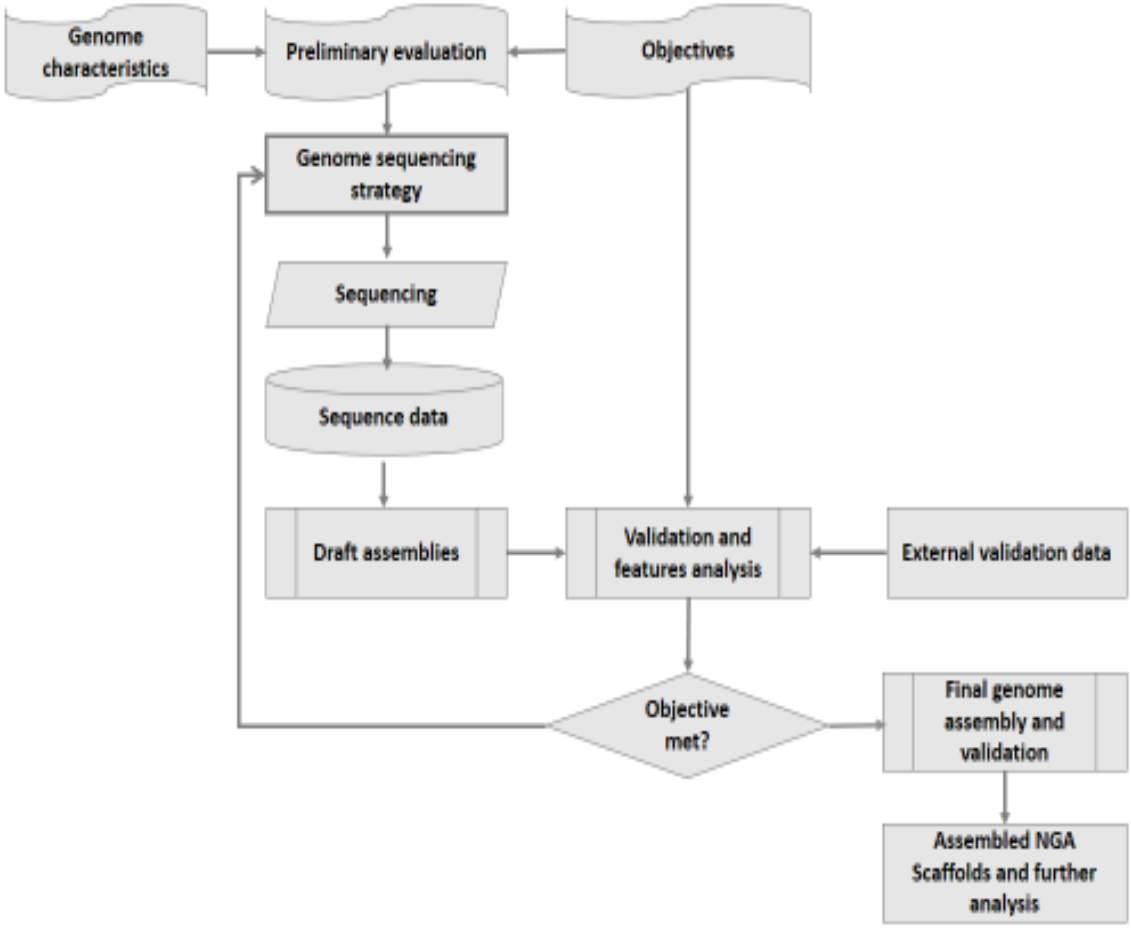


Figure 1: Conceptual framework illustrating key steps in performing sequencing with quality control checks at each stage

Despite these advancements, our understanding of the spectrum and nature of genetic alterations that characterized drug resistance in the MTBC is rudimentary and fragmented (Farhat *et al*., 2013). Further efforts to characterize novel and/or additional causative mutations are required to fully explicate the evolution of drug-resistant TB and its prevalence, particularly in countries with low advancements in technology such as Tanzania.

1.6 Biology of the pathogen

Mycobacterium tuberculosis is a gram-positive and rod-shaped non-spore forming, an aerobic and an obligate human pathogen (Potdar and Thakur, 2013). The characteristic features of the tubercle bacillus include dormancy, slow growth, complex cell wall, genetic homogeneity, and intracellular pathogen with a generation time of 24 hours (Cole *et al*., 1998; Sanz *et al*., 2011). Mycobacteria are categorized as AFB with a complex cell wall that have a thick lipid-rich and mycolic acid which is covalently attached to the peptidoglycan polysaccharide

providing an extraordinary lipid-barrier (Knechel, 2009). The lipid barrier is attributed to persistence and drug resistance that insure survival of the person inside the human host (Cole *et al*., 1998; Knechel, 2009). The composition cell wall affects the virulence and growth rate the pathogen (Lee *et al*., 2004). The peptidoglycan provides cell wall rigidity and permeability barrier of pathogen (Knechel, 2009). Additionally, the cell wall possess a lipoarabinomannan (carbohydrate) that is immunogenic and assists the survival of pathogen within macrophages (Lee *et al*., 2004). The dormancy state and slow generation time of the bacillus within infected cells contribute to the chronic nature and length of the treatment regimens (Cole *et al*., 1998). The genetic basis for dormancy and reactivation of TB remain unclear; however, it is presumed to be genetically programmed with the involvement of intracellular signalling pathways (Cole *et al*., 1998). The genome of *M. tuberculosis* has been entirely sequenced (reference strain H37Rv) and found to be 4.41 Megabytes long with an estimated 4000 predicted genes (Cole *et al*., 1998). However, the current body of literature on whole genome sequencing demonstrate that the information on the species genetic variability particularly in clinical isolates is underrepresented (Farhat *et al*., 2013; Köser *et al*., 2012).

1.7 Strain epidemiology

Advances in genetic techniques facilitate a better understanding of epidemiological and bio-physiological attributes of the pathogen. Introduction of sequencing and genotyping methods such as IS6110 restriction fragment length polymorphism, spoligotyping and mycobacterial interspersed repetitive units have been useful in understanding molecular epidemiology of mycobacteria as they generate new insight on drug resistance profile, transmission pathways, relapse, and re-infection of TB (Alland *et al*., 1994; Small *et al*., 1994a). However, information on the TB strain epidemiology and transmission dynamics in Tanzania remain limited, and where available, restricted to small geographical localities. One molecular study characterized 487 isolates from 23 regions reported a diverse range of spoligotype families in all zones of the country. The most common families were Central Asian (40.9%), the Latin American Mediterranean (18%), the East-African Indian families (11.5%), and the Beijing family (7%) (Mfinanga *et al*., 2014). In this study, Dar es Salaam contributed significantly to all main reported families. Whilst this study reported no spoligotype patterns were consistent with *M bovis*, notification of bovine TB incidence in Tanzania has significantly increased, with a varying prevalence ranging from 0.2 and 14 percent (Durnez *et al*., 2011; Katale *et al*., 2012).

Cattle are sought to be the principal source for *M. bovis* infection in a human, but other domesticated and wild animals can also be infected and may lead to cross-species transmission of the infection (Katale *et al*., 2017; Katale *et al*., 2012). The presence of diverse domesticated and wildlife reservoirs has been obstructing pathogen eradication efforts in most of the African countries (Katale *et al*., 2017). However, the zoonotic importance varies from country to country, signifying the need of country-specific interventions.

Although zoonotic TB contributes a small proportion of the overall human TB cases, the pathogen has serious implications for public health, hampering efforts to end TB by 2035. Bovine TB is well-known associated with extra-pulmonary TB (EPTB), and this has been evidenced in Tanzania (Katale *et al*., 2012), whereby EPTB cases are often reported in areas with high human-animal interactions. High bovine TB transmission is expected in a rural population where human interaction with livestock and wild animals is intense. However, the true incidence of bovine TB remains unknown because *M. tuberculosis* and *M. bovis* are indistinguishable unless the culture is performed in combination with biochemical and molecular characterization methods. Reports of *M. bovis* are based on localized studies, highlighting the need comprehensive surveys to accurately determine the true burden of bovine TB to public health. Furthermore, to adequately address the zoonotic importance of bovine TB, multi-level interventions are needed, that includes: testing of wildlife, livestock and water sources for pathogens where human and animal interaction is intense; introducing new rapid diagnostic techniques for disease detection; guidelines for infection screening *i*n domesticated animals; strengthening sensitization and communication strategies about zoonotic importance of the disease; and tough movement control of animals and biosecurity.

1.8 Burden of TB in Tanzania

Tanzania is ranked amongst 20 countries with highest TB incidence (WHO, 2016). Currently, national TB control programs in Tanzania are coordinated by the National Tuberculosis and Leprosy Programme (NTLP) established in 1977 under the Ministry of Health. The National TB and Leprosy Programme is organized at the central, regional, and district levels, working in collaboration with the National AIDS Control Programme, and other stakeholders to synergistically reduce the burden of TB-HIV co-infection. According to NTLP, the number of TB cases has gradually increased from 11 753 cases in 1983 (MoHSW, 2012) to 65 902 cases in 2016, almost six-fold (MoHCDGEC, 2016). Similarly, the number of HIV/AIDS cases increased from 3 cases in 1983 to 2 199 809 cases in 2012, thus, the rapid increase of

TB is primarily attributed to the HIV/AIDS epidemic. The majority of these cases appear in the young adult population aged 15-45 years, the same age group affected by HIV/AIDS. Several socio-economic factors such as poverty or malnutrition, smoking and excessive alcoholism have also recently received recognition (WHO, 2014).

The 2011-2012 Tanzania HIV/AIDS Indicator Survey reported 5.1% of the adult population were infected with HIV, whilst 36% of TB patients were co-infected with HIV (MoHCDGEC, 2016). One recent study in Tanzania reported five-fold and seven-fold mortality risk among TB-HIV patients on antiretroviral (ART) and ART-naïve TB-HIV patients, respectively (Nagu *et al*., 2017). The epidemic is fueled by high reactivation of latent TB infection among TB-HIV co-infected individuals (30-50%), compared to a lifetime risk of 5-10% in HIV- negative individuals, rapid progression to disease if HIV infected; and difficult to diagnose and cure (Sonnenberg *et al*., 2001). As a result of this co-morbidity has led integrated public health programs to synergistically address the two diseases. Like many other poor TB high endemic countries, data on latent TB infection prevalence is scarce; one study in urban Mwanza reported high latent TB infection prevalence among household contacts (59%) and neighbourhood of TB patients (41%) (Jensen *et al*., 2013). The overall prevalence of latent TB infection in sub Saharan Africa is estimated at 35% (Corbett *et al.*, 2006). Latent TB infection remains difficult clinical and public health issues that urgently need strategies to diagnose and treat it. Intensifying existing interventions for HIV and TB prevention, care and treatment, recording and reporting is of paramount important.

The 2012 National TB Prevalence Survey reported a prevalence of bacteriological confirmed pulmonary TB of 295 per 100 000 adult populations (MoHSW, 2013). In 2014, the WHO estimated the prevalence of all forms of TB to be 528 per 100 000 population with an annual incidence of 327/100 000 population (95% CI 155-561) and a case detection rate of only 36%, with nearly 60 000 new cases of TB diagnosed per year (WHO, 2015). In 2015, an overall 90% of 59 293 TB new and relapse TB cases were successful treated, 5.8% (3 531) died, 2% (1091) lost to follow up, and 0.2% (106) failed treatment (MoHCDGEC, 2016). Similarly, in 2016, there were 65 902 notified TB cases, 64 404 (95.5%) being new and relapse cases of which only 27 655 (39%) were bacteriologically confirmed, with overall treatment success rate of over 85%. Joint efforts that's involves expanding linkages to reach out academic and other research institutions, young academia, non-governmental institutions,

industries, practitioners, parliamentarians and Local Governments for an intensified TB research and interventions that are people-centered as drivers of public health.

Following an increasing trend of MDR-TB cases in Tanzania, in 2009, Ministry of Health designated Kibong'oto Infectious Diseases Hospital (KIDH) as the referral hospital for all MDR-TB cases in the country (MoHSW, 2009). In the same year, the GeneXpert diagnostic assay was introduced for research purposes and at the program level in 2011/2012. In 2013, pilot installation of the internet-based GxAlert initiative started, connecting the GeneXpert devices for automated real-time reporting of diagnostic results to facilitate case management.

As in many sub-Saharan Africa countries, the true burden of MDR-TB in Tanzania remains unknown due to inadequate laboratory infrastructures and lack of resources to diagnose and report MDR-TB. Data on drug resistance and scanty and argued to be underreported (WHO, 2016). In 2010, the TB drug resistance survey reported 1.1% MDR-TB prevalence among new cases and 3.9% among retreatment TB cases (Chonde *et al*., 2010). In 2014, laboratory-confirmed rifampicin resistant cases (presumed MDR-TB cases) were estimated at 516, of which only 143 cases were started on MDR-TB treatment (WHO, 2016). In this MDR-TB cohort, 108 (76%) patients were successful treated, 25 (17%) died, 10 (7%) defaulted and 2 (1%) were not evaluated (MoHCDGEC, 2016). In the 2016 MDR-TB cohort, there was a significant enrolment increase of MDR-TB patients on treatment, in which 196/158 (82%) patients were initiated with MDR-TB treatment. Such an increase is attributed to the ongoing decentralization of MDR-TB treatment which started in 2016 and by the end of this year, 22 sites had already started offering initiation of MDR-TB regimens (MoHCDGEC, 2016).

Although TB transmission and occurrence of MDR-TB has increased (WHO, 2016), and relative increase in numbers of MDR-TB patients enrolled each year, MDR-TB case detection rates are far still very low (WHO, 2016), suggesting that plateauing or exhaustion of current identification and prevention strategies. Whilst the dormancy state and slow generation time of the bacillus within infected cells contribute to the chronic nature and the lengthy of the treatment regimens with multiple drugs hence normally associated with poor adherence, information on adherence to national guidelines for management and follow-up of TB/MDR-TB patients in the country are scarce, highlighting the need of robust database for clinical data. A recent study on MDR-TB found only one-third of MDR-TB patients diagnosed by the GeneXpert between 2013 and 2016 were placed on an MDR-TB therapy and serious inadequacy data recording was observed (Mnyambwa *et al*., 2018). Delay in

diagnosis and management of MDR-TB poses a threat to the success of TB control programs and creates an enormous financial burden. The primary focus of the present study was to characterize the genetics of clinically drug resistant tuberculosis isolates from pulmonary TB patients in Tanzania.

1.9 Scientific rationale and justification

Evolving novel and unfamiliar mutations are revolutionizing the pathways of antibiotic resistance of clinical TB, challenging both diagnostic tools and treatment. Although known mutations explain much resistance in TB, several causative mutations in clinically resistant isolates are unknown and, even where such mutations have been identified, there may be additional mutants that nurture drug resistance (Farhat *et al*., 2013). It is estimated that between 10 and 40 percent of causative mutations in clinically resistant isolates have not yet been identified (Farhat *et al*., 2013). Furthermore, an accumulation and interactions of these poorly characterized mutations fuel the complexity of resistant pathogenic strains and raise public health concerns (Mnyambwa *et al*., 2017a). And consensus is growing that mycobacterial genotype may be strongly interconnected with a specific human population in a narrow geographical region (Bos *et al*., 2014; Couvin *et al*., 2015; Firdessa *et al*., 2013), which reframes global TB pandemic as a totality of genetically diverse outbreaks (Mnyambwa *et al*., 2017a). Despite the facts that Tanzania has high TB burden, diagnostic challenges and minimal information on resistant TB complicates building effective management strategies and posing a gap in achieving TB targets as indicated in the Sustainable Development Goals. To contribute to minimization of the gap, the present study employed Next Generation Sequencing and genotyping methods to characterize genetics of drug resistance amongst TB strains circulating in Tanzania.

1.10 Research questions

(i) What is the clinical implication of novel drug resistance-conferring mutations in resistant TB?

(ii) What is the drug sensitivity pattern of each identified TB isolates?

(iii) Are the identified clinical isolates of TB share the same DNA fingerprint?

(iv) What kind of genetic changes characterize drug-resistant tuberculosis in Tanzania?

(v) What is the proportion of notified MDR-TB cases by GeneXpert were linked to MDR-TB care?

1.11 General and specific objectives

The primary goal of this study was to characterize genetics of drug resistance and assessment of diagnostic challenges of clinical TB in Tanzania. This objective was attained by pursuing the following specific objectives:

(i) To assess the clinical relevance of novel resistance-conferring mutations in clinically resistant TB
(ii) To determine the drug susceptibility patterns of clinical MTBC isolates to first-line anti-TB drugs
(iii) To determine the DNA fingerprint of the strains causing tuberculosis as a measure of recent transmission of the disease and rapid progression to active TB
(iv) To determine and characterize mutations (genes, Inter-genic regions (IGRs) and/or genetic variants) associated with drug resistance in MTBC isolates
(v) Assessment of GeneXpert diagnostic system of MDR-TB in Tanzania

1.12 Significance of the study

The genetic data found in this study is important for both effective diagnosis and treatment of resistant TB. Data on DNA clustering is useful in addressing important epidemiological question such as; estimating recent TB infection-versus-reactive TB; rapid progression to active tuberculosis; the degree of exogenous re-infection; and transmission dynamics of TB. Measuring disease frequency due to recent transmission and rapid progression is important for the evaluation of TB control activities designed to prevent transmission. The general principle is that patients with isolates of *M. tuberculosis* that share the same DNA fingerprint results (clustered cases) are part of a chain of transmission and therefore developed disease as a consequence of recent transmission and rapid evolution to active disease (Lockman *et al*., 2001; Small *et al*., 1994b). The study results generally, contribute to the ongoing efforts to improve prevention, diagnosis and treatment of clinically resistant TB/MDR-TB in Tanzania.

CHAPTER TWO

Clinical implication of novel drug resistance-conferring mutations in resistant-tuberculosis[1]

Nicholaus Peter Mnyambwa[1, 2], Dong-Jin Kim[1], Esther S Ngadaya[2], Rudovick Kazwala[3], Pammla Petrucka[1,4], Sayoki G Mfinanga[2]*

[1]School of Life Sciences and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania. [2]National Institute for Medical Research, Muhimbili Medical Research Centre, Dar es Salaam, Tanzania. [3]Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania, [4]College of Nursing, University of Saskatchewan, Saskatoon, Canada

**Corresponding author's: Nicholaus Peter Mnyambwa:(lodnicho@gmail.com)*

Abstract

Evolving novel and/or unfamiliar mutations are revolutionizing the pathways of antibiotic resistance of clinical tuberculosis. Accumulation and interaction of these poorly characterized mutations augment the complexity of resistant pathogenic strains and raise public health concerns. This article reviews our current understanding of the genetic changes that characterize drug resistance in tuberculosis and highlight the imperative for further investigations focusing on the effects of an individual mutation and interacting mutations with detailed strain epidemiology, particularly as these pertain to technology-limited countries with high tuberculosis incidence rates. Concomitantly, there is a need for development, testing, and uptake of new tools for studying the effects of these mutations in drug resistance and fitness cost of the pathogen. Such genetic data is critical for effective localized and global tuberculosis control interventions and for accurate epidemiological predictions.

Keywords: Drug-resistant tuberculosis; *Mycobacterium tuberculosis* complex (MTBC); novel mutation; anti-tuberculosis agent; multi-drug resistant tuberculosis (MDR-TB): drug resistance conferring-mutation.

[1] *European Journal of Clinical Microbiology & Infectious Diseases 2017: **11** (36) 2021-2028*

2.1 Introduction

One of the most formidable challenges in modern public health is the emergence and pervasiveness of drug-resistant diseases with TB, caused by species and subspecies of Mycobacterium, known as the MTBC, being of high concern (Bouakaze *et al*., 2010). In 1993, the WHO officially declared TB as a global health concern, launching a series of concerted efforts to avert and combat the epidemic (WHO, 2013). These efforts have contributed to a plummeting global TB incidence, with an annual rate decline of 1.3% since 2000, and a reduction in mortality from 1.8 million in 2000 to 1.4 million in 2015 (WHO, 2016).

Globally 5% (480 000) of TB cases are categorized as multi-drug resistant (MDR-TB) with an additional 100 000 rifampicin-resistant (RR) cases which are eligible treated as MDR-TB. Nearly 1 in 10 of the combined total 580 000 MDR-TB cases are classified as XDR-TB (WHO, 2016). Typically, drug-resistant TB is categorized as mono-resistant; poly-resistant; MDR-TB; XDR-TB; and/or RR-TB (WHO, 2016). In the first two instances, drug-resistance is linked to one or more first-line anti-TB drugs (other than isoniazid and rifampicin) (WHO, 2013; WHO, 2016). MDR includes resistance to both isoniazid and rifampicin, whereas XDR includes resistance to at least one fluoroquinolone plus one or more second-line injectable drug. RR may be in conjunction with or without resistance to any other anti-TB drug and is detected using phenotypic or genotypic methods (WHO, 2013).

Complications stemming from the development of MDR-TB and XDR-TB fuelled by global epidemics of Human Immunodeficiency Virus (HIV) have hampered eradication efforts (Gandhi *et al*., 2010; WHO, 2013). The complexities inherent in these emerging disease patterns have challenged diagnostic and intervention approaches. More concerning is that recent evidence indicates a more complex genetic cause of drug-resistant tuberculosis than previously envisioned (Farhat *et al*., 2013; Zhang *et al*., 2013).

Advances in genetic techniques have allowed penetration and insights into many epidemiological and bio-physiological attributes of this infectious disease. In 1998, the first whole sequenced genome of *M. tuberculosis* (H37Rv) was published (Cole *et al*., 1998), which revolutionized our understanding of the biology of this pathogen. The penetration of high-throughput sequencing platforms has fundamentally changed how scientists study biological attributes, including drug resistance-conferring mutations of these pathogens. To

date, public databases have been populated with hundreds of genomes representing the global array of genetically diverse tuberculosis genotypes. Using sequencing technologies, assessment of bacterial heterogeneity in an individual patient and prediction of the pathways to drug-resistant mutations becomes plausible (Ford *et al*., 2012). Despite these advancements, our understanding of the spectrum and nature of genetic alterations that characterize drug resistance in the MTBC is incomplete and fragmented (Farhat *et al*., 2013). Further, efforts to characterize novel and/or additional causative mutations are required to fully explicate the evolution of drug-resistant TB. This review reflects current knowledge of the genetic signatures of resistance, their interconnectedness to the molecular diversity of tuberculosis strains and human demography, and elements of the unresolved genetic basis of resistance in tuberculosis.

2.2 Evolution of drug resistance in MTBC strains

Drug resistance is a complex phenomenon determined by multiple biochemical, physiological, and environmental factors. Clinically, drug-resistant TB is primarily attributed to non-compliance with prescribed treatment among patients with active TB (WHO, 2013). As a result, spontaneous sequential chromosomal mutations (with fitness benefits) are selected in genes that classically encode drug targets or drug-metabolizing enzymes (Philip Supply *et al*., 2003). Alterations in other classes of genes have been associated with a selective advantage in the presence of drugs through: (i) down-regulation of cell-wall permeability or promote drug-efflux pumps (Farhat *et al*., 2013); or (ii) compensation for the reduction in fitness costs of other drug resistance-conferring mutations (Farhat *et al*., 2013; Zhang *et al*., 2013). The fitness cost of specific mutations is an important determinant of the spread of drug-resistant strains (Gagneux *et al.,* 2006). Both single nucleotide polymorphisms (SNPs) and large sequence polymorphisms (LSPs) may result in drug-resistant mycobacterial strains (Gagneux *et al.,* 2007). The mutation rate at which resistance emerges varies significantly between and within genes (Gillespie *et al.,* 2002).

The most reported and well-characterised classical drug resistance-conferring mutations in MTBC spoligotypes occur in genes: *katG, kasA, inhA* and *ndh* for isoniazid (INH); *ethA* and *inhA* for ethionamide; *rpoB* and *rpoC* for rifampicin (RIF); *rpsL* and *rrs* for streptomycin (SM); *embCAB* for ethambutol (EMB); *pncA* for pyrazinamide (PZA); *gyrA* for fluoroquinolone (FQ); and *eis* for amikacin (AMK) and kanamycin (KAN) (Campbell *et al., 2011; Müller et al*., 2013) and intergenic regions (IGRs): *proA–ahpC, embC–embA, eis–*

Rv2417c and *Rv1482c–fabG1* (Almeida *et al*., 2011; Ramaswamy and Musser, 1998; Sekiguchi *et al*., 2007). Additional listing of partially characterized mutations strongly associated with drug resistance in TB strains is reported in Table 1. A discussion of the novel drug resistance-conferring mutations in MTBC is offered in the next sections.

Table 1: Partially characterized mutations (gene/IGR) that strongly correlate with resistant-tuberculosis

Anti-TB Drug	Gene/IRG	Description of the mutation
Isoniazid (INH)	*kasA, inhA, ahpC* and *ndh* and intergenic *oxyR-ahpC*	kasA mutations are observed in both susceptible and resistant strains (Mdluli *et al.*, 1998; Ramaswamy and Musser, 1998), while *inhA* is associated with low-level INH-resistance (McMurry *et al*., 1999). *oxyR-ahpC* is speculated to reduce the expression of inhA (Dalla Costa *et al*., 2009); whereas *ndh* confers resistance to INH and ethionamide in *M. tuberculosis bovis* (Vilchèze *et al*., 2005).
Rifampicin (RIF)	*rpoA ponA1* and *pncA*	Putative mutations in the *rpoA* are associated with low-level cross-resistance to RIF and other RIF-derivatives (de Vos *et al*., 2013). Gene *ponA1* demonstrated in vitro growth advantage in the presence of the RIF (Farhat *et al*., 2013), while mutations in pncA are linked as the main causative of pyrazinamide-resistance (Juréen *et al*., 2008). Mutations in *katG, ahpC,* and *inhA* are associated with RIF resistance (Hazbón *et al*., 2006a).
Streptomycin (SM)	*gidB*	*gidB* causes low-level resistance in some clinical SM-resistant strains that exhibited no mutation in either rrs or rpsL (Spies *et al*., 2011a), with susceptible strains confirmed as harbouring similar mutations (Spies *et al*., 2011a; Wong *et al*., 2011).
Ethambutol (EMB)	*ubiA*	Mutations in *katG315* favour ethambutol-resistance (Hazbón *et al*., 2006a), while those in *ubiA* are implicated to high-level EMB-resistance (Lingaraju *et al*., 2016).
Pyrazinamide (PZA)	*RpsA* and *panD*	Some PZA-resistant strains show no mutations in pncA neither in its promoter region (Sreevatsan *et al*., 1997). *M. bovis* isolates are reported to be naturally resistant to pyrazinamide (Jong *et al*., 2005).
Fluoroquinolone (FQ)		Amino acid substitutions in gene *gyrB* were recently implicated with FQ resistance; however, the genetics of these mutations is poorly characterized (Singh *et al*., 2014).
Aamikacin (AMK), Capreomycin (CAP), Kanamycin (KAN) and Viomycin (VIM)	*tlyA, RpsA and panD*	Drug resistance-conferring mutations in the *rrs* A1401G is commonly associated with cross-resistance between injectable resistance-conferring mutations in the *rrs.* A1401G is commonly associated with cross-resistance between injectable drugs (Alangaden *et al*., 1998). Recently, mutations in the gene *tlyA* conferring resistance to CAP and VIM were identified (Campbell *et al*., 2011), and mutations in *RpsA* observed in some PZA-resistant strains. The role of *panD* mutations in some resistant strains remains inconclusive (Zhang *et al*., 2013).

2.3 Transmissibility and fitness cost of the MDR/XDR-TB

A complex and hard-to-cure MDR/XDR-TB is emerging worldwide. A deadly form of TB called totally drug-resistant TB, which do not to respond to both first-line and second-line anti-TB agents, has been increasingly reported among patients with MDR-TB [32]. Thus, an accurate prediction of the course of the future TB epidemic is difficult to anticipate, calling for more attention from the scientific community.

The first-line anti-TB agents are INH, RIF, PZA and ETH while FQ and injectable drugs are classified as second-line. Acquired drug-resistant TB occurs during treatment when a drug regimen is ineffective in destroying a particular mycobacterial strain. Alternatively, primary drug-resistant TB describes the direct inter-personal transmission of a resistant strain. Accumulation of mutations in different genomic regions may result into MDR/XDR-TB and which may further get enhanced with epistatic interactions (Fig. 2). A varying fitness cost and transmissibility of MDR/XDR-TB have been suggested (Cohen *et al*., 2015), which depends on a number of biological factors, such as genetic background, compensatory mutations, and epistatic interactions as reviewed in the next sections. A high proportion of MDR/XDR-TB is reported amongst immunocompromised patients with HIV co-infection (WHO, 2016) suggesting that responsible strains might have reduced fitness cost compared to counterpart strains. Prolific MDR/XDR strains have been reported in Russia and Northern Europe despite low rates of HIV infection (Couvin and Rastogi, 2015). This drug resistance pattern can be attributed to a long-standing host-pathogen association, which is indicative of natural selection of the pathogen through compensatory mutations. Such strains constitute the most significant public health concern regarding future untreatable TB epidemics.

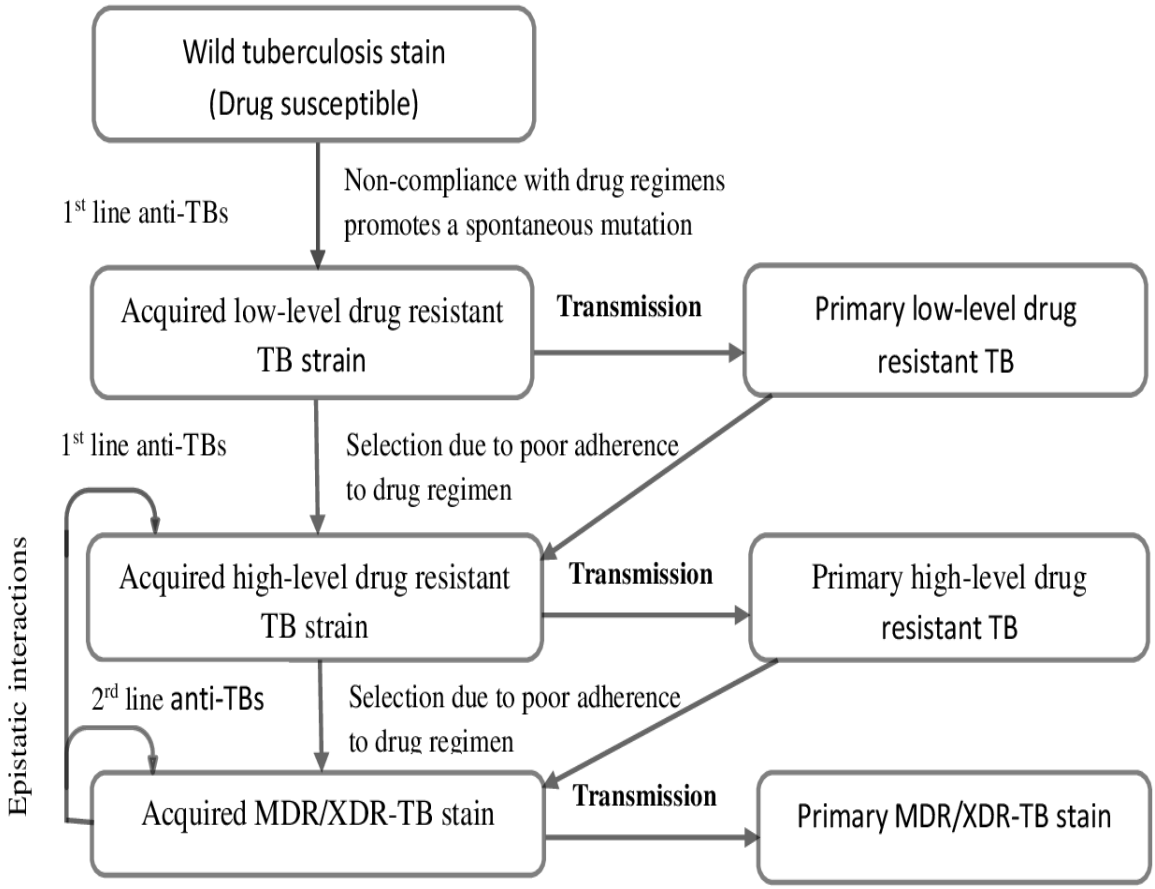


Figure 2: A simplified theoretical step-wise development of drug resistance in MTBC due to mutations.

From Fig. 2, a high-level drug resistance is a result of an accumulation of chromosomal mutations in drug targets, which may be enhanced by epistatic interactions. Accumulation of mutations in and out of the drug targets, working in concert, may yield or catalyse evolution of the MDR/XDR–TB strains.

2.4 Novel mutations in MTBC: clinical relevance and controversies

The evolution of Mycobacterium species usually includes deletion of non-functional genes, insertion of aberrant genes, and/or a combination of these events, which potentiates their survival under different environmental conditions or geographical niches (Rahman *et al*., 2014). In MTBC, mutations have been observed in limited hotspots of coding genes (Goldberg *et al*., 2011). However, it is estimated that between 10 and 40% of causative mutations in clinically resistant-isolates have not yet been identified (Campbell *et al*., 2011). MTBC mutations (SNPs, insertion or deletion) have a role in the development of drug resistance if the outcome of mutation is a beneficial phenotype in the face of drug selection. A hierarchy of beneficial mutation types is more likely to benefit MDR/XDR development in

the pathogen as demonstrated in Fig. 2.

In 2013, two studies (Farhat *et al*., 2013; Zhang *et al*., 2013), which analysed hundreds of TB genomes from clinical settings recovered all classical mutations and identified new genome-wide signatures and SNPs that were strongly correlated with drug resistance. Using sequences of 123 TB isolates from a global pool of samples, Farhat and his colleagues reported 39 novel genomic regions of positive selection, of which 11 genomic regions have annotated functions and the remaining 28 regions, of which 16 belong to the PE/PPE gene family, have unknown functions. A recently published study by (Cui *et al*., 2016), provides additional evidence implicating genes from PE/PPE family in drug-resistant tuberculosis. PE/PPE genes are highly polymorphic - unique sets of genes that constitute 10% of the *M. tuberculosis* genome (Fishbein *et al*., 2015). Despite considerable research interest in the PE/PPE families, their roles in the survival of the pathogen remain unclear. Presumably the PE/PPE genes code for cell surface-associated proteins that may play an important role in the modulation of host immune responses through antigenic variation (McEvoy *et al*., 2012). Zhang and his colleagues sequenced 161 isolates from Chinese patients and identified novel mutations in 72 new genes, 28 IGRs, and 21 other variants, which were strongly correlated with drug resistance (Zhang *et al*., 2013). IGRs have received little research attention; hence, their involvement in the evolution of drug resistance in MTBC remain elusive (Kuan *et al*., 2015). Both studies (Farhat *et al*., 2013; Zhang *et al*., 2013) demonstrated the presence of putative mutations likely to influence some resistant genes by enhancing their activity or compensating for their detrimental outcomes and survival. More recently, comparative gene functional analysis revealed novel mutations in *ald* that encodes L-alanine dehydrogenase and a loss of its function (Safi *et al*., 2013). The mutations were significantly associated with unexplained drug resistance, including to a toxic cycloserine drug. Cycloserine remains a cornerstone in the treatment of both MDR-TB and XDR-TB, despite its adverse neurological health effects. Hence, increasing resistance to cycloserine would contribute to increased incidences of untreatable tuberculosis. One study (Desjardins *et al*., 2016) presented observations linking ETH resistance to an accumulation of multiple mutations in a step-wise manner, initiated in *embB,* then in *Rv3806c* or *Rv3792,* and finally in *embC*. This exemplar of a multistep emergence of drug-resistance supports previous findings that reported a subpopulation of resistant-TB strains can spontaneously evolve in the presence of anti-TB agents and may progress step-wise to become multidrug resistant (Gandhi *et al*., 2010).

Table 2 summarizes a number of current novel genes and IGRs, primarily from within PE/PPE, *fadD, mmpL,* and *pks* families, which are strongly associated with drug resistance in MTBC clinical isolates. Representatives of *fadD, mmpL* and *pks* families are highlighted by Zhang and his colleagues (Zhang *et al*., 2013). More recently, findings suggested that mutations in *fadD, mmpL,* and *pks* genes may play a vital adaptive and survival role for the pathogen (Kuan *et al*., 2015). While Fishbein and his colleques highlighted regions *Rv2670c* and *mmpL11* as possible targets of drug resistance in the Beijing lineage (Fishbein *et al*., 2015).

Generally, uncertainty remains on how the reported mutations could have directly impacted resistance and/or conferred compensatory effects for the fitness of a resistant strain. Additionally, it is unclear how these mutations interact with each other and influence the resistance pattern of a particular strain. However, there is an extensive evidentiary base that a number of mutations modulate drug resistance and in some cases, working in concert. Thus, addressing resistance in TB is not sufficient within past individualist, situational approaches, instead of requiring a system-wide response rooted in current knowledge of the pathogen at the genome level.

Table 2: Key genomic regions with interesting novel drug resistance-conferring mutations in MTBC

Novel Gene/IGRs	**Resistance**	**Description**	***Novel Gene/IGRs***	**Resistance**	**Description**
fadE5	KAN	Acyl-CoA dehydrogenase fadE5	*PE_PGRS10*	Not Specified	PE-PGRS family protein
mmpL1	ETH, CPM	Transmembrane transport protein mmpL1	*PPE55*	PZA	PPE family protein
fadD30	OFX, ETH	Fatty-acid-CoA ligase fadD30	*PE_PGRS3*	Not Specified	PE-PGRS family protein
fadA	OFX, KAN	Acyl-CoA thiolase fadA	*pks12*	RIF, INH, ETH, KAN	Polyketide synthase 12
fadD14	EMB	Medium chain fatty-acid-CoA ligase fadD14	*PE_PGRS1*	Not Specified	PE-PGRS family protein
PPE20	KAN	PPE family protein	*pks3*	Not specified	polyketide beta-ketoacyl synthase
pks8	EMB	polyketide synthase pks8	*PPE54*	PZA	PPE family protein
pks17	OFX	Polyketide synthase pks17	*PPE3*	Not specified	PPE family protein
pks15	EM, KAN	Polyketide synthase pks15	*PE_PGRS48*	Not specified	PE-PGRS family protein
fadE33	ETH	Acyl-CoA dehydrogenase fadE33	*PE_PGRS50*	Not specified	PE-PGRS family protein
pks2	OFX	Polyketide synthase pks2	*PPE47*	Not specified	PPE family protein
PE_PGRS4	Not specified	PE-PGRS family protein	*PE_PGRS47*	Not specified	PE-PGRS family protein
PE_PGRS6	Not specified	PE-PGRS family protein	*PPE60*	KAN, CAP, PZA	PPE family protein
PE_PGRS9	PZA	PE-PGRS family protein	*PE_PGRS53*	Not specified	PE-PGRS family protein
PPE9	Not specified	PPE family protein	*Ald*	CYS	L-alanine dehydrogenase

ETH: Ethionamide, OFX: Ofloxacin, LEVO: Levofloxacin, CIP: Ciprofloxacin, CYS: Cycloserine. The rest of drug abbreviations defined previously. The first 11 genes were considered in reference (Zhang *et al*., 2013), the remaining in (Farhat *et al*., 2013) with exception of *ald* which is described (Safi *et al*., 2013)

2.5 Strain diversity and drug resistance

For years, species and subspecies of human-adapted MTBC have been considered highly genetically-related (Musser *et al*., 2000; Sreevatsan *et al*., 1997), a popular classification based on the phylogeny putting the global population structure and diversity into six main lineages (Hershberg *et al*., 2008). Lineages 1–4 are geographically distributed worldwide while lineages 5–6 are restricted to select parts of Africa (Comas *et al*., 2013). A novel lineage 7 was recently described in Ethiopia (Firdessa *et al*., 2013). However, mounting evidence suggests that strains of the MTBC rubric are highly genetically-diverse (Hershberg *et al*., 2008; Homolka *et al*., 2010), and a new standard universal nomenclature based on whole-genome sequencing is urgently sought (Comas *et al*., 2009). The genetic diversity may have important implications, not only for the drug-sensitivity pattern but also to the nature of mutations selected during treatment (Köser *et al*., 2012). Evidence suggests human-adapted MTBC spoligotypes form ecotypes that are strongly interconnected to human demography aligning along a narrow geographical scale (Bos *et al*., 2014; Couvin and Rastogi, 2015; Firdessa *et al*., 2013), which might, in turn, affect the nature of drug resistance-conferring mutations. One prospective study reported tuberculosis strains in Russia are becoming more resistant to antibiotics than in any other parts of the world (Casali *et al*., 2014). This study, which sequenced 1000 clinical isolates from patients living in Samara, Russia, reported about half (48%) were MDR-TB strains, of which 16% had additional drug resistance resulting in reclassification as XDR-TB (Casali *et al*., 2014). The Beijing genotype was found to have a higher proportion of drug resistance than Euro-American strains, clear evidence that the rate of acquisition of drug resistance-conferring mutations and transmissibility is also accelerated in Beijing strains.

Study by Fenner and colleagues (Fenner *et al*., 2012), demonstrated that amongst strains harbouring the same kind of drug resistance-conferring mutations to INH, there were different drug resistance patterns. In this scenario, epistatic interactions due to either strain genetic background or other compensatory mutations may have operationalized in the variable modulation of drug resistance profiles among TB strains. The influence of pathogen-host interaction to both strains epidemiology and the emergence of drug resistance-conferring mutations is not well established. Despite a global presence of *M. tuberculosis*, only 1 in 10 of the infected population will ever succumb to active TB (Murray *et al*., 1990), suggesting the relevance of host genetic predisposition to the disease. The evidence suggests a long-term

interaction between host-pathogen could lead to adaptive genetic changes of the interacting host-pathogen populations (Comas *et al*., 2013). Recently, genetic analyses focusing on the interaction of the human-genetic variant and MTBC genotypes has provided evidence that a mycobacterial genotype is strongly interconnected to a specific human population in a narrow geographical region (Möller and Hoal, 2010). The implications of such association in the emergence of drug resistance-conferring mutations need to be established. However, a consensus is growing that successful resistant-TB control strategies need to integrate host-genetics (Martínez *et al*., 2011), in addition to pathogen genetics, as well as environmental and social factors. Concurrently, acknowledging that a global TB pandemic is a totality of genetically-diverse outbreaks; hence, a call for well-defined localized programs at national and regional levels is necessary to determine and invoke meaningful global interventions.

Both experimental and clinical findings have established varying infection outcomes (Merker *et al*., 2015; Ribeiro *et al*., 2014) and acquisition of drug resistance among diversifying TB strains (Merker *et al*., 2015). A high prevalence of drug resistance is observed in the Beijing family of *M. tuberculosis,* which is globally widespread and most prevalent in East Asia (Merker *et al*., 2015). The lowest frequency of drug resistance is reported among lineage 1 strains, which normally are associated with low-level resistance (Fenner *et al*., 2012; Singh *et al*., 2015). Lineage 2 acquired drug resistance in *vitro* was found to be more rapid than in lineage 4 (Ford *et al*., 2013), thus increasing the likelihood of patients infected with a drug-susceptible lineage 2 strains to develop an MDR strain at the time of multiple-drug treatment. The study also highlighted the role of nonsynonymous SNPs in the emergence of resistance to second-line anti-TB agents. A more recent study reported five (two being new discoveries) non-synonymous mutations in resistant TB (Cui *et al*., 2016), and other novel mutations. Thus, suggesting a significant contribution of non-synonymous SNPs in purifying selection in MTBC, particularly in the survival of the pathogen in the face of drug selection.

2.6 Conclusion

A large number of emerging novel drug resistance-conferring mutations pose a new challenge in defining the genetic cause of drug resistance in TB, presenting many and complex genetic routes to the development of drug resistance in MTBC. It is unclear to what extent the continuous evolution of drug resistance-conferring mutations is represented at the phylogeographic scale. Developing new methods of studying the effects of a single mutation and groups of mutations in drug resistance, transmissibility, and fitness of the pathogen is of

paramount importance. Additional research is needed to decipher the genetic signatures with detailed strain epidemiology and their implications to the emerging clinical drug resistant-TB. This evidentiary base will promote effective localized and global interventions, including improved early diagnosis, tailored therapies, and accurate prediction of the course of future resistant TB epidemics, particularly MDR/XDR-TB cases.

Conflicts of interest: Authors declare no conflicts of interest

Acknowledgement: NPM is immensely grateful to the Nelson Mandela African Institution of Science and Technology for providing adequate learning resources and invaluable support.

CHAPTER THREE

Genetic characterization of clinical multi-drug resistant tuberculosis strains in Tanzania reveals possible putative mutations

Nicholaus Peter Mnyambwa[1,2*], Dong-Jin Kim[1], Esther Ngadaya[2], Kennedy Addo[3] Rudovick Kazwala[4], Sayoki Mfinanga[2]

[1]School of Life Sciences and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania. [2]National Institute for Medical Research, Muhimbili Medical Research Centre, Dar es Salaam, Tanzania. [3]Bacteriology Department, Noguchi Memorial Institute for Medical Research, Accra, Ghana. [5]Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania.

[*]Corresponding author: Nicholaus Peter Mnyambwa (lodnicho@gmail.com)

Abstract

Tanzania is ranked among the 20 countries with the highest number of tuberculosis cases and trends of tuberculosis cases have been increasing for the past 10 years, however, minimal information of drug resistance complicates building effective control strategies. This study characterized 40 MDR-TB clinical isolates in Tanzania with respect to the first-line anti-TB drugs (rifampicin, isoniazid, streptomycin and ethambutol). Drug susceptibility Testing demonstrated that: 12/ 40 (30%) were resistant to all 4 drugs; 36 /40 (90%) rifampicin resistance; 35/40 (87.5%) isoniazid resistance; 24/40 (60%) streptomycin; and 19/40 (47.5%) ethambutol resistance. Whole genome sequencing recovered all known drug-resistant genes: *katG*, *inhA*, *ethA*, *inhA*, *embCAB*, *rpoB*, *rpoC*, *rpsL*, *gyrA*, *eis*, and *pncA*; the majority of these genes had detectable additional polymorphisms. Of particular note, additional mutations in 3 other genes (*ndhC, ndhI* and *ndhK*) were strongly linked to drug resistance. Species identification revealed that 37 were *M. tuberculosis*; one *M. yongonense*; one *M. africanum* and one isolate had insufficient sequence read counts and it was excluded in this analysis. These findings expand the spectrum of potential drug resistance-conferring mutations in MTBC clinical isolates, suggesting the need of molecular diagnostic for delineating the drug susceptibility profiles of MTBC isolates in clinical settings.

3.1 Introduction

Tuberculosis is the ninth leading cause of death globally and the leading cause of a single infectious agent, ranking second above HIV/AIDS (WHO, 2017). The emergence and widespread of drug-resistant TB, especially multidrug resistant-TB (MDR-TB) and XDR is one of the most formidable obstacles of the modern public health (WHO, 2014a). MDR-TB infection is caused by bacteria that are resistant to treatment with least two of the most potent first-line anti-TB medications, rifampicin and isoniazid while XDR is a form of MDR with additional resistance to at least anti-TB drug in both of the two most potent classes of medicines in an MDR-TB regimen: fluoroquinolones and second-line injectable agents (amikacin, kanamycin or capreomycin) (WHO, 2016). The global epidemics of highly lethal MDR-TB/XDR-TB fuelled by Human Immunodeficiency Virus (HIV) have substantially challenged diagnostic and management of the disease (Gandhi *et al*., 2010; Lange *et al*., 2014; WHO, 2016). A particularly alarming aspect is the recent evidence implicates a large number of novel mutations in resistant TB, suggesting a more intricate genetics underpinning resistance in pathogenic TB strains than previously anticipated (Farhat *et al*., 2013; Zhang *et al*., 2013).

In 2016, there were 490 000 (5%) new MDR-TB cases worldwide, with an additional 110 000 people with rifampicin-resistant TB, who are also eligible for MDR-TB treatment. Of the combined 600 000 MDR cases, nearly 1 in 7 were categorized as XDR-TB. During the same period: 54% of MDR-TB cases (2014 cohort) were successfully treated; 16% died; 15% were lost to follow-up; 8% treatment failed; 7% had no outcome information with only 30% treatment success for XDR-TB cases (WHO, 2017). Clinically, drug-resistant TB is mainly attributed to incorrect or inadequate treatment among patients with active TB (WHO, 2013, 2014a), which may trigger spontaneous sequential chromosomal mutations in genes that classically encode drug-metabolizing enzymes or drug-targets (Philip *et al*., 2003). A serial of beneficial mutations (SNPs, insertion or deletion) is more likely to promote the evolution of MDR/XDR in the pathogen (Mnaymbwa *et al*., 2017a).

Although data on drug-resistance are limited in Tanzania, MDR-TB cases have been reported to increase since 2005, with more than 700 MDR-TB cases notified by the GeneXpert® during the period 2013-2016, forming only a proportion of all MDR-TB cases diagnosed during the same time period in the country (Mnyambwa *et al*., 2018). Here we have used

genomics to investigate the genetic cause of resistance in MDR-TB clinical isolates in Tanzania.

3.2 Materials and Methods

3.2.1 Study setting and population

This study utilized clinical *M. tuberculosis* isolates collected between 2014 and 2016 at CTRL. Central TB Reference Laboratory is the only reference laboratory in Tanzania receiving referral TB samples through routine surveillance of TB drug resistance in the country. Patients' demographic variables (age and sex) and clinical information (HIV status and TB treatment history) were extracted from the laboratory register books and tools at the sites and CTRL. Human immunodeficiency virus tests were performed at a respective health facility level as per standard care as recommended by the national HIV and testing algorithm, in order to determine the HIV status of the consented participants.

3.2.2 Culture and drug susceptibility testing

Culture and DST were performed using a standard protocol (WHO, 2009a) in a contained laboratory Biosafety Level 3 at CTRL, Dar es Salaam. In summary, bacterial isolates were inoculated onto the Lowenstein-Jensen (LJ) media slants and incubated at 35-37^{o}C until growth was observed or discarded as negative after 8 weeks. All drug sensitivity assays were performed in triplicate, and mutants resistant to a particular drug were determined by standard proportion method. Any critical proportion score of 1% or more of bacilli resistant to any of the 4 first-line drugs: isoniazid (0.2/1.0 mcg/mL), rifampicin (5.0 mcg/mL), streptomycin (4.0 mcg/mL) and ethambutol (2.0 mcg/mL) were classified as a resistant strains to that drug or if no growth observed were regarded as drug-sensitive/susceptible strains after the duration of 4 weeks. Resistance to both isoniazid and rifampicin were termed as MDR-TB case.

3.2.3 Genomic DNA extraction

The genomic DNA was extracted from heat-killed *M. tuberculosis* using standard protocol (van Soolingen *et al*., 1991) performed at CTRL, Dar es Salaam, Tanzania. Briefly, using a sterile loop, viable colonies were picked from the LJ slants and emulsified them in the appropriate tubes containing 400µL 1X Tris-EDTA buffer, pH 8.0, then heat-killed in the waterbath at 80°C for 20 minutes. Addition of 50µL 10 mg/mL Lysozyme in each tube was

followed by incubation at 37°C in the incubator overnight. Ten percent 70µL sodium dodecyl sulphate (SDS) and 5µL 10 mg/mL proteinase K was added followed by incubation at 65°C for 10 minutes. An aaddition of 100µL 5M NaCl was followed by a mixture 100µL N-acetyl-N,N,N-trimethyl ammonium bromide and NaCl, CTAB/NaCl (10 % CTAB in 0.7 M NaCl, pre-warmed to 65°C) and incubated at 65°C for 10 minutes. Seven hundred fifty microliters of chloroform-isoamyl alcohol (24:1) were added to each tube, and the tubes were centrifuged at 10,000 g for 5 minutes. The aqueous supernatants were transferred into the tubes containing 450µL of ice-cold isopropanol and placed at -20°C for 30 at least minutes and then microfuge at 10 000 g for 15 minutes. Supernatants were removed and the pellets washed with 1mL ice-cold 70% ethanol. After microfuge at 10 000 g for 5 minutes, ethanol was discarded. The tubes were then laid/tilted with open lids to allow the pellets to air-dry for at least 15 minutes. The pellets of genomic DNA were rehydrated in 80µL TE and left overnight at 4°C. The gDNA was then temporarily stored at -20°C before shipped to Kilimanjaro Clinical Research Institute (KCRI) for sequencing procedure.

3.2.4 Whole genome sequencing

The quality and quantity of gDNA were confirmed using Qubit 2.0 fluorometer (Thermal Fisher Scientific, Waltham, MA USA). Dual-indexed (24) DNA libraries preparation were constructed by shearing genomic DNA ~ 300 bp using NexteraXT DNA sample preparation kit according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA). Genomic DNA libraries were sequenced using a paired-end 2x250 bp protocol (MiSeq Illumina) available at KCRI, Moshi, Tanzania.

3.2.5 Sequencing data quality control

FastQC was used to assess the quality of the raw sequence data. The key metric to be monitored here is "Per base sequence quality". This is a plot of the entire sequence reads per sequencing cycle using Phred quality scores ($Q = -10 \log_{10} P$; where P is the base calling-error probability). The Lander/Waterman (Lander and Waterman, 1988) equation was used to calculate sequencing coverage.

3.2.6 Data pre-processing

The presence of poor quality or technical sequences such as adapters in next-generation sequencing data can easily result in suboptimal downstream analyses. Data preprocessing is an important and critical process in the data analysis. Inappropriate data preprocessing can influence and affect the whole downstream analysis and may mislead to a wrong biological conclusion. Trimmomatic (Bolger *et al*., 2014) and FastX ToolKit (Hannon Lab) were used to remove Illumina adapter sequences and filter out low quality sequenced bases and reads.

3.2.7 *de novo* genome assembly

In this study, we used widely used software for *de novo* assembly, contig ordering and annotation. Preprocessed sequence reads were assembled with SPAdes (Bankevich *et al*., 2012) using default parameters to generate contigs and scaffolds. Scaffold-level assembly was evaluated using QUAST (Gurevich *et al*., 2013) with the reference genome (*M. tuberculosis* H37Rv) to generate summary statistics, for example, N50, maximum scaffold length, G+C content and others. The assembled scaffold sequences were ordered using ABACAS (Assefa *et al*., 2009) with *M. tuberculosis* H37Rv as a reference genome for 38 samples and *M. yongonense* 05-1390^{T} and *M. yongonense* for the N18 sample.

3.2.8 Annotation

Prokka (Seemann, 2014) software tool was used to annotate genomes and produce standards-compliant output files. Prokka identifies protein-encoding, tRNA genes, tmRNA and assigns functions to the predicted genes.

3.2.9 Gene presence/absence analysis

Gene content (presence/absence) data together with clustering was used to generate a heat map (Chun's bioinformatics lab). The construction utilizes tetra-nucleotides of gene encoding proteins (CDSs) through the whole genome. Jaccard coefficient was used to calculate similarities between genomes based on presence/absence data.

3.2.10 Variant detection

Alignment to reference genomes: A total of 38 samples were aligned to the *M. tuberculosis* H37Rv genome sequence and only sample N18 was aligned to *M. yongonese* using BWA aligner (Li and Durbin, 2009). The references were indexed using BWA aligner and read1

and read2 were aligned to the reference using the BWA-MEM tool to generate a '*'.sam file. This file was then converted to '*'.bam and subsequently sorted by coordinates to '*'_sorted.bam then indexed using samtools (Li *et al*., 2009).

Variant detection and annotation: Genome Analysis Toolkit (McKenna et al., 2010) was used for variant detection against the respective reference genomes using standard parameters as recommended. The resulting VCF file was then annotated using SnpEff (Cingolani *et al*., 2012) to indicate synonymous vs non-synonymous variants in all the genic regions.

Phylogenetic Analysis: The multi-sample VCF file for the 38 samples subsequently used for neighbour joining phylogenetic re-construction using VCF-kit, (Andersen lab).

3.3 Results

3.3.1 Socio-demographic and clinical information

Deoxyribose nucleic acid samples were prepared from 40 MDR -TB patients aged (SD) 37.7 (13.4) years. Two-thirds (75%) of the 40 MDR-TB patients were males with an average age of 36.6 (12.0). The mean age of 10 female patients was 40.8 (12.2) years. Out of 40 MDR-TB cases, 18 (45%) were classified as relapse cases, while 10 (25%) of the study participants were HIV-positive. Nearly half (45%) of these cases were from Dar es Salaam.

3.3.2 Drug susceptibility patterns

Drug susceptibility testing to first-line anti-TB drugs revealed; 12/ 40 (30%) were resistant to all four drugs; 36 /40 (90%) rifampicin resistance; 35/40 (87.5%) isoniazid resistance; 24/40 (60%) streptomycin resistant and 19/40 (47.5%) ethambutol resistance (Fig. 3).

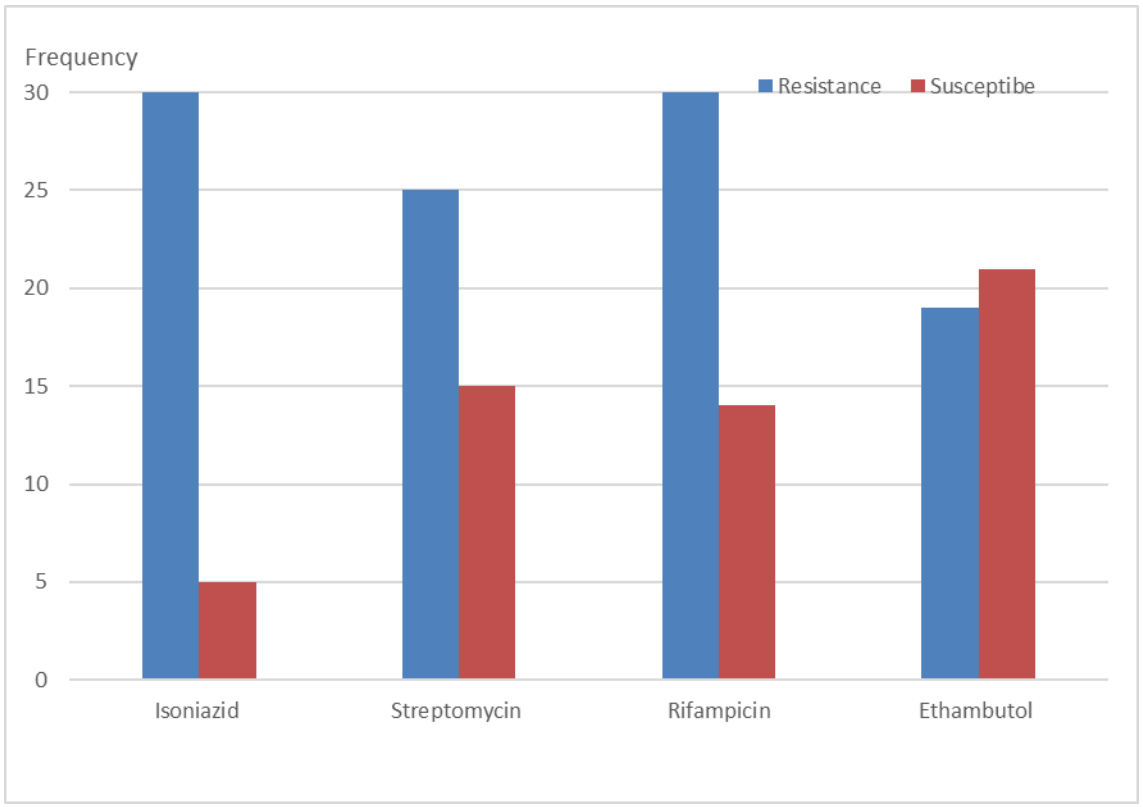


Figure 3: Frequency of phenotypic drug resistance profile of the four first-line anti-TB drugs

3.3.3 Pre-processing quality control

Pre-processed read QC showed that quality of each sequence bases (Phred Score >25) suitable for downstream analysis. Overall mean quality scores for all the preprocessed data is sumarized in Fig. 4.

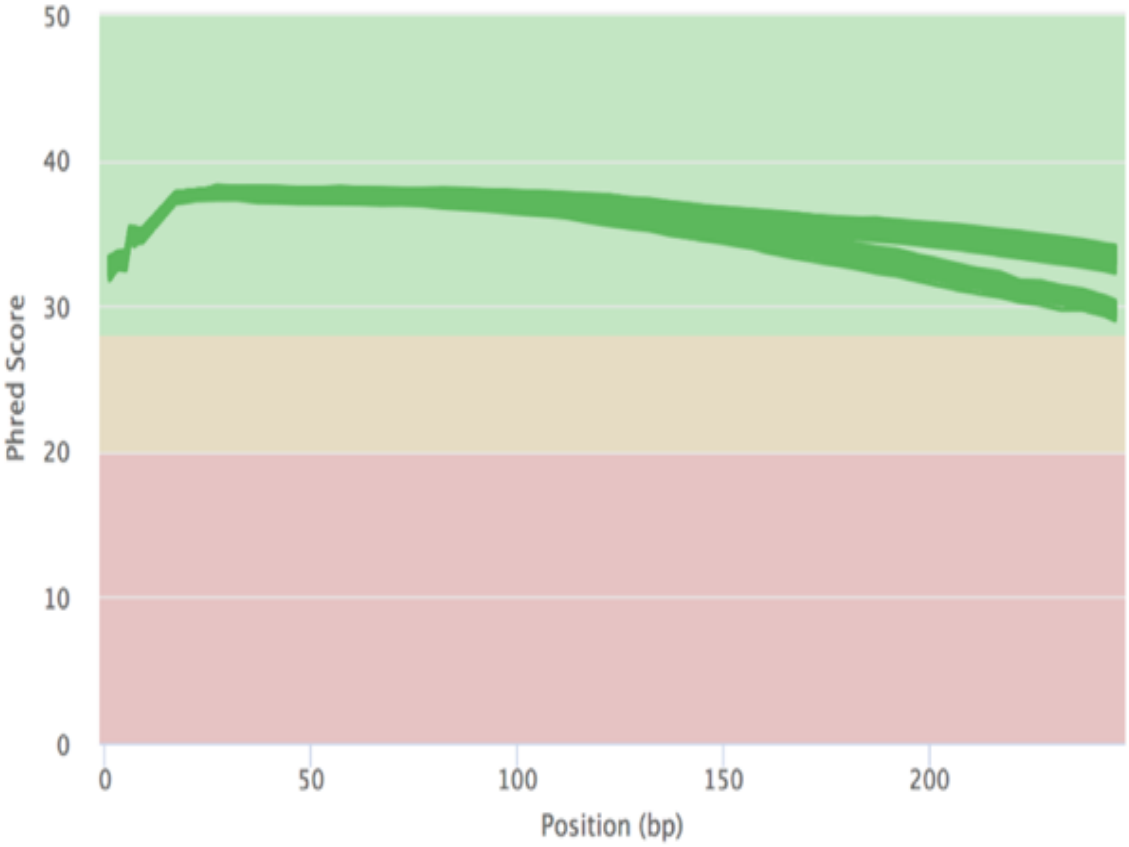


Figure 4: Overall mean quality scores for all the 39 pre-processed data

3.3.4 *de novo* assembly and annotation

ABACAS uses MUMmer to find alignment positions and identify genomic synteny and then rearranges the scaffolds of the query assembly to output a pseudomolecule that approximates the organism's chromosome. One isolate had low read counts and failed to assemble hence it was excluded in analysis. Supplementary 3.1 summarizes the pseudomolecule length of each sample while Supplementary 3.2 summarizes the predicted CDS in assembled genome following annotation.

3.3.5 Gene presence/absence analysis (phylogenomic analysis)

Figure 5 represents clustering analysis of gene content information as expressed by the presence (blue) and absence of genes (red). Here, "gene" means "gene encoding protein (CDS)". Surprisingly, isolate N18 seems to be quite different from the others in the composition of tetra-nucleotides and it has a large number of CDSs that are not present in the other isolates. Therefore, it was recovered as an outlier in this UPGMA dendrogram. Figure 6 represents the same dendrogram without detailed gene content information and *M.*

yongonense 05-1390^T is included as the reference genome. Similarity analysis of the genome by orthoANI suggested that the isolate was *M. yongonense* and not a member of MTBC as presumed. Further analysis of this species is detailed in chapter 5.

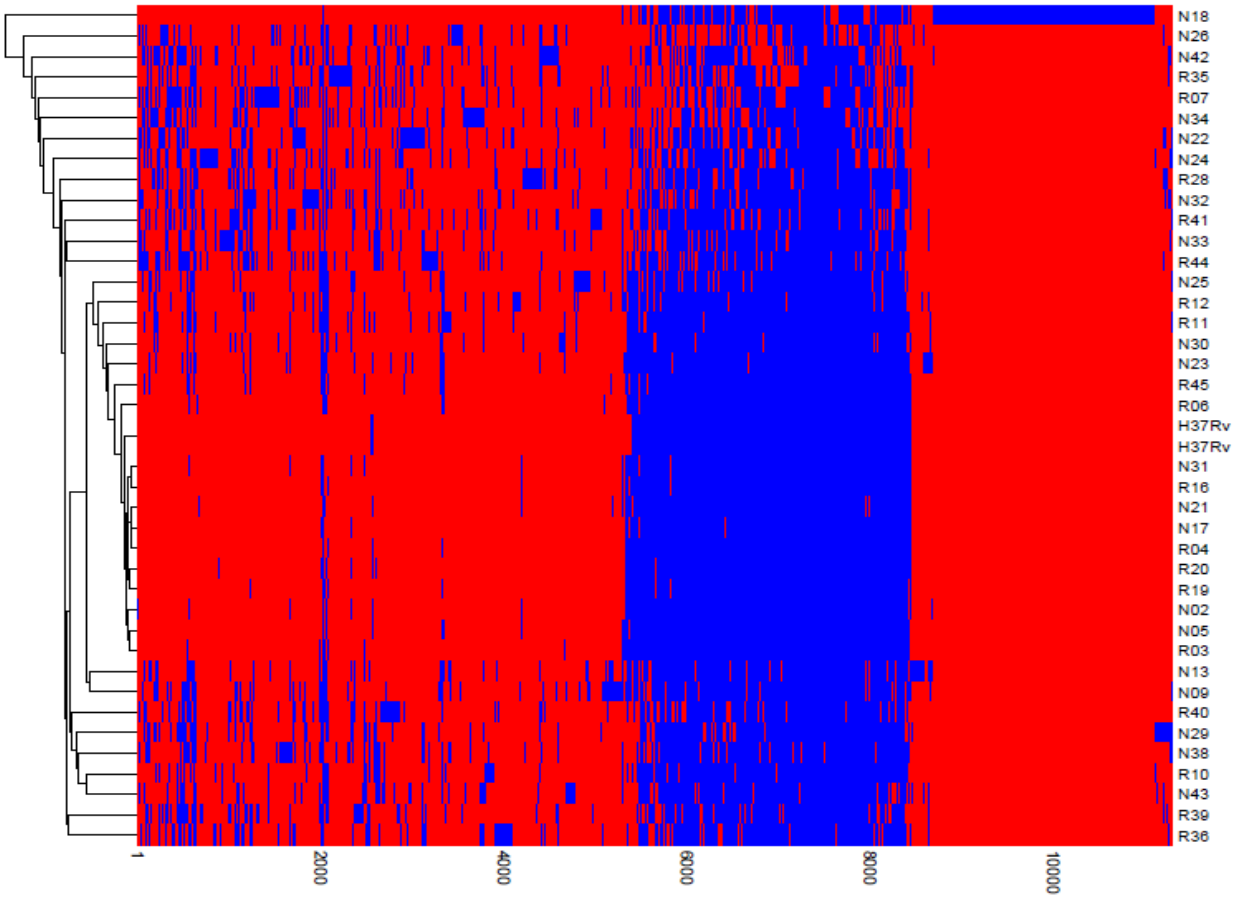


Figure 5: Heatmap showing clustering analysis of the gene content among strains

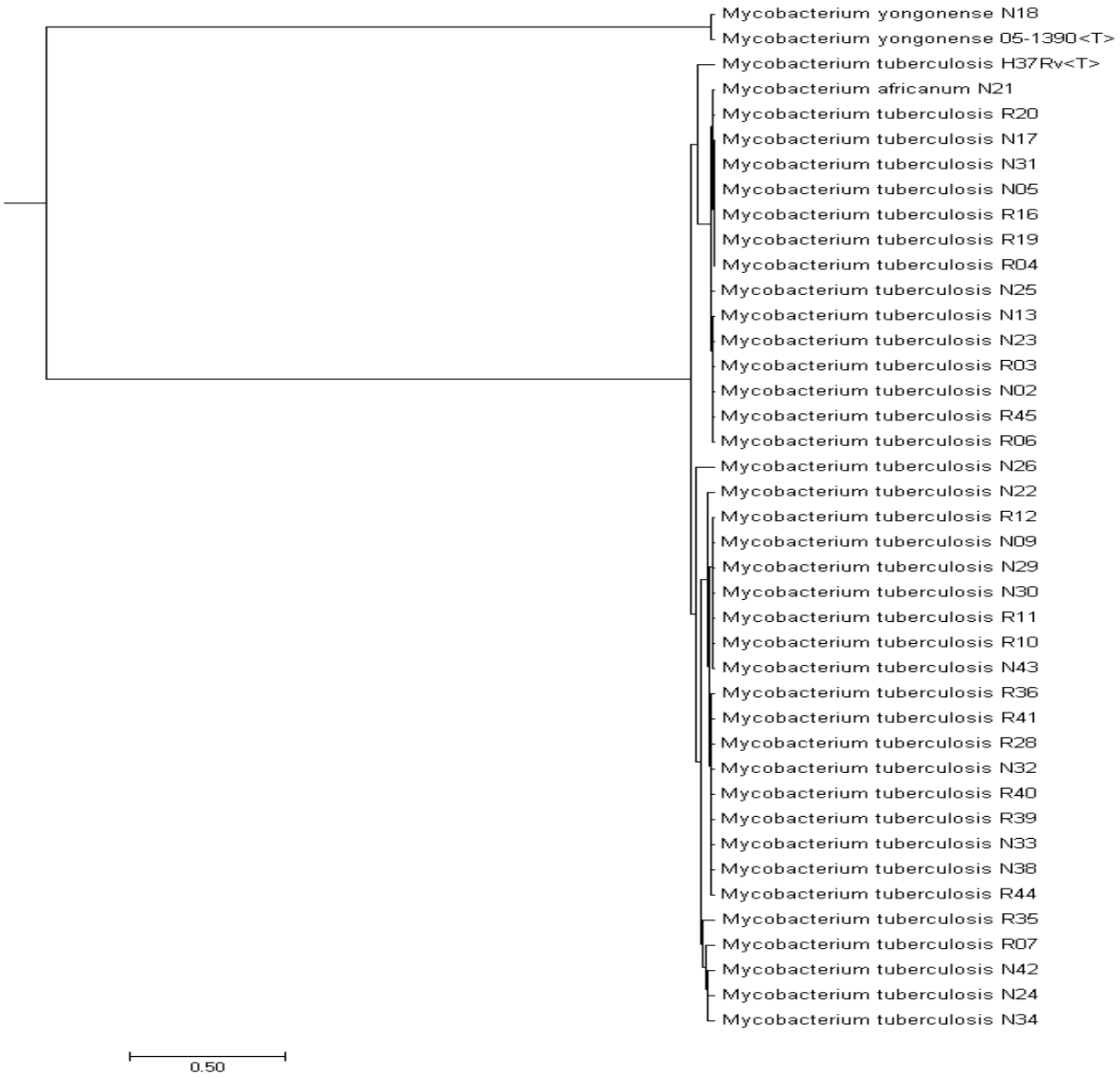


Figure 6: The UPGMA clustering of the strains based on tetra-nucleotides compositions of the 39 isolates.

3.3.6 Stain distribution

Of the 38 isolates, one was *M. yongonense* (N18), one was *M. africanum* (N21) and the remaining were *M. tuberculosis* (Fig. 7). Phylogenetic tree analysis identified a number of stains sharing identical genetic patterns hence forming clusters. Few strains had distant genetic patterns and are hereby referred as “nonclustered strains”.

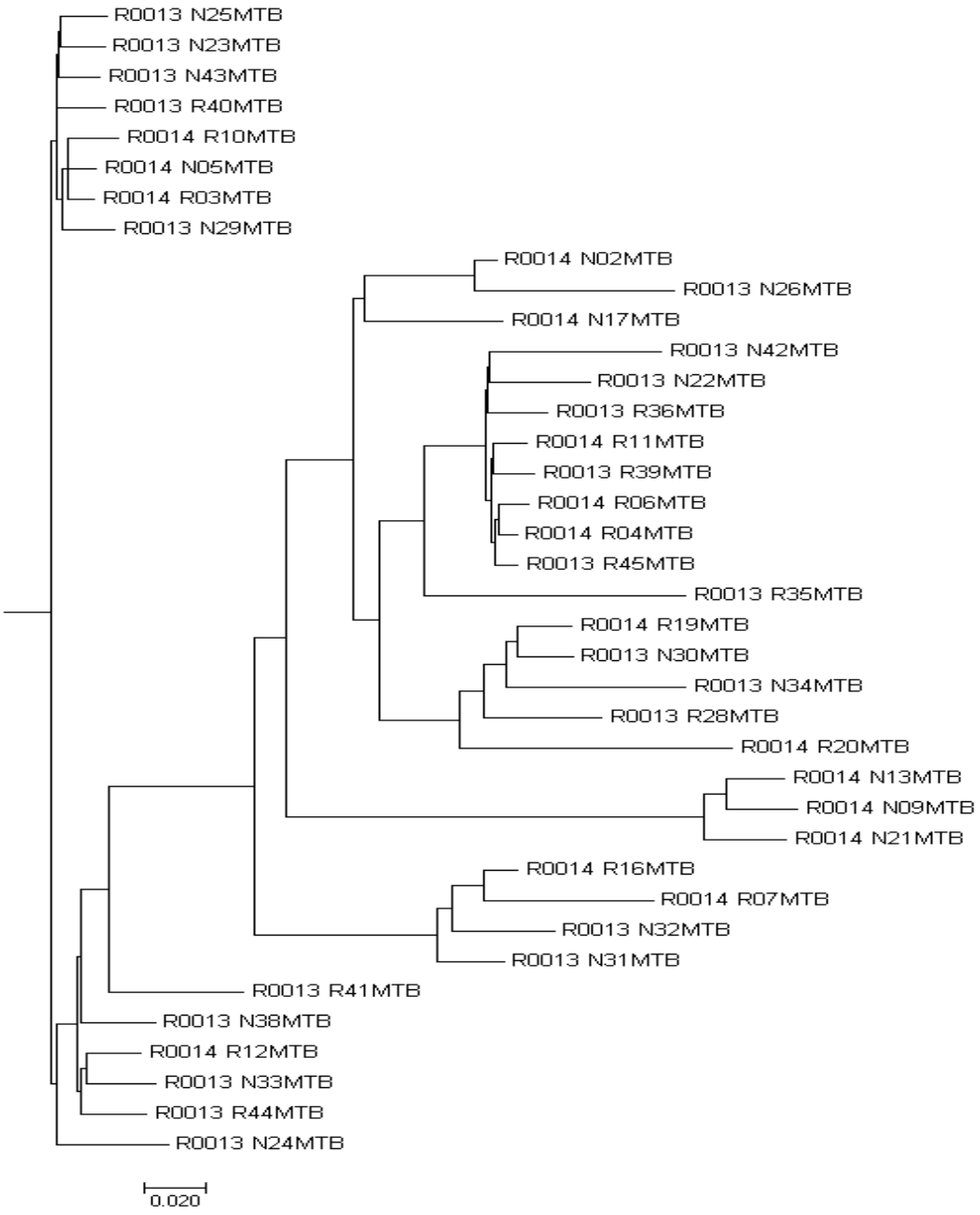


Figure 7: Phylogenetic tree showing the genetic patterns of the 38 MTBC strains

3.3.7 Detection of genetic markers associated with drug-resistant TB

Following variant calling, annotation and variant filtering of 38 genomes to prioritize drug resistant-causative variants, we recovered 10 known drug-resistant genes namely: ethA, gyrA, *inhA, rpoB, embCAB, rpoC, rpsL*, *katG*, *kasA*, *eis*, *pncA and embCAB.* A number of additional variants were also detected in most of these resistant genes. Intergenic regions detected were *Rv1482c-fabG1*, *embC-embA* and *embC-embA.* Additionally, possible

compensatory mutations were observed in other 3 genes, *ndhC, ndhI* and *ndhK* that might confer a selective advantage to drug-resistant strains, not previously associated with resistance.

3.3.8 Distributions of variants

The distribution of variants and the three additional genomic areas will be published together with this manuscript as supplementary materials.

3.4 Discussion

Clinically drug-resistant TB primarily results from the acquisition of point mutations in genes that classically encode drug-metabolizing enzymes or drug-targets within the TB genome (Supply *et al*., 2003). In 2005, Tanzania identified the rise in multi-drug resistant tuberculosis (MDR-TB) (MoHSW, 2009), highlighting the need to characterize drug resistance genetics for improved rapid diagnostic assay and management of the diseases. Whole genome sequencing of 38 clinical isolates from patients who were classified as MDR, recovered 11 of most common genetic markers (some with several additional variants) of drug resistance plus 3 novel genomic regions. Previous studies have reported that several causative mutations remain unknown and, even where such mutations have been identified, additional mutants that nurture drug resistance in MTBC may exist (Farhat *et al*., 2013). Furthermore, literature estimate between 10 and 40% of causative mutations in clinically resistant isolates have not yet been identified and that mycobacterial genotype may be strongly interconnected with a specific human population in a narrow geographical region. The implication of such interconnectedness in the evolution of drug resistance remains unclear.

The most known drug conferring mutations occur in *rpoB* and *rpoC* for rifampicin; *kasA, katG, inhA* and *ndh* for isoniazid; *rpsL* and *rrs* for streptomycin; *embCAB* for ethambutol; *ethA* and *inhA* for ethionamide; *pncA* for pyrazinamide; *gyrA* for fluoroquinolone; and *eis* for amikacin and kanamycin (Campbell *et al*., 2011; Müller *et al*., 2013); and intergenic regions: *embC–embA, proA–ahpC, eis–Rv2417c* and *Rv1482c–fabG1* (Almeida Da Silva and Palomino, 2011; Ramaswamy and Musser, 1998; Sekiguchi *et al*., 2007). None of the intergenic regions were detected in this study. Sequence analysis recovered the hot-spot regions of *ethA, gyrA, inhA, embB rpoB, rpoC, rpsL, katG, kasA, eis* and *pncA*. Detection of

mutations in *ndhC, ndhI* and *ndhK* might have compensatory effects in the evolution of MDR-TB strains a in the face of drug selection.

The findings presented in this study also suggest that resistance-conferring mutations were not always associated with resistance phenotypes. For example, DST results demonstrated that 36/38 isolates were resistant rifampicin and no genetic markers were detected in *embCAB* in 4/19 resistant isolates. However, all four Ethambutol isolates had *katG* mutations. Mutations in *katG* have been reported to cause cross-resistance and are commonly found in MDR-isolates (Hazbón *et al*., 2006b). Similarly, one streptomycin resistant isolate had no detectable mutation in both *rr* and *rpsL* and three resistant isolated had no detectable mutation in both *rr* and *rpsL*. Streptomycin-susceptible strains have been previously described to harbour similar mutations. These findings are in accord with previous evidence that suggests that about 30% of clinical streptomycin-resistant TB present no mutation in either of these genes (Perdigão *et al*., 2014; Spies *et al*., 2011b) and mutations in gene *gidB* may cause low-level resistance in some clinical streptomycin-resistant strains that displayed no mutation in *rrs* and *rpsL* (Spies *et al*., 2011b).

Phylogenetic analyses suggest a shared DNA fingerprint (forming clusters), indicative of recent transmission and rapid progression to active TB. However, these results remain suggestive rather than conclusive because of the nature of the study that used convenient samples (hospital-based samples) and small sample size hence hindering the analysis of clustering/transmission tree. Therefore, population genetics studies are warranted to provide enough evidence and inform better control strategies. Population genetics studies have demonstrated that patients with isolates of *M. tuberculosis* that share the same DNA fingerprint (clustered cases) are part of a chain of transmission and therefore, developed the disease as a consequence of recent transmission and rapid evolution to active disease (Lockman *et al*., 2001; Small *et al*., 1994b). One study in San Francisco has shown that the presence of a substantial reduction in the incidence of tuberculosis, there has been a disproportionate decrease in the amount of disease that is attributable to recent transmission (clustered cases) (Lockman *et al*., 2001). Clustered cases represent failures of the tuberculosis control system, and an analysis of these failures can be used to guide control efforts. With targeted interventions, these programmatic failures can be avoided and tuberculosis caused by recent transmission of *M. tuberculosis* can be greatly reduced (Lockman *et al*., 2001).

3.5 Conclusion

This data revealed additional genomic areas and variants which might have positive effects on the survival of the pathogen under drug selection. Clustering of strains was observed among the isolates which is the evidence of shared same DNA fingerprints hence high possibility of recent transmission. More localized population genetics studies are needed to accurately define the spectrum of drug resistance, strain clustering and transmission dynamics of TB/MDR-TB in Tanzania so as to inform better diagnostic and disease management strategies.

Ethical Consideration: The study protocol received ethical approval from the Medical Research Coordinating Committee under Nation Institute for Medical Research (Tanzania). Sample handling followed the 4th WHO report on "the guidelines for surveillance of drug resistance in tuberculosis" and "the International Conference on Harmonization guideline for Good Clinical Practice E6".

Authors' contributions: NPM, DJK, SM and RK conceived the study idea and co-designed the protocol. DJK, SM, AK, RK, and EN co-supervised the study. NPM performed a culture of the isolates, prepared DNA isolates, performed whole-genome sequencing, analysed the sequence data and drafted the paper. SGM, ESN, RK, DJK, and AK revised the paper.

Funding: This work was mainly supported by the Government of Tanzania through NM-AIST and partial support from the DELTAS Africa Initiative [Afrique One-ASPIRE /DEL-15-008]. Afrique One-ASPIRE is funded by a consortium of donors including the Alliance for Accelerating Excellence in Science in Africa (AESA), African Academy of Sciences (AAS), the New Partnership for Africa's Development Planning and Coordinating (NEPAD) Agency, the Wellcome Trust [107753/A/15/Z] and the UK government.

CHAPTER FOUR

Genome sequence of *Mycobacterium yongonense* RT 955-2015 isolate from a patient misdiagnosed with multidrug-resistant tuberculosis: First clinical detection in Tanzania[2]

Nicholaus Peter Mnyambwa[a.b*], Dong-Jin Kim[a], Esther Ngadaya[b], Jongsik Chun[c.d], Sung-min Ha[c.d], Pammla Petrucka[a.e], Kennedy Kwasi Addo[f], Rudovick R. Kazwala[g], Sayoki G. Mfinanga[b]

[a]School of Life Science and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania. [b]National Institute for Medical Research Muhimbili Medical Research Centre, Dar es Salaam, Tanzania. [c]School of Biological Sciences, Seoul National University, Seoul, Republic of Korea. [d]ChunLab, Inc., Nambusunhwan-ro, Seocho-gu, Seoul, Republic of Korea. [e]University of Saskatchewan, Saskatoon, Canada. [f]Bacteriology department, Noguchi Memorial Institute for Medical Research, Accra, Ghana. [g]Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania

[*]Corresponding author: Nicholaus Peter Mnyambwa (lodnicho@gmail.com)

[2] International Journal of Infectious Diseases 2018:71 82–8.

Abstract

Background: *Mycobacterium yongonense* is a recently described novel species belonging to *Mycobacterium avium* complex, which is the most prevalent aetiology of non-tuberculous mycobacteria associated with pulmonary infections, and poses tuberculosis diagnostic challenges in high-burden, resource-constrained settings.

Methods: Whole genome shotgun sequencing and comparative microbial genomic analyses were used to characterize the isolate from a patient diagnosed with multidrug-resistant tuberculosis (MDR-TB) after relapse.

Results: Sequence analysis revealed that the RT 955-2015 strain had a high similarity to *M. yongonense* 05-1390(T) (98.74%) and *Mycobacterium chimaera* DSM 44623(T) (98%). Its 16S rRNA showed similarity to *Mycobacterium paraintracellulare* KCTC 290849(T) (100%), *Mycobacterium intracellulare* ATCC 13950(T) (100%), *M. chimaera* DSM 44623(T) (99.9%), and *M. yongonense* 05-1390(T) (98%). The strain exhibited a substantially different rpoB sequence to that of *M. yongonense* 05-1390 (95.16%), but closely related to that of *M. chimaera* DSM 44623(T) (99.86%), *M. intracellulare* ATCC 13950(T), (99.53%), and *M. paraintracellulare* KCTC 290849(T) (99.53%).

Conclusion: In light of the OrthoANI algorithm and phylogenetic analysis, it was concluded that the isolate was *M. yongonense* Type II genotype, which is an indication that the patient was misdiagnosed with TB/MDR-TB and received inappropriate treatment.

Keywords: *Mycobacterium yongonense; Mycobacterium avium complex* (MAC); multi-drug resistant tuberculosis (MDR-TB); nontuberculous mycobacteria (NTM); Tanzania

4.1 Introduction

Mycobacterium yongonense is a recently reported species from human specimens in South Korea (Kim *et al*., 2013a, b), with the *16S rRNA* gene showing a high degree of similarities to members of the *Mycobacterium avium* complex (MAC) namely: *M. marseillense* (100%), *M. chimaera* (99.9%), and *M. intracellulare* (99.8%) with a distinct *rpoB* gene sequence that resemble that of *M. parascrofulaceum* (Kim *et al*., 2013a). Other members of MAC include *M. timonense, M. bouchedurhonense* (Salah *et al*., 2009), *M. vulneri* and *M. colombiense* (Tortoli *et al*., 2004), as well as *M. arosiense* (van Ingen *et al*., 2009). MAC is the most common group of nontuberculous mycobacteria (NTM), that frequently associated with pulmonary infections in humans, particularly in immunocompromised persons (Gordin e*t al*., 1997). A complete genome sequence of *M. yongonense* that was isolated in Korea have been published, with an estimate of a circular DNA of 5.5 Mb and 5 222 protein-coding genes CDSs (Kim *et al*., 2013b). Subsequently, other two clinical cases of *M. yongonense* were reported from Italian patients with lung infections (Enrico *et al*., 2013). Both strains had *rpoB* gene showing high similarity to that of *M. intracellulare* (99.4%), but not *M. parascrofulaceum*.

Recent study (Kim *et al*., 2016) has suggested two distinct genotypes of *M. yongonense* based on the origin of the *rpoB* gene: Type I genotype with *rpoB* gene acquired through Horizontal Gene Transfer (HGT) event from *M. parascrofulaceum* and Type II genotype with the *M. intracellulare rpoB* gene (without HGT events). More recent publication by the same authors provided additional evidence that the entire *M. yongonense* Type I *rpoBC* operon resulted from a distantly related species of *M. parascrofulaceum* (Kim *et al*., 2017). The authors also revealed that members of *M. yongonense* Type I genotype harbor a unique DNA mismatch repair gene *MutS4* family, that potentially serves as a putative driving force for the suggested Horizontal gene transfer between the *M. parascrofulaceum* and *M. yongonense* Type I genomes through homologous recombination events (Kim *et al*., 2017). The occurrence of NTM in Tanzania has been frequently reported (Hoza *et al*., 2016; Kilale *et al*., 2016; Mfinanga *et al*., 2014; Mnyambwa *et al*., 2017b), but, characterization of the isolates has always relied on suboptimal typing methods that hinder accurate species identification.

In 2008, the WHO endorsed the molecular tests known as Line Probe Assays and more recently the GeneXpert for rapid detection of clinically resistant TB. MTBDR*plus* is one of the line probe assays which have been endorsed by the WHO for rapid detection (within 24

hrs) of MDR-TB (WHO, 2013). MTBDR*plus* can detect drug resistance-conferring mutations for rifampicin and isoniazid. The MDR-TB treatment guidelines in Tanzania recommends the use of MTBDR*plus* and a patient should be placed in MDR-TB treatment if there is proof of rifampicin resistance either by Line Probe Assay, GeneXpert or conventional culture-based DST (proportional) methods (MoHSW, 2012). Here, we describe a whole genome sequence of *M. yongonense* isolate RT 955-2015, which is the first case of infection to be reported in Tanzania. The organism was isolated in sputum specimen of a patient who was diagnosed with MDR-TB by GenoType MTBDR*plus* (Hain Lifescience GmbH, Nehren, Germany) after TB relapse.

4.2 Materials and methods

4.2.1 Culture and drug susceptibility testing

The GenoType MTBDR*plus* was performed in accordance with manufacturer's instructions. Culture and DST were performed in a contained laboratory Biosafety Level 3 laboratory at the CTRL, in Dar es Salaam, Tanzania. The bacterial isolates were inoculated onto the Lowenstein-Jensen media slants and incubated at 35-37ºC and growth was observed in the sixth week. All drug sensitivity assays were performed in triplicate and susceptibility to a particular drug was determined by the standard proportion method as recommended by WHO 2009.

4.2.2 Genomic DNA extraction

The genomic DNA was extracted from heat-killed pathogen by the Cetyltrimethyl Ammonium Bromide (CTAB) method as described in detail elsewhere (van Soolingen *et al*., 1991). In summary, bacterial isolates were grown on Lowenstein-Jensen media slants at CTRL. Using a sterile loop, viable colonies were lifted from the LJ slants and emulsified into tubes containing 1X Tris-EDTA buffer, pH 8.0, then heat-killed in waterbath at 80°C for 20 minutes. This was followed by addition of 10 mg/mL lysozyme in each tube and incubation at 37°C overnight. In the following day, the DNA was then extracted with chloroform-isoamyl alcohol (24:1) and the pellets of genomic DNA were rehydrated in 80 µL TE and left overnight at 4°C. The gDNA was then temporarily stored at -20°C before shipping to KCRI Moshi, Tanzania for sequencing procedure.

4.2.3 Whole genome sequencing

Dual-indexed (24) DNA libraries preparation were constructed by shearing genomic DNA ~ 300 bp using NexteraXT DNA sample preparation kit according to the manufacturer's instructions (Illumina Inc., San Diego, CA, USA). Genomic DNA libraries were sequenced using a paired-end 2x250 bp protocol (MiSeq Illumina) at KCRI.

4.2.4 Genome assembly, annotation, and identification

The pre-processed sequencing reads were assembled by means of SPAdes 3.9.1, using default parameters to generate contigs and scaffolds. Scaffold-level assembly was evaluated using QUAST with the reference genome (*M. yongonense* 05-1390 (T) to generate summary statistics, for example, N50, maximum scaffold length, G+C content, and others. The assembled scaffold sequences were ordered using ABACAS (Assefa *et al*., 2009) with *M. yongonense* 05-1390 as a reference genome. Gene prediction was performed using Prodigal and annotation was conducted using a homology search against the Clusters of Orthologous Groups (COG), EggNOG, SEED systems, Swiss-Prot and KEGG databases. A species-level identification was performed using OrthoANI and *16S rRNA* with 98% and 97% as a threshold value, respectively (Chun's Bioinformatics Lab [ChunLab]).

Gene content (presence/absence) data was used to generate a Heatmap (ChunLab). The construction of the Heatmap utilizes tetra-nucleotides of CDSs through the whole genome (pan-genome). Jaccard coefficient was used to calculate similarities between genomes based on presence/absence data. Sequence alignment and syntenic assessment were performed using Mugsy and an approximately-maximum-likelihood phylogenetic tree (UBCG: Up-to-date Bacterial Core Genome-based tree) was reconstructed using FastTree programme. The phylogenetic tree was derived from the alignment of pseudomolecule with reference sequences. *Mycobacterium tuberculosis* H37Rv was included as a reference genome representing members of MTBC while *M. intracellulare* ATCC 13950 and *M. parascrofulaceum* were included to provide additional evidence for the classification of the isolate into a specific genotype (Type I or Type II genotype). Additionally, blast analysis against the NCBI refseq database was performed.

4.2.5 Multilocus sequence typing

Multilocus sequence typing (MLST) were analyzed using 10 public *M. yongonense* genomes from the EzBioCloud database (ChunLab). The Kruskal's algorithm, based on the allelic distance matrix of the shared loci, was used to construct a minimum spanning tree.

4.2.6 Quality control

The quality and quantity of gDNA were confirmed using Qubit 2.0 fluorometer (Thermal Fisher Scientific, Waltham, MA USA). FastQC and K-mers spectra analysis by K-mer Analysis Toolkit (KAT 2.3.2) were used to assess the quality of the raw sequence data. Trimmomatic Version 0.36 and FastX ToolKit (Hannon Lab) were used to remove Illumina adapter sequences and filter out low quality sequenced bases and reads. The *de novo* genome assembly was validated against a complete genome sequence of *M. yongonense* 05-1390(T). Genome contamination was checked using the ContEst16s tool and coverage was checked based on bacterial core genes.

4.3 Results

4.3.1 Demographic and clinical characteristics of the patient

The patient was male 60-year-old and HIV negative living in Dar es Salaam, Tanzania who was diagnosed with TB relapse by smear-examination and mycobacterial culture (2015). Additionally, drug-conferring mutations for rifampicin and isoniazid were detected using a rapid GenoType MTBDR*plus* VER 2.0 test. Subsequently, the conventional phenotypic culture-based DST (proportion method) on LJ media was performed and the isolates were susceptible to all first-line anti-TB drugs. Based on the MTBDR*plus* test, the patient was classified as having MDR-TB, hence placed into the MDR-TB treatment regimen. The patient converted to sputum smear and culture negative status in the third month of MDR-TB treatment and thereby considered cured on completion of the treatment.

4.3.2 Sequence analysis

MiSeq sequencing generated 400 000 reads, which passed quality checks with an average of ~28X depth coverage. The genome was free from contamination. The *de novo* assembly generated 202 scaffolds containing (402 contigs). Our findings revealed 5 528 170 bp genome size with 99. 61% degree of genome coverage. The *M. yongonense* RT 955-2015 genome has 5331 open reading frames (ORFs), 46 tRNA genes, 4 rRNA genes and 1

tmRNA. We obtained a mean length of CDS (SD) 913.8 (652.3) bp and 807 bp median of CDS length. The mean length of IGR (SD) was 137.7 (194.3) bp. The genome of *M. yongonense* RT 955-2015 has a G+C content of 67.96%. A comparison of predicted ORFs of *M. yongonense* RT 955-2015 and that of *M. yongonense* 05-1390(T) showed that the two genomes share 4763 ORFs with an average genome identity of 98.74%. A large number of unshared genes among *M. yongonense* genomes were observed. Comparative functional analysis of genes showed that about one-third of genes were assigned unknown function and the majority of remaining genes were assigned to transcription (6.7%), lipid transport and metabolism (6.4%), energy conversion and production (6.0%), secondary metabolite biosynthesis, transport and catabolism (5.5%), replication, and amino acid transport and metabolism (see Fig. 8).

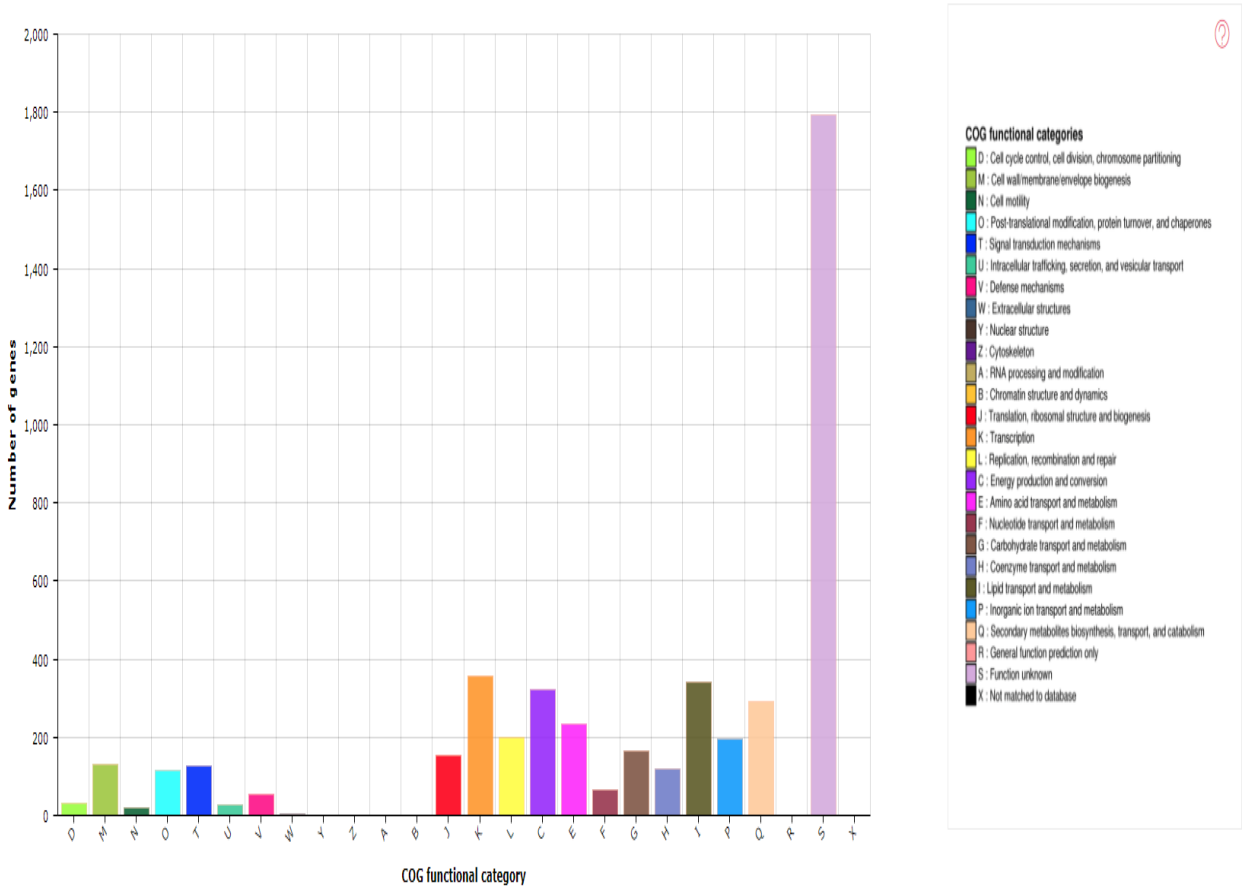


Figure 8: Distribution of CDSs based on function prediction

4.3.3 Species identification

Genome similarity analysis of the RT 955-2015 isolate has shown the highest similarity to *M. yongonense* 05-1390 (T), while both *16S rRNA* and *rpoB* showing high similarity to *M. intracellulare* ATCC 13950 and its related members (*M. paraintracellulare* KCTC 290849 and *M. chimaera* DSM 44623) (Table 3).

Table 3: Similarity values generated by the OrthoANI showing the similarity of isolate *RT* 955-2015 to the closest members of MAC.

Taxonomy name	**Genome identity (%)**	***16S rRNA* similarity (%)**	***hsp65* similarity (%)**	***rpoB* similarity (%)**	**ITS1 similarity (%)**
M. yongonense 05-1390(T)	98.74	99.8	99.12	95.16	98.23
M. chimaera DSM 44623(T)	98.00	99.9	99.26	99.86	94.00
M. paraintracellulare KCTC 290849(T)	97.50	100	99.26	99.53	99.29
M. intracellulare ATCC 13950(T)	97.50	100	99.45	99.53	99.29
M. marseillense DSM 43457(T)	92.10	99.8	98.22	97.24	98.27
M. parascrofulaceum ATCC BAA-614(T)	83.83	97.89	94.46	95.07	89.36

Gene content analysis showing the evolutionary relationship of the isolate suggests that the RT 955-2015 isolate shares a recent ancestry with *M. yongonense* species. Clustering analysis of gene encoding protein (CDS) content information as expressed by the presence (blue) and absence (red) of genes is reflected in Fig. 2 (A). The RT 955-2015 strain is closely related to *M. yongonense* species (I and II) and quite different from *M. tuberculosis* H37Rv(T) in the composition of tetranucleotides. Fig. 2(B) represents the same dendrogram without detailed gene content information. The whole genome-based phylogenetic tree also supports the findings that the RT 955-2015 strain is a member of *M. yongonense* (Fig. 9) and in reference to previous n this genome-based phylogenetic tree, group I and II reflects the previously classification into two distinct genotypes (*M. yongonense* Type I genotype and *M. yongonense* Type II genotype). BLAST analysis against the NCBI refseq database of all hits and top hits suggested that the isolate RT 955-2015 was *M. yongonense.*

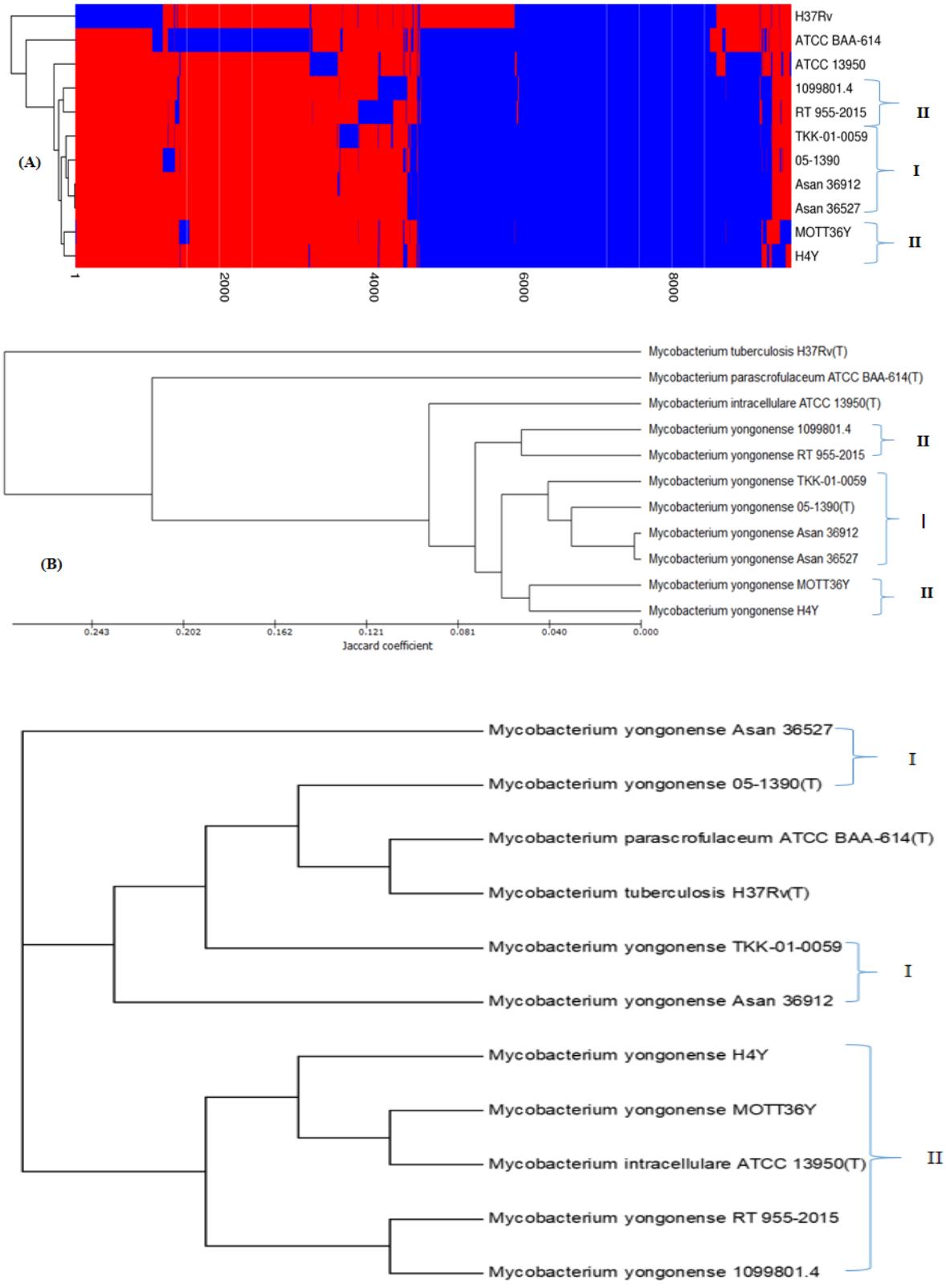


Figure 9: Phylogeny based on presence/absence of genes among genomes. (A) Heat¬map showing clustering analysis of the gene content among genomes. (B) The UPGMA clustering of the strains based on tetra-nucleotides compositions.

4.3.4 Mutations

A number of single nucleotide polymorphism (SNPs) were detected in six genes, namely: - *rboB, rpoC, kasA, katG, inhA,* and *pnacA*. More polymorphisms (including nonsynonymous mutations) were common in the *rboB* gene.

4.3.5 MLST of *M. yongonense*

A total of 4 019 core genes were retrieved from 10 public *M. yongonense* genomes and the RT 955-2015 isolate for strain typing and a minimum spanning tree was generated (Fig. 10). Of the 10 genomes retrieved from public databases, 9 genomes are reclassified as *M. yongonense* in the EzBioCloud database. The profile was then used to search the core genes in each genome including RT 955-2015 strain. The colored circles in the tree represent the country where a strain was isolated and the distance between the nodes indicates the genetic distance of the isolates. The shorter the distance the closer the strains are; thus, RT 955-2015 strain seems to be very close to *M. yongonense* 1099801.4 strain.

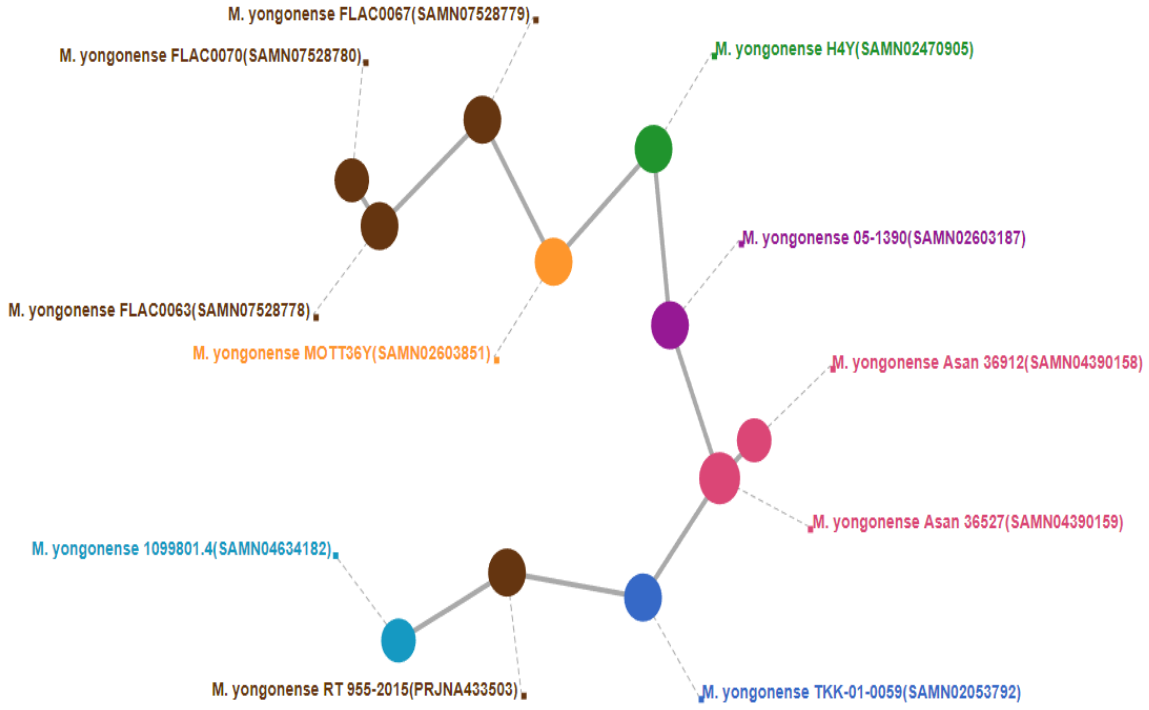


Figure 10: Minimum spanning tree based on MLST showing genetic distances among 11 strains *of M. yongonense* species. BioSample accession code of each strain is provided in parenthesis**).**

4.4 Discussion

This study presents the whole genome sequence of the clinical isolate which is hereby confirmed as *M. yongonense* RT 955-2015, *a* first detected case in Tanzania. Our findings are based on an orthoANI algorithm, gene content (presence/absence) and complete genome sequence-based phylogenetic analyses, which further suggest that the *M. yongonense* RT 955-2015 belongs to Type II genotype. Interesting, heatmap and UPGMA dendrogram based on the gene content (gene presence/absence) showed similar strain clustering pattern to that of the complete genome sequence-based phylogenetic (approximately-maximum likelihood) analysis. The isolate identification based on similarity analysis of *16S rRNA* gene sequence alone was not possible due to its high degree of similarity (>99%) with other members of MAC. Previous findings have demonstrated that the *16S rRNA* gene has no good resolution power above 98% similarity (Lee *et al*., 2016).

The substantial difference between *M. yongonense* RT 955-2015 *rpoB* sequence and that of *M. parascrofulaceum* and *M. yongonense* 05-1390, while showing high similarity to *M. intracellulare* and its closely related species (*M. chimaera* DSM 44623, *M. paraintracellulare* KCTC 290849), is an additional evidence that RT 955-2015 stain belongs to Type II genotype group. Kim and colleagues use genome-based phylogenetic analysis to classify *M. yongonense species* into two distinct genotypes (Type I genotype and Type II genotype) (Kim *et al*., 2016), and demonstrated that Type I genotype has laterally acquired *rpoB* from *M. parascrofulaceum* while Type II genotype processing the *M. intracellulare rpoB* gene (Kim *et al*., 2016, 2017,). Kim and colleague, classified *M. yongonense* 05-1390, *M. yongonense* Asan 36912, *M. yongonense* Asan 36527 as members of Type I genotype while *M. yongonense* MOTT36Y and *M. yongonense* H4Y were classified under the Type II genotype (Kim *et al*., 2016): our findings on these strains are consistent with their classification. Furthermore, our findings suggest that *M. yongonense* TKK-01-0059 (formerly known as *Mycobacterium* sp. TKK-01-0059) from South Africa, belongs to *M. yongonense* Type I genotype. The pronounced difference in a number of unshared genes among the *M. yongonense* genomes highlights the possibility of extensive HGT from outside the cell in these genomes. Evidence of HGT among NTM members has been previously provided (Fedrizzi *et al*., 2017; Kim *et al*., 2013a; Kim *et al*., 2016, 2017).

Mulitilocus sequence typing demonstrated that *M. yongonense* RT 955-2015 strain was more closely related to *M. yongonense* 1099801.4, which was formerly classified as a member of *M. intracellulare*, sampled in Mozambique in 2012. The strain is placed in the same group with *M. yongonense* RT 955-2015 (Type II genotype) based on pan-genome and phylogenetic analysis.

In this study, *M. yongonense* TKK-01-0059 (formerly *Mycobacterium sp*. TKK-01-0059) from South Africa (2009) was also placed in Type I genotype. A similar classification has been proposed previously (Kim *et al*., 2017) based on the presence of *MutS4* and supported by *hsp65* and genome sequence-based phylogenetic analyses. The first three stains in the top of the spanning tree were isolated in the United States of America, and the remaining 5 isolates were isolated from South Korea. The clustering of these isolates in the minimum spanning tree based on the three geographical regions where they were isolated (Africa, Korea, and the United State of America) may infer an important evolutionary relationship among the strains.

Mycobacterium yongonense is a recently described species (Kim *et al*., 2013a, b) of NTM belonging to the MAC members that are associated with pulmonary infections in humans. The steady increase in health complexities and mortality rate caused by several NTM species have recently been recognized (Yeung *et al*., 2016) and are raising a public health concern. As there was no evidence of genome contamination, the involvement *M. yongonense* RT 955-2015 as the causative agent of disease and that the patient was misdiagnosed with TB/MDR-TB seems incontrovertible. The epidemiology and clinical implications of *M. yongonense* remains a mystery as only a few cases have been described. This can be attributed to the use of suboptimal diagnostic tests, particularly, in developing countries like Tanzania. Failure to detect NTMs due to use of suboptimal diagnostic tests may lead to false positive results (Maiga *et al*., 2012), which normally do not respond to anti-TB drugs hence are easily classified as MDR-TB (Lu *et al*., 2016; Shahraki *et al*., 2015). Most of the MAC members are naturally resistant to anti-TB drugs (Mpagama *et al*., 2013; Shahraki *et al*., 2015), which leads to the unnecessary administration of toxic second-line anti-TB treatments, thereby increasing risks for the evolution of drug-resistance.

It is estimated that between 7% and 12% of the adult population has been previously infected with MAC, but the rate of the disease varies significantly by geographic region (Benson, 1994; Gordin *et al*., 1997). MAC disease is typically prevalent in patients with compromised

immune systems (Gordin *et al*., 1997). A high incidence rate, between 20 and 40%, has been reported in patients with severely low CD4 cell count associated with HIV infections, especially when there is an ineffective antiretroviral treatment or chemoprophylaxis (Gordin *et al*., 1997). MAC infections have been reported to cause serious disease in HIV patients globally (Gordin *et al*., 1997; Kwak *et al*., 2017), and evidence of such mycobacterial infections have been reported in Tanzania (Crump *et al*., 2009). Patients with misdiagnosed infections normally receive incomplete therapy which contributes to repeated visits to health facilities, potential increases in mortality and the development of drug-resistance. Misdiagnosis can also lead to misallocation of resources for health care and disease control. Thus, proper tests are needed for an effective health care system, providing critical information for the right medical decisions and management of a particular health condition. Such diagnostic tests should also help to address epidemiological concerns regarding the *M. yongonense* species and other underreported members of NTM.

It is well-known that the detection of acid-fast in NTM can result in misdiagnosis of TB cases by smear examination. Although the GenoType MTBDR*plus* assay is designed to detect common resistance mutations of *rpoB* gene for rifampicin and *katG* and *inhA* for isoniazid in MTB complex strains (Lacoma *et al*., 2008), sequence analysis of *M. yongonese* RT 955-2015 revealed a catalog of nonsynonymous mutations in same genes and in three other genes namely:- *rpoC, kasA,* and *pnacA.* Pyrazinamide resistance-conferring mutations are known to occur in gene *pnacA,* while mutations in gene *rpoC* are associated with drug resistance to rifampicin drug in MTBC (Mnyambwa *et al*., 2017a). Although we provide evidence of genetic markers in genes that are known to cause drug resistance in MTBC strains, it remains a mystery whether these genetic signatures play the same role in *M. yongonense*.

In conclusion, sequence analysis suggests that the isolate RT 955-2015 was *M. yongonense* Type II genotype, a member of MAC and is the first detected case in Tanzania. These findings strongly affirm the hypothesis that the condition was misdiagnosed as TB/MDR-TB and that the patient received inappropriate treatment. This finding highlights the need for accurate diagnostic assays to ensure effective use of drugs and minimize risks for the development of drug resistance and worsening patient's condition.

This whole genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number PSQD00000000 and assigned name as *M. yongonense* RT 955-2015. The version described in this paper is version PSQD01000000.

Conflict of interest: The authors declare that there is no conflict of interest.

Ethical consideration: This study was part of the research study aiming at investigating the genetic basis of drug-resistant pattern in *Mycobacterium tuberculosis* complex in clinical isolates from tuberculosis patients in Tanzania, which was approved by the National Institute for Medical Research, Tanzania.

Consent for publication: Not applicable

Availability of data and material: All important datasets used and/or analyzed in this study are available at https://www.bioiplug.com/genome/YongonenseRT9552015Set

Authors' contributions: NPM, DJK, SM, AKK, and RK conceived the study idea and co-designed the protocol. D-JK, SM, RK, AKK, and EN co-supervised the study. NPM performed a culture of the isolates, prepared DNA isolates, and performed genome sequencing. NPM, DH, and JC analyzed the sequence data and drafted the paper. SGM, ESN, RK, D-JK, and PP revised the paper. All authors approved the final version.

Funding: This work was mainly supported by the Government of Tanzania through Nelson Mandela African Institution of Science and Technology and partially supported by the DELTAS Africa Initiative [Afrique One-ASPIRE /DEL-15-008]. Afrique One-ASPIRE is funded by a consortium of donor including the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA), the New Partnership for Africa's Development Planning and Coordinating (NEPAD) Agency, the Wellcome Trust [107753/A/15/Z] and the UK government. The sponsors had no involvement in the study design, data collection nor in interpretation and the views expressed in this publication those of the authors and not necessarily those of Afrique One-ASPIRE or Government of Tanzania

Acknowledgement: The authors extend sincere appreciations to Happiness Kumburu (Kilimanjaro Clinical Research Institute), Daphne Mtunga and Salim Bossy (CTRL) for their support during laboratory work, and Edson Ishengoma (Mkwawa University College of Education) for his technical assistance during sequence analysis.

CHAPTER FIVE

Assessment of sputum smear-positive but culture-negative among newly diagnosed pulmonary tuberculosis patients in Tanzania[3]

Nicholaus Peter Mnyambwa[1,2], Esther S Ngadaya[2], Kimaro Godfather[2], Dong-Jin Kim[1], Rudovick Kazwala[3], Pammla Petrucka[1,4], Sayoki G Mfinanga[2]

[1]School of Life Sciences and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania. [2]National Institute for Medical Research, Muhimbili Medical Research Centre, Dar es Salaam, Tanzania. [3]Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania, [4]College of Nursing, University of Saskatchewan, Saskatoon, Canada.

Correspondence: Nicholaus Peter Mnyambwa (lodnicho@gmail.com)

Abstract

Diagnosis of pulmonary tuberculosis (TB) in technology-limited countries is widely achieved by smear microscopy, which has limited sensitivity and specificity. The frequency and clinical implication of smear-positive but culture-negative among presumptive TB patients remain unclear. A cross-sectional substudy was conducted which aimed to identify the proportion of nontuberculous mycobacteria (NTM) infections among 94 "smear-positive culture-negative" patients diagnosed between January 2013 and June 2016 in selected health facilities in Tanzania. Out of 94 sputa, 25 (26.60%) were GeneXpert® mycobacteria TB positive and 11/94 (11.70%) repeat-culture positive; 5 were Capilia TB-Neo positive and confirmed by GenoType MTBC to be *M. tuberculosis/M. canettii*. The remaining 6 Capilia TB-Neo negative samples were genotyped by GenoType® CM/AS, identifying 3 (3.19%) NTM, 2 Gram-positive bacteria, and 1 isolate testing negative, together, making a total of 6/94 (6.38%) confirmed false smear-positives. Twenty-eight (29.79%) were confirmed TB cases, while 60 (63.83%) remained unconfirmed cases. Out of 6 (6.38%) patients who were HIV positive, 2 patients were possibly coinfected with mycobacteria. The isolation of NTM and other bacteria among smear-positive culture-negative samples and the presence of over two-thirds of unconfirmed TB cases emphasize the need of both advanced differential TB

[3] *International Journal of General Medicine 2017:10 199–205*

diagnostic techniques and good clinical laboratory practices to avoid unnecessary administration of anti-TB drugs.

Keywords: Nontuberculous mycobacteria (NTM); Pulmonary TB; smear-positive; culture-negative; Tuberculosis (TB) diagnosis

5.1 Introduction

In 2014, the WHO adopted a new strategy to reduce TB deaths by 90% and TB incidence by 80% within the 15 years window from 2015 to 2030; however, diagnosis of TB remains challenging (WHO, 2016), particularly to resource-constrained countries and contexts of high TB-incidence rates. The widely used simple and inexpensive sputum smear microscopy for diagnosis of pulmonary TB (McNerney and Daley, 2011; Pai and Schito, 2015) suffers from both lacks of sensitivity and specificity (McNerney and Daley, 2011; Parsons *et al*., 2011). The performance of smear microscopy is further reduced in patients with extra-pulmonary TB, co-infected with HIV (Parsons *et al*., 2011; Weyer *et al*., 2013), and with NTM (Claude *et al*., 2016; Weyer *et al*., 2013).

Nontuberculous mycobacteria are ubiquitous bacteria found in water and soil (Maiga *et al*., 2012; Raju *et al*., 2016), that are well known to cause lung diseases (Henkle *et al*., 2016; Johnson and Odell, 2014). In recent years, pulmonary NTM infections have dramatically increased worldwide (Kendall and Winthrop, 2013; Maiga *et al*., 2012; Raju *et al*., 2016) and widely varies across geographical scale and human population. Defining the clinical implication of NTM infection in settings with endemic TB like Tanzania requires discrimination of NTM from *Mycobacterium tuberculosis* complex suspects (Hoza *et al*., 2016). Shared morphological characteristics between NTM and MTBC contributes to the poor performance of smear microscopy (Raju *et al*., 2016). Thus, NTM infections are commonly misdiagnosed as pulmonary TB particularly in resource-limited countries (Maiga *et al*., 2012; Raju *et al*., 2016). Despite these limitations, sputum smear microscopy remains a mainstay technique for clinical diagnosis of pulmonary TB in Low and Middle-Income Countries and remains an integral part of the global TB control strategy (Claude *et al*., 2016). Current WHO TB guidelines recommend immediate initiation of anti-TB drugs and a regular follow-up schedule after smear-positive results. Thus, the limited capacity of smear examination to differentiate MTBC from NTM infection is not taken into account, resulting in unnecessary TB treatment.

Mycobacterial culture is more sensitive and highly specific than smear examination hence being the gold standard for diagnosis of TB (Agrawal *et al*., 2016). Identification of species from culture is based on phenotypic and biochemical tests which are not definitive (Johnson and Odell, 2014), and non-viable bacterial cells result in culture-negative (Ryu, 2015). NTM infections can falsely present as culture-positive and can even be misdiagnosed as MDR-TB

(Maiga *et al*., 2012; Shahraki *et al*., 2015). The clinical implication of culture false-positive has remained elusive. Nucleic acid amplification assays like GenoType CM/AS (Hain Lifescience, Nehren, Germany) and GeneXpert (Cepheid, Sunnyvale, California, USA) can be applied for quick identification and confirmation of TB, though the techniques are expensive for routine diagnosis. This study used both phenotypic and molecular techniques, to identify the likelihood of NTM or other TB-like infection among smear-positive culture-negative.

5.2 Materials and methods

5.2.1 Study design

This cross-sectional study, nested in the East Africa Public Health Laboratory Network (EAPHLN) project, utilized sputum specimens collected between Feb 2013 and June 2016. The broad objective of the EAPHLN project was to establish a network of efficient, high quality, accessible public health laboratories for the diagnosis and surveillance of TB and other communicable diseases in East Africa.

5.2.2 Settings

The EAPHLN project collected samples from the following selected sites: Kibong'oto Infectious Diseases Hospital (in Kilimanjaro Region), Musoma Regional Referral Hospital (in Mara Region) Mnazi Mmoja Referral Hospital (in Zanzibar), Levolosi Health Centre (in Arusha Region), Nyamagana Health Centre (in Mwanza Region), St Vicent Health Centre in Mkuranga (in Pwani Region) and Shinyanga Regional Referral Hospital (in Shinyanga Region). For the period Feb 2013-June 2016, a total of 530 patients aged 15 years and above were diagnosed smear-positive. Eighteen percent (94/530) were "smear-positive and culture-negative". Laboratory analysis to identify the proportion of NTM species among presumptive TB cases performed at the Central TB Reference Laboratory (CTRL), unless otherwise stated. CTRL is the only reference laboratory in the country, located in Dar es Salaam that serves health facilities in all regions of the country. CTRL performs smear microscopy and microbial culture, participates in epidemiological research on TB, and evaluates and provides diagnostic services for the quality assurance of the National Tuberculosis Control Programme. As a part of quality assurance and laboratory strengthening, CTRL corroborates with a number of supranational TB laboratories like UK NEQAS, Uganda Supranational Laboratory as well as the Supranational Laboratory of the Antwerp Institute of Tropical

Medicine, Belgium.

5.2.3 HIV testing

HIV tests were performed at study site level as per standard care as speculated in the national HIV and testing algorithm (NACP, 2013), to determine the HIV status of the consented participants.

5.2.4 Smear examination using Ziehl-Neelsen (ZN) stain

Sputum sample collection, processing, transportation, and detection of acid-fast bacilli (AFB) were performed according to the National guidelines for TB management (NTLP, 2015). Smear-positive sputum was sent to CTRL for smear re-examination and culture. Laboratory scientists/technicians read the smears blindly and grading the results using the WHO system.

5.2.5 TB detection by GeneXpert MTB/RIF assay

Using unprocessed sputa samples, GeneXpert tests (Cepheid, Sunnyvale, California, USA) were performed to all 94 sputa in accordance with the manufacturer's instructions for detection of TB. GeneXpert MTB/RIF assay is an automated hemi-nested real-time PCR assay for detection of both MTBC and rifampicin-resistance in two hours estimate. It amplifies an MTB specific sequence of the 81-bp *rpoB* gene, which is probed with molecular beacons for mutations within the rifampin-resistance determining region.

5.2.6 Repeat-culture

Sputum samples from each study site were processed, packaged and transported to CTRL, with a transit time of 2-3 days. At CTRL, samples were homogenized (using N-acetyl-cysteine) to free the bacilli from the mucus, cells or tissue, followed by decontamination procedure (using NaOH) that liquefies the organic debris and eliminates the unwanted normal flora. Processed sediments were used to inoculate on LJ culture media according to the standard operating procedure. Repeat-culture were performed after completion of sample collection period, using sputum specimens that were archived at -80 °**C** after initial culture. The inoculated culture was incubated and observed every week for 8 weeks before being regarded as negative. A culture was considered as contaminated following observation of overgrowth of microorganisms that were lacking characteristics of mycobacteria. Repeat cultures were performed to all 94 smear-positive but culture-negative sputa. Re-examination

with ZN stain was performed to confirm the presence of acid-fast bacilli among NTM suspects.

5.2.7 Species identification

The Capilia TB-Neo assay (an immunochromatographic test) were performed to all repeat culture positive in accordance with the manufacturer's instructions (TAUNS Laboratories, Inc. Japan) to differentiate MTB complex from NTM infections. The Capilia TB-Neo detects MPT64 protein secreted by species of MTB complex. Isolates of MTB complex detected by Capilia TB-Neo were further subjected to a GenoType MTBC assay (Hain Lifescience, Nehren, Germany) for confirmation and identification to the species level. In order to identify species that were Capilia TB-Neo negative (NTM suspects), two additional genotyping methods: the GenoType CM/AS (Hain) were performed according to the manufacturer's instructions. GenoType CM can simultaneously detect and differentiate up to 27 clinically relevant NTM species from MTBC. GenoType AS can detect and differentiate an additional 19 NTM species.

5.2.8 Demographic and clinical data

Patients' demographic variables (age and sex) and clinical information (HIV status and TB treatment history) were extracted from the EAPHLN project register books and tools at the sites and CTRL.

5.2.9 Definitions

As stated earlier, a definitive diagnosis of pulmonary TB is difficult to achieve, however, based on this study, we establish that:-

(i) True TB smear-positive (confirmed TB infection= smear-positive and confirmation by any of the nucleic acid amplification tests (GeneXpert, or GenoType CM/AS or GenoType MTBC).

(ii) False smear-positive (confirmed not TB infection) = smear-positive but nucleic acid amplification confirmed NTM or other acid-fast infection.

(iii) Unconfirmed TB case = smear-positive but negative in all other tests (culture and nucleic acid amplification).

5.2.10 Ethical consideration

The study was nested within the EAPHLN project which was approved by the National Health Research Ethics Sub-Committee in Tanzania. All study participants aged 18 years and above provided a written informed consent and those who were younger than 18 years, a written informed consent was obtained from parents or guardians.

5.2.11 Results

A total of 94 (17.74%) smear-positive culture-negative samples of presumptive TB patients presented with TB symptoms from a pool of 530 smear-positives collected between Jan 2013 and June 2016 were analyzed. The remaining 436 (82.26%) of 530 samples were smear-positive culture-positive and were not included in this analysis.

The overall mean age (SD) of the 94 patients was 40.54 (15.63) years. The majority 54/94 (57.45%) of the participants were males with mean age of 41.8 (17.04) years. The mean (SD) age of 40 (42.55%) female participants was 38.85 (13.54) years. Table 4 reports on the distribution of the participants by study sites.

Table 4: Distribution of participants by study site

Health facility	**Smear-positive, Culture-negative**		
	Female	**Male**	**Total**
Kibong'oto	6	16	22
Levolosi	2	8	10
Mkuranga	1	2	3
Musoma	9	12	21
Mwanza	9	3	10
Shinyanga	5	9	14
Zanzibar	8	4	12
Grand Total	**40**	**54**	**94**

Out of 94 sputum samples, 25 (26.60%) were GeneXpert positive (no RIF resistance detected), of which 2 were culture positives. A total of 69 (73.40%) sputa were GeneXpert negatives, of which only 9 were culture positives. Only 2/25 GeneXpert positives were also culture-positives.

Repeat LJ culture yielded 11 (11.70%) growths, 2 were from a pool of 25 GeneXpert positives and the remaining 9 from 69 GeneXpert negative sputa. Additionally, two culture growth were discarded due to contamination during culture preparation. Eighty-three (88.30%) remained negative on culture to the termination point that was after 8 weeks.

Of the 11 culture-positive, 5 tested Capilia TB-Neo positive for MTB complex and confirmed by GenoType MTBC test all being *M. tuberculosis/M. canettii*. GenoType MTBC assay cannot distinguish *M. tuberculosis* from *M. canettii*. The remaining 6 Capilia TB-Neo negative isolates were typed by GenoType CM which identified 3/94 (3.19%) NTM infections (2 *Mycobacterium fortuitum*, 1 *Mycobacterium abscessus*, 2-gram +ve bacteria and 1 isolate was negative. The 2 gram +ve bacteria and the negative isolate were subjected to GenoType AS assay, all 3 were negative. Sixty (63.83%) sputa were negative in both GeneXpert and culture. Laboratory workflow and outcomes are summarized in Fig. 11. Six (6.38%) of the 94 patients were HIV-positive. Of the 2 *M. fortuitum* infection, 1 patient was HIV co-infected and 1 patient from a pool of 5 *M. tuberculosis/M. canettii* cases had also HIV co-infection.

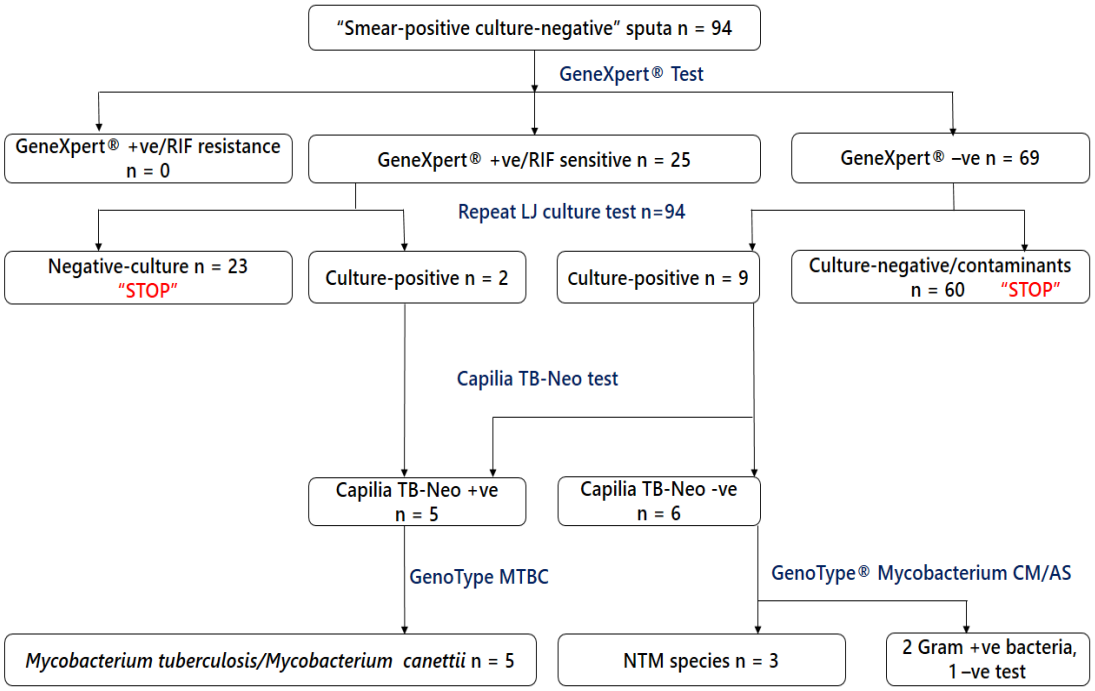


Figure 11: Flowchart showing laboratory procedure and outcomes

Based on the tests performed and case definition established earlier, a total of 28/94 (29.79%) patients sputa (GeneXpert positive plus GenoType MTBC positives), were referred as "confirmed TB cases" or "true smear-positive" while 6 (6.38%) cases were referred as "confirmed not TB infection" or "false smear-positive"; among these 3 were NTM infections, 2 gram +ve bacteria and 1 not identified by the tests. The remaining 60 (63.83%) samples that were negative in both GeneXpert and culture, were referred as "unconfirmed cases".

5.2.12 Discussion

Early detection and effective treatment are central components for successful management and control of every infectious disease including pulmonary TB. The diagnosis of pulmonary TB by sputum smear microscopy is seriously obstructed by its limited sensitivity and specificity (Pai and Schito, 2015; Parsons *et al*., 2011). Although low sensitivity in the examination of stained acid-fast smears has been given high priority, little consideration is given to false smear-positive as it pertains to specificity.

Despite the fact that the sensitivity of smear microscopy is inferior to both, the GeneXpert and mycobacterial culture, this study report 60/94 smear-positive sputa but negative in both GeneXpert and culture. Due to retrospective nature of this study, we could not to confirm the actual cause of these cases. However, laboratory technical failures and presence of non-viable bacilli can result in no growth in culture (Lee *et al*., 2008). Toxicity of reagents for decontamination, centrifugation, transportation and storage are all known to affect the viability of the bacilli if not performed properly (Parsons *et al*., 2011). Similar reasons that decontamination procedure and centrifugation, if not accurately performed, might have contributed to the miss of 5/11 cases of *M. tuberculosis/M. canettii* on the initial culture but identified by the repeat-culture. Demonstration studies documented high sensitivity and specificity (97–100%) of GeneXpert to detect TB and RIF resistance; however, the technology requires stable electricity, annual recalibration, and compliance with temperature ceiling (Trébucq *et al*., 2011). The impact of these on the sensitivity of the instrument for routine diagnosis pulmonary TB in places where there is a shortage and unstable power supply like Tanzania are not well established. However, due to the high specificity of this nucleic acid amplification technology, we conclude that 25 (26.60%) GeneXpert positive (2 confirmed by GenoType CM), was referred as "true pulmonary TB cases".

Recently, there has been increasing trends of respiratory infection worldwide due to NTM which are significantly varying across different geographical localities (Hoza *et al*., 2016; Maiga *et al*., 2012; Raju *et al*., 2016). However, older adults and immunocompromised persons like HIV-infected persons are at more risk of the infection (Hatta *et al*., 2010; Henkle *et al*., 2016). NTM infections present similar clinical and radiographic manifestations like those of pulmonary TB (Koh *et al*., 2006; Maiga *et al*., 2012), hence resulting false in positives in both smear and culture. In Tanzania, it is commonly assumed that patients present with pulmonary infection are infected with *M. tuberculosis* (Hoza *et al*., 2016; Kilale *et al*., 2016), with the possibility of other respiratory infection like NTM infection not taken into consideration. In this study, 3 (3.19%) were confirmed NTM infections (*2 M. fortuitum and M. abscessus*) but may have been mistreated as pulmonary TB, which may result in worsening patient's health, repeated visits to health facilities, and raise a risk of drug-resistance. NTM infection does not respond to anti-TB agents, however, the clinical impact of the indiscriminate use of anti-TB drugs in the tremendously recent emerging drug-resistant TB remain unassessed (Mnyambwa *et al*., 2017a; Pai and Schito, 2015). Infection due to species of *M. fortuitum* and *M. abscessus*, which both have the capacity to cause respiratory infections in human, have previously been reported in Tanzania (Mfinanga *et al*., 2014; Mpagama *et al*., 2013). Although this study did not conduct drug sensitivity testing, both *M. fortuitum* and *M. abscessus* can be misdiagnosed as resistant-TB/MDR-TB due to their natural resistant pattern in anti-tuberculosis drugs (Mpagama *et al*., 2013; Shahraki *et al*., 2015), which may result in unnecessary administration of toxic second-line anti-TB treatment. The prevalence of NTM reported in the present study which seems to be lower than recently reported in the country (Hoza *et al*., 2016; Kilale *et al*., 2016), is not conclusive because of limited sample size and which constitute a small proportion of pulmonary TB diagnosis scenarios. Moreover, due to retrospective nature of this study, we could not certainly implicate the isolated NTM as the cause (and the only cause) of the infection in these patients.

Additionally, various particles that are acid-fast (e.g. food particles, dye precipitates, and non-mycobacterial species) can cause false smear-positive. For example, *Nocardia sp* and *Rhodococcus equi* (Hatta *et al*., 2010), have a similar appearance to that Mycobacterium on smear examination and can be misdiagnosed as pulmonary TB. Recently, one study in Tanzania isolated different species of Nocardia among TB suspects, suggesting that Nocardia species are an important cause of pulmonary infections that are merely underdiagnosed

and/or ignored (Hoza *et al*., 2017). Furthermore, poor quality of smear microscopy and reading errors may result in either failure to detect active TB or reporting non-TB cases (Manalebh *et al*., 2015; Mfinanga *et al*., 2008). In this study, 3/11 isolates (2 Gram +ve bacteria and 1 isolate) could not be identified to species level due to the limitations of the typing methods used in this study. GenoType CM can detect gram +ve bacteria with a high G+C content, however, careful evaluation is needed to rule out the possibility of coexistence with mycobacterial species (Richter *et al*., 2006).

Limitations: This study could not assess the association of NTM infection with chronic diseases (e.g. HIV) and socio-demographic characteristic like age, sex due to a small study population. It was also difficult to assess treatment outcomes among patients due to retrospective nature of this study. Furthermore, this study could not rule out the possibility of bacterial colonization and/or contamination among the reported NTM or bacterial cases.

5.2.13 Conclusion

Our study reports that, out of 94 smear-positive cases, 64% were both GeneXpert and culture-negative and identified non-tuberculous mycobacteria among TB suspects. Less consideration of false-smear positives may result in mismanagement of a health condition. It is important that the national TB program to consider strengthening quality assurance system and track the performance of both smear microscopy and mycobacterial culture system regularly. Use of advanced differential TB diagnostic techniques, reviewing and emphasizing better clinical practices that ensure accuracy in diagnosis would help to avoid putting patients' health and lives at risk.

Conflict of interest: The authors declare no conflicts of interest

Acknowledgment: This work was mainly supported by the EAPHLN project, which was appraised in February 2010 and implemented regionally within the East Africa countries with the soft loan from the World Bank. The work was partly supported by the DELTAS Africa Initiative [Afrique One-ASPIRE /DEL-15-008]. Afrique One-ASPIRE is funded by a consortium of donors including the African Academy of Sciences (AAS), Alliance for Accelerating Excellence in Science in Africa (AESA), the New Partnership for Africa's Development Planning and Coordinating (NEPAD) Agency, the Wellcome Trust [107753/A/15/Z] and the UK government. The authors extend sincere appreciations to the Nelson Mandela African Institution of Science and Technology for providing adequate

learning resources and invaluable support, and to the management of the CTRL where this piece of work was carried out. The views expressed in this publication are those of the authors and not necessarily those of Afrique One-ASPIRE or the World Bank.

CHAPTER SIX

Assessment of GeneXpert-GxAlert platform for multi-drug resistant tuberculosis diagnosis and patients' linkage to care in Tanzania

Nicholaus Peter Mnyambwa[1,2*,] Issack Lekule[3], Esther S Ngadaya[2], Godfather Kimaro[2], Pammla Petrucka[1,4], Dong-Jin Kim[1], Johnson Lymo[5], Rudovick Kazwala[6], Fausta Mosha[7], Sayoki G Mfinanga[2]

[1]School of Life Sciences and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, Tanzania. [2]National Institute for Medical Research, Muhimbili Medical Research Centre, Dar es Salaam, Tanzania. [3]Kibong'oto Infectious Diseases Hospital, Kilimanjaro, Tanzania. [4]College of Nursing, University of Saskatchewan, Saskatoon, Canada. [5]National Tuberculosis and Leprosy Programme, Tanzania. [6]Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania. [7]National Health Laboratory, Ministry of Health Community Development Gender Elderly and Children, Republic of Tanzania

**Corresponding author: lodnicho@gmail.com*

Abstract

Objective: The gap between patients diagnosed with multi-drug resistant tuberculosis (MDR-TB) and enrolment in treatment is one of the major challenges in tuberculosis control programmes. A 4-year (2013–2016) retrospective review of patients' clinical data and subsequent in-depth interviews with health providers were conducted to assess the effectiveness of the GeneXpert GxAlert platform for MDR-TB diagnosis and its impact on linkage of patients to care in Tanzania.

Results: A total of 782 new rifampicin resistant cases were notified, but only 242 (32.3%) were placed in an MDR-TB regimen. The remaining 540 (67.07%) patients were not on treatment, of which 103 patients had complete records on the GxAlert database. Of the 103 patients: 39 were judged as untraceable; 27 died before treatment; 12 were treated with first-line anti-TBs; 9 repeat tests did not show rifampicin resistance; 15 were not on treatment due to communication breakdown, and 1 patient was transferred outside the country. In-depth interviews with health providers suggested that the pre-treatment loss for the MDR-TB patients was primarily attributed to health system and patients themselves.

Conclusion: Strengthening the health system by developing and implementing well-defined interventions to ensure all diagnosed MDR-TB patients are accurately reported and timely linked to treatment are highly recommended.

Keywords: Multi-drug resistant tuberculosis (MDR-TB); MDR-TB treatment; Tuberculosis (TB) diagnosis; GeneXpert; GxAlert; rifampicin resistant TB.

6.1 Introduction

Tuberculosis remains an important health problem worldwide. Effective TB control interventions require early identification of TB infected persons and tailored therapies. As a result of limited performance of smear microscopy (Claude *et al*., 2016; Mnyambwa *et al*., 2017b) and a slow generation time of mycobacteria on culture media (Tyrrell *et al*., 2012), the WHO endorsed the use of GeneXpert MTB/RIF for rapid detection of TB and rifampicin resistance (WHO, 2013). As rapid diagnostic tools become available, the main challenge remains on how responsive the health system is to ensure effective linkage of diagnosed patients to requisite healthcare.

Tanzania is ranked among 20 countries with the highest TB incidences worldwide (WHO, 2016). An increase in MDR-TB cases within the country was first noted in 2005, resulting in the concerted efforts by the Ministry of Health, which include designation Kibong'oto Infectious Diseases Hospital (KIDH) as the first centre for treatment of MDR-TB (MoHSW, 2009). In 2013, a pilot GxAlert health platform was installed connecting GeneXpert diagnostic facilities with the aim of monitoring GeneXpert results and improving the linkage of TB patients to care as facilitated by the real-time reporting of diagnosis results hence decision making. The GxAlert initiative is an electronic system for data management, designed to enable a range of TB, HIV, and Ebola diagnostic devices to connect to the mobile network/channels (Abt Associates, 2013). Results from the GeneXpert device is automatically sent to the GxAlert, in which a short message is generated and sent through mobile channels or over internet as e-mail to referring clinicians, treatment centres, and the country's existing health information systems. Ideally, this systematic approach would contribute to the patients being diagnosed and managed in a timely and consistent manner. This study was conducted to assess the effectiveness of GeneXpert-GxAlert health platform for MDR-TB diagnosis and its facilitation of the linkage of patients to healthcare services.

6.2 Materials and methods

6.2.1 Design

This is a retrospective review of routine clinical data of a cohort of MDR-TB patients diagnostically confirmed by the GeneXpert and complemented with interviews with health providers: Regional TB Leprosy Coordinators (RTLCs) and District TB and Leprosy Coordinators (DTLCs) from all parts of the country where patients were not enrolled in

treatment. TB services are provided free of charge and coordinated by DTLCs and RTLC coordinators at district and regional levels, respectively.

Additionally, efforts were made to reach the patients through by phones and/or physically whenever necessary.

6.2.2 Study population

A cohort of MDR-TB patients with a confirmed GeneXpert test, diagnosed between January 01, 2013 and September 21, 2016.

6.2.3 Data collection

A list of rifampicin-resistant patients with complete information from the GxAlert system was formulated from the CTRL MDR-TB database, which is composed of test results from all networked GeneXpert within Tanzania. Patients' information was extracted from TB treatment (KIDH) and laboratory registries as well as from the GeneXpert machines. Information extracted included patient demographic information (age, sex, and contact/physical address), place of diagnosis, diagnosis results and whether the patient was placed in treatment or not.

6.2.4 Qualitative data

Before analysis, all data from hospital records were cleaned by checking for accuracy, and completeness. Recorded audio of in-depth interviews (30-45 minutes) with key informants were translated from Swahili to English. Thematic analysis was conducted for themes, patterns and organized into main thematic categories. Finally, description of each theme was established supported with few quotations.

Key informants Knowledge of MDR-TB diagnosis and management: Assessment of level of knowledge of the MDR-TB was based on diagnosis and management of MDR-TB. A good knowledge included whether the coordinator was able to interpret GeneXpert results, aware that sputum sample of every TB patient should be sent to CTRL for culture, and Drug Sensitivity Testing, and arranging placing transport order from KIDH to pick an MDR-TB patient up to KIDH for treatment initiation.

6.3 Results

By the end of 2016, there were 70 GeneXpert devices, of which 37 were connected to GxAlert. Fig. 12 summarizes the MDR-TB notification process to clinicians and TB program officials in Tanzania. A monthly summary of diagnosis results was produced from each diagnostic centre with GeneXpert not connected to GxAlert. These results were aggregated into the GxAlert database using Xpert tracking tool. A total of 878 rifampicin resistant cases were registered in the GxAlert database between January 01, 2013 and September 21, 2016. Out of the 878 patients, 58 were identified as duplicates and 38 samples for External Quality Assurance processes. The remaining 782 were verified unique patients registered in the GxAlert; however, only 242 (32.3%) were found to be in MDR-TB treatment regimen at KIDH, leaving 540 (67.7%) patients outside the requisite continuum of care. Amongst the 540 patients, only 103 had all required information (including patients' name and identification) at CTRL. The remaining 437 were considered incomplete records; of which twenty-five (25) were missing both names and identification.

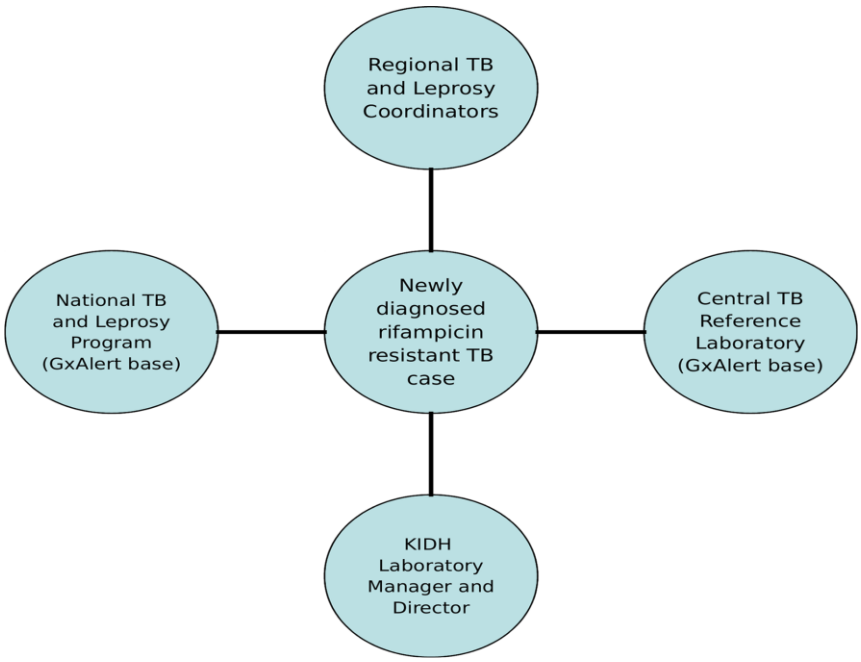


Figure 12: The visualization of the TB 'spill out' of the GxAlert notification process in Tanzania

Of the 103 patients with complete records, 27 died before MDR-TB treatment initiation with 69 days mean duration from the diagnosis date; 39 patients had no medical records available, and 12 patients were treated with first-line anti-TBs. Of the 12 patients treated with first-line drugs, 5 had repeat tests rifampicin resistance not detected; 4 had no records available; 2

results were not correctly recorded in the laboratory register; and 1 case was confirmed MDR-TB, but the DOT nurse did not interpret the result correctly, and the patient was put on first-line TB drug. Nine patients were not placed on MDR-TB treatment because GeneXpert repeat tests showed no resistance. Of all 103 patients, 11 patients were successful traced by the study team and were not in treatment due to communication breakdown among health providers (laboratory technicians, DOT nurses, DTLCs) and patients. These 11 patients were subsequently referred to their respective DTLC for treatment procedure. Another 4 patients were not in treatment because results were not communicated to DTLCs. One patient from Morogoro was initiated with first-line anti-TB agents and later transferred by family to South Africa for treatment (Table 5).

Table 5: Distribution of patients with complete records

Region	Died	Not in treatment due to miscommunication	Not traceable	Placed on 1st line anti-TBs	Repeat test RIF res not detected	Total
Tabora	0	0	0	1	0	**1**
Shinyanga	2	1	1	0	2	**6**
Simiyu	1	0	0	0	0	**1**
Mwanza	2	0	0	0	0	**2**
Geita	0	0	2	1	0	**3**
Kagera	1	0	0	1	0	**2**
Songea	0	1	1	1	1	**4**
Mbeya	3	0	0	0	5	**8**
Katavi	1	0	1	0	1	**3**
Tanga	1	1	0	0	0	**2**
Manyara	0	0	0	2	0	**2**
Pwani	0	0	0	1	0	**1**
Lindi	0	0	1	1	0	**2**
Mtwara	1	1	0	1	0	**3**
Morogoro	3	2	0	0	0	**5**
Dar es Salaam	12	9	33	3	0	**57**
Total	**27**	**15**	**39**	**12**	**9**	**102**

Medical records for 39 patients were missing and patients were not reachable though mobile phones and physical addresses. Of the 103 patients, only 11 were interviewed by the research team as a complement to factors associated with pre-treatment loss. To avoid bias, due to small number of patients interviewed, results are not shown in this work.

Key informants (KIs): Twenty-seven (27) TB coordinators at the district and regional levels were interviewed to gather information on their knowledge of MDR-TB management and challenges for effective program implementation. Out of these 21 (77.8%) were males with mean age of 47.5±8.8 and average of 7.7±6.1 years work experience as TB coordinators. Of all TB coordinators, 17 (63.0%) were DTLCs while 10 (39.0%) were RTLCs. At the time of data collection, only RTLCs were connected to GxAlert notification system: all acknowledged receiving short text notifications via mobile channels when an MDR-TB case was detected by GeneXpert, and communicated with respective DTLC and KIDH. All had good knowledge of MDR-TB patient management, while most (70.4 %, n =20) were aware of the classification of GeneXpert results and treatment regimens.

Following an in-depth interview with key informants and review of patients' records, factors associated pre-treatment loss were either related with the health-system, patient, social issues, or death of patient before taken to KIDH. Inadequate patient's records, miscommunication between health providers, failure to assess "low" and "high" likelihood of MDR-TB and lack of transportation to KIDH were noted during interviews.

> *One KI said 'Results obtained on September 28, 2015, and the patient died on November 2, 2015, before taken to KIDH… another MDR patient who was an antiretroviral therapy defaulter, he did not start the treatment as he was not reported to the RTLC/DTLC after the results'.*

Some MDR-TB patients were reluctant to provide correct contact addresses fearing disclosure of their health condition to relatives. In some cases, the illness is attributed to superstition.

> *'…. this patient was a traditional healer, after a couple of days, his condition became worse; he decided to contact RTLC for help. RTLC made a quick logistical arrangement for a car from KIDH, but the patient died before taken to KIDH'. KI said.*

By the end of 2015, there was only one hospital mandated to initiate MDR-TB treatment, so social issues (e.g., parenting) related to separation for prolonged MDR-TB treatment and hospitalization away patient's residence was highly complicated. For example, some worried losing their jobs.

'You know some patients are heads of households and children depend on them, they deny to leave their families to KIDH for a long duration of treatment'. KI said.

6.4 Discussion

The findings of this study demonstrate that the installation of the GxAlert and GeneXpert enhance rapid diagnosis of TB and improve communication of the results among health providers. This linkage could ideally improve patient's follow-up and timely access to healthcare services. However, of the MDR-TB patients diagnosed by the GeneXpert between January 2013 and September 2016, only one-third were placed into care. This reflect a small proportion of a total number of MDR-TB patients diagnosed during the same time period in the country. Similarly, in 2015 the WHO reported only 125 000 (20%) of 580 000 new patients eligible for MDR-TB treatment globally were placed on treatment (WHO, 2016). This persistent and complex gap perpetuates negative implications on the transmission, severity, and prognosis as well as an ethical dilemma since any diagnosed patient deserves to be provided with appropriate treatment.

Despite the added advantage of access to real-time information through GeneXpert-GxAlert platform, obstacles such as inaccurate or incomplete patient data entry during the operation of GeneXpert "garbage in, garbage out", and communication breakdown among health providers (laboratory technicians, referring clinicians, TB coordinators), hinder an effective program implementation. More than two-thirds of patients had incomplete information in the GxAlert database. This scenario contributes to a delay, pre-treatment loss and/or non-standardized therapy, which, in turn, may contribute to continued transmission, development of drug resistance and high mortality rates among MDR-TB patients (Xu *et al*., 2017). Treating MDR–TB patients with first-line anti-TBs may worsen patient's health, make the disease difficult to cure and death. Limited knowledge among laboratory personnel on the new TB coding systems resulted in misreporting and inconsistency in reporting the GeneXpert results. Misreporting of patients refrained some of the patients from getting proper health care. Successful MDR-TB treatment requires adherence to the MDR-TB guideline such as early detection and referring an MDR-TB patient to a specialized TB centre for treatment initiation with second-line anti-TB drugs, with close monitoring of the treatment outcome (WHO, 2009) as the Ministry of Health works to scale up diagnostic platform, harmonization of the recording system, inclusion of DTLCs in the GxAlert system, and training to all DOT nurses, TB coordinators, and laboratory technicians on the new

diagnostic algorithm and management of TB/MDR-TB could significantly improve provision of healthcare services.

Traditional beliefs and superstitions, as well as limited knowledge on transmission and treatment of the disease, stigma related to TB/MDR-TB, opting alternative treatment, might have contributed to pre-treatment loss. Healthcare seeking behaviors among TB patients correlate with their level of knowledge and awareness of TB (Khan *et al*., 2006). Stigmatization of MDR-TB patients, due to fear of spread of infection, has been elsewhere reported as contributing to treatment delay (Courtwright and Turner, 2010; Daftary, 2012).

For the period 2013-2015, KIDH was the only specialized reference hospital for the treatment of MDR-TB. The ongoing effort to decentralize MDR-TB therapy should aim at expanding services to the district level, hence providing choice for MDR-TB patients where to get treatment and making 'nearer' to home services a reality. The duration for MDR-TB treatment takes at least 20 months (WHO, 2014), and a patient remain hospitalized until the sputum smear become negative, which takes more than 2 months (Lange *et al*., 2014).

Our findings revealed that MDR-TB patients often spend over two months with their families and die before getting appropriate treatment. If these patients were to receive treatment in timely manner, most of these deaths could have been averted. The delays can be attributed to both patients, and the health system (Belay *et al*., 2012).

Limitation: Patients untraceability prevented interviews that could preciously establish factors associated with pre-treatment loss; hence there is potential bias. Diagnosis rate and timing, as well as treatment outcomes, were not assessed.

6.5 Conclusion

The study revealed a significant proportion of MDR-TB patients were not enrolled in treatment often attributable to; substantial inconsistencies and deficiencies in the reporting of diagnosis results from the GeneXpert. Training of laboratory technicians and clinicians involved in the TB program and strengthening the health community workers for TB patients tracing is recommended.

Ethics approval and consent to participate: This study was approved by the National Institute for Medical Research (Tanzania). Permission to conduct the study was obtained

from the District Medical Officer of each respective district. A written informed consent was obtained from all key informants who consented that their quotes can be anonymously published.

Consent for publication: Not applicable

Availability of data and material: The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Competing interests: The authors report no financial conflict of interests

Authors' contributions: NPM, IL, ESN, SGM, and GK designed the study, ESN and SGM coordinated and supervised the study. NPM, IL, ESN, and SGM analyzed the data and drafted the manuscript. The remaining authors PP, DJ, JL, RK, and FM contributed to the implementation and write-up of the manuscript.

Funding: This work was mainly funded by the World Bank through the East Africa Public Health Laboratory Network (EAPHLN) project and partially supported by the DELTAS Africa Initiative [Afrique One-ASPIRE/DEL-15-008]. Afrique One-ASPIRE is funded by a consortium of donors including the Alliance for Accelerating Excellence in Science in Africa (AESA), African Academy of Sciences (AAS), the New Partnership for Africa's Development Planning and Coordinating (NEPAD) Agency, the Wellcome Trust [107753/A/15/Z] and the UK government

Acknowledgement: The authors extend sincere appreciations to Said Mfaume (Country GeneXpert focal person) and CTRL Management for their support during this study.

CHAPTER SEVEN

General discussion, conclusion and recommendations

7.1 Introduction

This forms the sixth and last chapter of this dissertation. It summarizes key findings of the study and its programmatic implication for tuberculosis control interventions and recommendations for future investigations.

7.2 General discussion

The present study reports the following key findings: (i) recovered 11 known drug-resistant genes; no mutation was detected in *emB* gene (ii) Drug resistance genetic markers in novel genes and new genetic variants in some known drug-resistant genes (iii) isolation of *M. yongonense* case in a patient who was diagnosed as having MDR-TB. To our best of knowledge, this is the first case in Tanzania and Africa at large (iv) Misdiagnosed cases of TB (v) Nearly two-thirds of MDR-TB patients diagnosed by GeneXpert between 2013 and 2016 were not placed in treatment.

Chapter one: This one gives background information, research questions and objectives of the study. The chapter also describes why is study the important and should take advantage of advances molecular techniques for better understanding of the genetic basis of resistance and answering important epidemiological questions as well as delineating strain epidemiology of the pathogen. While Genetic data on the pathogen has been accumulating in developed countries, the same information is lacking in resource-limited countries and where the burden of TB is high.

Chapter two is literature review highlighting how the emerging novel drug resistant-conferring mutations are revolutionizing the pathways of antibiotic resistance of clinical tuberculosis in clinical settings. This article reviews our current understanding of the genetic changes that characterize drug resistance in tuberculosis and highlight the imperative for further investigations focusing how individual mutation works and interact each other that contributes to complex genetics of resistance Generally, information on the genetics of clinically resistant TB and detailed strain epidemiology is limited, particularly as these pertain to technology-limited countries with high tuberculosis incidence rates thereby

preventing effective control interventions of the disease. This calls for localized TB programs to build effective evidence-based interventions that contribute to meaning global intervention.

Chapter three hypothesized that drug-resistant genes or IGRs in Mycobacterial tuberculosis infections should be diversifying by harbouring a higher density of mutations (nonsynonymous SNPs and IGR SNPs). A high proportion of variants were detected in *ethA*, *inhA* and *katG* genes. These findings expand the spectrum of potential drug resistance-conferring mutations in MTBC clinical isolates suggesting the need of molecular diagnostic for delineating the drug susceptibility profiles of MTBC isolates in clinical settings. Both known, additional variants/SNPs and novel genetic markers for resistant tuberculosis is described. The study did not demonstrate the presence of IGRs and *embCAB* gene, but this can be described by the low average sequence coverage of this study. High-coverage sequencing is recommended to maximize SNP discovery, genotype and variants calling accuracy to the wrong biological conclusion (Sims *et al*., 2014).

Chapter four: This chapter describes unexpected findings of isolation of *M. yongonense* from a patient was diagnosed with MDR-TB. *Mycobacterium yongonense* is a recently reported novel species of mycobacterium belonging to MAC (Kim *et al*., 2013a, b); to the best of our knowledge, the findings present the first isolation *M. yongonense* in Tanzania. Effective diagnostic tools should also help to answer epidemiological questions regarding the *M. yongonense* species. This is highlighting the importance of elucidating the epidemiology and clinical implication of *M. yongonense* which remain unknown because of a few cases have been reported so far. To achieve this, more advanced differential molecular diagnostic tools should be developed.

Chapter five: This chapter employed molecular tools to investigate the observation of high rate of culture-negative (a gold standard for TB diagnosis) among smear-positive samples. Our findings present possible misdiagnosed TB cases that resulted into the unnecessary administration of anti-TB drugs to these patients, suggesting that diagnosis of tuberculosis remain challenging particularly in technology-limited countries like Tanzania. We confirmed the presence of non-tuberculous mycobacteria that were misdiagnosed as TB and over two-thirds of the cases were not confirmed as TB cases. This emphasizes the need for both advanced differential TB diagnostic techniques and good clinical laboratory practices to avoid unnecessary administration of anti-TBs which might contribute to chronic nature of the disease and development of drug-resistant strains.

Chapter six: This chapter describes assess the operation of GxAlert health system for MDR-TB diagnosis and explores possible challenges for effective implementation of the program in Tanzania. We found a tremendous number (2/3) of patients who were diagnosed with MDR-TB between 2013 and 2016 were not in MDR-TB treatment. Such gap between diagnosed with MDR-TB and enrolled in treatment carries a negative implication on the transmission, severity and prognosis of the diseases. Urgently well-defined interventions are needed to strengthen TB program and health system to ensure timely detection of the disease and appropriate treatment.

7.3 Conclusion

This study reports potential novel drug resistance-conferring mutations (genes and additional variants) in clinical MTBC clinical isolates suggesting the need for effective molecular diagnostic tests for delineating the drug susceptibility profiles of MTBC isolates in clinical settings. We also identified report the first case of *M. yongonense* in Tanzania, isolated from a patient who was possibly misdiagnosed with MDR-TB. Misdiagnosis was also demonstrated by isolation of a number of NTMs among presumptive TB patients who were treated with anti-TB drugs. Urgent efforts are needed to strengthen TB diagnosis and treatment programs to ensure effective infection control strategies in order interrupt the transmission chain of TB/MDR-TB. These finding highlights the need for accurate diagnostic assays to ensure the effective use of drugs and to minimize the risk of the development of drug resistance and worsening of the patient's health condition.

7.4 Recommendations

(i) It is important that the national TB program to consider strengthening quality assurance system and track the performance of both smear microscopy and mycobacterial culture system regularly.

(ii) Use of advanced differential TB diagnostic techniques, reviewing and emphasizing better clinical practices that ensure accuracy in diagnosis would help to avoid putting patients' health and lives at risk

(iii) Training laboratory technicians and clinicians involved in TB programmes on the current molecular diagnostic methods and laboratory practices as well as strengthening the health community workers involved in TB patients tracing and health education is highly recommended.

(iv) Effective infection control strategies are urgently needed to interrupt the transmission chain of TB/MDR-TB and prevent the emergence of new resistant TB cases.

(v) Population genetics studies are needed to accurately determine the actual spectrum of drug resistance, strain epidemiology and transmission dynamics of the pathogen to inform well-defined TB control programs in the country.

REFERENCES

Abt Associates (2013). The GxAlert Initiative Connecting Rapid TB Diagnosis with a Faster Health System Response for Better Health Outcomes Faster Data Delivery Better Care of TB Patients Improved TB logistics. [http://www.abtassociates.com /AbtAssociates /files/be/be800553-1a76-4d3b-aa9b-13ce2b86babb.pdf]. Accessed on June 14, 2017

Agrawal, M., Bajaj, A., Bhatia, V. and Dutt, S. (2016). Comparative Study of GeneXpert with ZN Stain and Culture in Samples of Suspected Pulmonary Tuberculosis. *Journal of Clinical and Diagnostic Research.* **10**(5): DC09-12. doi: 10.7860/JCDR /2016 /18837. 7755.

Alangaden, G., Kreiswirth, B. and Aouad, A. (1998). Mechanism of resistance to amikacin and kanamycin in *Mycobacterium tuberculosis*. *Antimicrobial Agents*. [http://aac.asm.org/content/42/5/1295.short]. Accessed on May 10, 2018.

Alexander, K. A., Laver, P. N., Michel, A. L., Williams, M., van Helden, P. D., Warren, R. M. and van Pittius, N. C. (2010). Novel *Mycobacterium tuberculosis* complex pathogen, M. mungi. *Emerging Infectious Diseases*. **16**(8): 1296–1299. doi.org/10.3201/eid1608.100314.

Alland, D., Kalkut, G. E., Moss, A. R., McAdam, R. A., Hahn, J. A., Bosworth, W., … Bloom, B. R. (1994). Transmission of Tuberculosis in New York City. An Analysis by DNA Fingerprinting and Conventional Epidemiologic Methods. *New England Journal of Medicine*. **330**(24). 1710–1716. doi.org/10.1056/NEJM199406163302403.

Almeida, P. E. A. and Palomino, J. C. (2011). Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis*: classical and new drugs. *The Journal of Antimicrobial Chemotherapy*. **66**(7). 1417–1430. doi.org/10.1093/jac/dkr173.

Ani, A., Okpe, S., Akambi, M., Ejelionu, E., Yakubu, B., Owolodun, O., … Idoko, J. (2009). Comparison of a DNA based PCR method with conventional methods for the detection of *Mycobacterium tuberculosis* in Jos, Nigeria. *The Journal of Infection in Developing Countries*. **3**(06). 470–475. doi.org/10.3855/jidc.420.

Assefa, S., Keane, T. M., Otto, T. D., Newbold, C. and Berriman, M. (2009). ABACAS: algorithm-based automatic contiguation of assembled sequences. *Bioinformatics*. **25**(15). 1968–1969. doi.org/10.1093/bioinformatics/btp347.

Baker, M. A., Harries, A. D., Jeon, C. Y., Hart, J. E., Kapur, A., Lönnroth, K., … Murray, M. B. (2011). The impact of diabetes on tuberculosis treatment outcomes: A systematic review. *BMC Medicine*. **9**(1), 81. doi.org/10.1186/1741-7015-9-81.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*. **19**(5). 455–477. doi.org/10.1089/cmb.2012.0021.

Behr, M. A. and Waters, W. R. (2014). Is tuberculosis a lymphatic disease with a pulmonary portal? *The Lancet Infectious Diseases*. **14**(3). 250–255. doi.org/10.1016/S1473-3099(13)70253-6.

Belay, M. Bjune, G., Ameni, G. and Abebe, F. (2012). Diagnostic and treatment delay among Tuberculosis patients in Afar Region, Ethiopia: A cross-sectional study. *BMC Public Health*. **12**(1). 369. doi.org/10.1186/1471-2458-12-369

Ben Salah, I., Cayrou, C., Raoult, D. and Drancourt, M. (2009). *Mycobacterium marseillense* sp. nov., *Mycobacterium timonense* sp. nov. and *Mycobacterium bouchedurhonense* sp. nov., members of the Mycobacterium avium complex. *International Journal Of Systematic and Evolutionary Microbiology*. **59**(11). 2803–2808. doi.org/10.1099/ijs.0.010637-0

Benson, C. A. (1994). Disease due to the *Mycobacterium avium* complex in patients with AIDS: epidemiology and clinical syndrome. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*. Suppl 3. S218-22.

Boillat-Blanco, N., Ramaiya, K. L., Mganga, M., Minja, L. T., Bovet, P., Schindler, C., … Probst-Hensch, N. (2016). Transient Hyperglycemia in Patients With Tuberculosis in Tanzania: Implications for Diabetes Screening Algorithms. *Journal of Infectious Diseases*. **213**(7). 1163–1172. doi.org/10.1093/infdis/jiv568.

Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**(15). 2114–2120. doi.org/10.1093/bioinformatics/btu170.

Bos, K., Harkins, K., Herbig, A., Coscolla, M. and Weber, N. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature*. Nature **514**. 494–497 (2014).

Bouakaze, C., Keyser, C., de Martino, S. J., Sougakoff, W., Veziris, N., Dabernat, H. and Ludes, B. (2010). Identification and genotyping of *Mycobacterium tuberculosis* complex species by use of a SNaPshot Minisequencing-based assay. *Journal of Clinical Microbiology*. **48**(5). 1758–1766. doi.org/10.1128/JCM.02255-09.

Campbell, P. J., Morlock, G. P., Sikes, R. D., Dalton, T. L., Metchock, B., Starks, A. M., … Posey, J. E. (2011). Molecular detection of mutations associated with first- and second-line drug resistance compared with conventional drug susceptibility testing of *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*. **55(**5). 2032–2041. doi.org/10.1128/AAC.01550-10.

Casali, N., Nikolayevskyy, V., Balabanova, Y., Harris, S. R., Ignatyeva, O., Kontsevaya, I., … Drobniewski, F. (2014). Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nature Genetics*. **46**(3). 279–286. doi.org/10.1038/ng.2878.

Chonde, T. M., Basra, D., Mfinanga, S. G., M., Range, N., Lwilla, F. Shirima, R. P., … van Leth, F. (2010). National anti-tuberculosis drug resistance study in Tanzania. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*. **14**(8). 967–972.

Cingolani, P., Platts, A., Wang, L. L., Coon, M. Nguyen, T., Wang, L., … Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*. **6**(2), 80–92. doi.org/10.4161/fly.19695.

Claude, J., Ngabonziza, S., Ssengooba, W., Mutua, F., Torrea, G., Dushime, A., … Mwaengo, D. (2016). Diagnostic performance of smear microscopy and incremental yield of Xpert in detection of pulmonary tuberculosis in Rwanda. *BMC Infectious Diseases*. 1–7. doi.org/10.1186/s12879-016-2009-x.

Cohen, K. A., Abeel, T., Manson McGuire, A., Desjardins, C. A., Munsamy, V. Shea, T. P., … Earl, A. M. (2015). Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of *Mycobacterium tuberculosis* Isolates from KwaZulu-Natal. *PLOS Medicine*. **12(**9). e1001880. doi.org/10.1371/journal.pmed.1001880.

Cole, S. T., Brosch, R. Parkhill, J., Garnier, T. Churcher, C., Harris, D., … Barrell, B. G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. **393(**6685). 537–544. doi.org/10.1038/31159.

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

Comas, I., Homolka, S., Niemann, S. and Gagneux, S. (2009). Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS ONE*. **4**(11). e7815. doi.org/10.1371/journal.pone.0007815.

Corbett, E. L., Marston, B., Churchyard, G. J. and De Cock, K. M. (2006). Tuberculosis in sub-Saharan Africa: opportunities, challenges, and change in the era of antiretroviral treatment. *The Lancet*. **367**(9514), 926–937. doi.org/10.1016/S0140-6736(06)68383-9.

Courtwright, A. and Turner, A. N. (2010). Tuberculosis and Stigmatization: Pathways and Interventions. *Public Health Reports*. **125**(Suppl 4). 34–42. doi.org/10.1177/00333549101250S407.

Couvin, D. and Rastogi, N. (2015). Tuberculosis – A global emergency: Tools and methods to monitor, understand, and control the epidemic with specific example of the Beijing lineage. *Tuberculosis*. **95**. S177–S189. doi.org/10.1016/j.tube.2015.02.023.

Couvin, D., Rastogi, N., Rastogi, N., Sola, C., Shabbeer, A., Ozcaglar, C., … Rastogi, N. (2015). Tuberculosis - A global emergency: Tools and methods to monitor, understand, and control the epidemic with specific example of the Beijing lineage. *Tuberculosis (Edinburgh, Scotland)*. **95** (Suppl 1). S177-89. doi.org/10.1016/j.tube.2015.02.023

Crump, J. A., van Ingen, J., Morrissey, A. B., Boeree, M. J., Mavura, D. R., Swai, B., … van Soolingen, D. (2009). Invasive disease caused by nontuberculous mycobacteria, Tanzania. *Emerging Infectious Diseases*. **15**(1), 53–55. doi.org/10.3201/eid1501.081093

Cui, Z. J., Yang, Q.Y., Zhang, H. Y., Zhu, Q. and Zhang, Q. Y. (2016). Bioinformatics Identification of Drug Resistance-Associated Gene Pairs in *Mycobacterium tuberculosis*. *International Journal of Molecular Sciences*. **17**(9). 1417. doi.org/10.3390/ijms17091417.

Daftary, A. (2012). HIV and tuberculosis: The construction and management of double stigma. *Social Science and Medicine*. **74**(10), 1512–1519. doi.org/10.1016/j.socscimed.2012.01.027

Dalla Costa, E. R., Ribeiro, M. O., Silva, M. S., Arnold, L. S., Rostirolla, D. C., Cafrune, P. I., … Riley, L. (2009). Correlations of mutations in katG, oxyR-ahpC and inhA genes and in vitro susceptibility in *Mycobacterium tuberculosis* clinical strains segregated by spoligotype families from tuberculosis prevalent countries in South America. *BMC Microbiology*. **9**(1). 39. doi.org/10.1186/1471-2180-9-39.

Dannenberg, Jr. A. M. and Rook, G. A. W. (1994). Pathogenesis of Pulmonary Tuberculosis: an Interplay of Tissue-Damaging and Macrophage-Activating Immune Responses—Dual Mechanisms That Control Bacillary Multiplication. *In Tuberculosis* (pp. 459–483). *American Society of Microbiology*. doi.org/10.1128/9781555818357.ch27.

de Jong, B. C., Onipede, A., Pym, A. S., Gagneux, S., Aga, R. S., DeRiemer, K. and Small, P. M. (2005). Does resistance to pyrazinamide accurately indicate the presence of *Mycobacterium bovis*? *Journal of Clinical Microbiology*. **43**(7). 3530–3532. doi.org/10.1128/JCM.43.7.3530-3532.2005

de Vos, M., Müller, B., Borrell, S., Black, P. A., van Helden, P. D., Warren, R. M., … Victor, T. C. (2013). Putative compensatory mutations in the rpoC gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. *Antimicrobial Agents and Chemotherapy*. **57**(2), 827–832. doi.org/10.1128/AAC.01541-12.

Desjardins, C. A., Cohen, K. A., Munsamy, V. Abeel, T., Maharaj, K. Walker, B. J., … Pym, A. S. (2016). Genomic and functional analyses of *Mycobacterium tuberculosis* strains implicate ald in D-cycloserine resistance. *Nature Genetics*. **48**(5). 544–551. doi.org/10.1038/ng.3548

Dooley, K. E., Tang, T. Golub, J. E., Dorman, S. E. and Cronin, W. (2009). Impact of Diabetes Mellitus on Treatment Outcomes of Patients with Active Tuberculosis. *The American Journal of Tropical Medicine and Hygiene*. **80**(4), 634–639. doi.org/10.4269/ajtmh.2009.80.634

Durnez, L., Katakweba, A., Sadiki, H. Katholi, C. R., Kazwala, R. R., Machang'u, R. R., … Leirs, H. (2011). Mycobacteria in terrestrial small mammals on cattle farms in Tanzania. *Veterinary Medicine International*. **2011**. 495074. doi.org/10.4061/2011/495074

Farhat, M. R., Shapiro, B. J., Kieser, K. J., Sultana, R., Jacobson, K. R., Victor, T. C., … Murray, M. (2013). Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. *Nature Genetics*. **45**(10), 1183–1189. doi.org/10.1038/ng.2747

Fedrizzi, T., Meehan, C. J., Grottola, A., Giacobazzi, E., Fregni Serpini, G., Tagliazucchi, S., … Segata, N. (2017). Genomic characterization of Nontuberculous Mycobacteria. *Scientific Reports*. **7**. 45258. doi.org/10.1038/srep45258.

Fenner, L., Egger, M. Bodmer, T. and Altpeter, E. (2012). Effect of mutation and genetic background on drug resistance in *Mycobacterium tuberculosis*. *Antimicrobial Agents*. **56**(6).3047-53. doi: 10.1128/AAC.06460-11.

Firdessa, R., Berg, S., Hailu, E., Schelling, E., Gumi, B., Erenso, G., … Aseffa, A. (2013). Mycobacterial Lineages Causing Pulmonary and Extrapulmonary Tuberculosis, Ethiopia. *Emerging Infectious Diseases*. **19**(3). 460–463. doi.org/10.3201/eid1903.120256

Fishbein, S., van Wyk, N., Warren, R. M. and Sampson, S. L. (2015). Phylogeny to function: PE/PPE protein evolution and impact on *Mycobacterium tuberculosis* pathogenicity. *Molecular Microbiology*. **96**(5). 901–916. doi.org/10.1111/mmi.12981.

Ford, C., Shah, R. Maeda, M. and Gagneux, S. (2013). *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nature Genetics*. **45**(7). 784–790. doi:10.1038/ng.2656.

Ford, C., Yusim, K., Ioerger, T., Feng, S. Chase, M., Greene, M., … Fortune, S. (2012). *Mycobacterium tuberculosis* – Heterogeneity revealed through whole genome sequencing. *Tuberculosis*. **92**(3). 194–201. doi.org/10.1016/J.TUBE.2011.11.003.

Gandhi, N. R., Nunn, P., Dheda, K. Schaaf, H. S., Zignol, M., van Soolingen, D., … Bayona, J. (2010). Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *The Lancet*. **375**(9728), 1830–1843. doi.org/10.1016/S0140-6736(10)60410-2.

Goldberg, D. E., Siliciano, R. F. and Jacobs, W. R. (2012). Outwitting Evolution: Fighting Drug-Resistant TB, Malaria, and HIV. *Cell*. **148**(6). 1271–1283. doi.org/10.1016/J.CELL.2012.02.021

Gordin, F. M., Cohn, D. L., Sullam, P. M., Schoenfelder, J. R., Wynne, B. A. and Horsburgh, C. R. (1997). Early manifestations of disseminated *Mycobacterium avium* complex disease: a prospective evaluation. *The Journal of Infectious Diseases*. **176**(1). 126–132.

Grzybowski, S. (1991). Natural history of tuberculosis. Epidemiology. *Bulletin of the International Union against Tuberculosis and Lung Disease*. **66**(4).193–194.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* **29**(8). 1072–1075. doi.org/10.1093/bioinformatics/btt086.

Harries, A. D., Kumar, A. M. V., Satyanarayana, S. Lin, Y., Zachariah, R. Lönnroth, K. and Kapur, A. (2015). Diabetes mellitus and tuberculosis: programmatic management issues. *The International Journal of Tuberculosis and Lung Disease: The Official Journal of the International Union against Tuberculosis and Lung Disease*. **19**(8), 879–886. doi.org/10.5588/ijtld.15.0069.

Harries, A. D., Lin, Y., Satyanarayana, S., Lönnroth, K. Li, L., Wilson, N., … Kapur, A. (2011). The looming epidemic of diabetes-associated tuberculosis: learning lessons from HIV-associated tuberculosis. *The International Journal of Tuberculosis and Lung Disease*. **15**(11). 1436–1445. doi.org/10.5588/ijtld.11.0503.

Hatta, M., Sultan, A. R., Tandirogang, N., Masjud, M. and Yadi, N. (2010). Detection and identification of mycobacteria in sputum from suspected tuberculosis patients. *BMC Research Notes*. **3**(1). 72. doi.org/10.1186/1756-0500-3-72

Hazbón, M. H., Brimacombe, M. Bobadilla del Valle, M., Cavatore, M. Guerrero, M. I., Varma-Basil, M., … Alland, D. (2006a). Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobacterium tuberculosis.* *Antimicrobial Agents and Chemotherapy*. **50**(8). 2640–2649. doi.org/10.1128/AAC.00112-06

Hazbón, M. H., Brimacombe, M., Bobadilla del Valle, M., Cavatore, M., Guerrero, M. I., Varma-Basil, M., … Alland, D. (2006b). Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*. **50**(8). 2640–2649. doi.org/10.1128/AAC.00112-06

Henkle, E., Aksamit, T., Barker, A., Daley, C. L., Griffith, D. Leitman, P., … Winthrop, K. L. (2016). Patient-Centered Research Priorities for Pulmonary Nontuberculous Mycobacteria (NTM) Infection. An NTM Research Consortium Workshop Report. *Annals of the American Thoracic Society*. **13**(9). S379–S384. doi.org/10.1513/AnnalsATS.201605-387WS

Hershberg, R., Lipatov, M., Small, P. M. Sheffer, H., Niemann, S., Homolka, S., … Gagneux, S. (2008). High Functional Diversity in *Mycobacterium tuberculosis* Driven by Genetic Drift and Human Demography. *PLoS Biology*. **6**(12). e311. doi.org/10.1371/journal.pbio.0060311

Homolka, S., Niemann, S., Russell, D. G. and Rohde, K. H. (2010). Functional Genetic Diversity among *Mycobacterium tuberculosis* Complex Clinical Isolates: Delineation of Conserved Core and Lineage-Specific Transcriptomes during Intracellular Survival. *PLoS Pathogens*. **6**(7). e1000988. doi.org/10.1371/journal.ppat.1000988

Hoza, A. S., Lupindu A. M., Mfinanga S. G., Moser, I. and König, B. (2016a) The role of non tuberculous mycobacteria in the diagnosis, management and quantifying risks of tuberculosis in Tanga, Tanzania. *Tanzania Journal of Health Research*. **8**(2) 18:1–9.

Hoza, A. S., Mfinanga, S. G., Rodloff, A. C., Moser, I. and König, B. (2016). Increased isolation of nontuberculous mycobacteria among TB suspects in Northeastern, Tanzania: public health and diagnostic implications for control programmes. *BMC Research Notes*. **9**(1). 109. doi.org/10.1186/s13104-016-1928-3.

Hoza, A. S., Mfinanga, S. G., Moser, I. and König, B. (2017). Isolation, biochemical and molecular identification of Nocardia species among TB suspects in northeastern, Tanzania; a forgotten or neglected threat? *BMC Infectious Diseases*. **17**(1), 407. doi.org/10.1186/s12879-017-2520-8.

Jensen, A. V., Jensen, L., Faurholt-Jepsen, D., Aabye, M. G., Praygod, G., Kidola, J., … Andersen, A. B. (2013). The Prevalence of Latent *Mycobacterium tuberculosis* Infection Based on an Interferon-γ Release Assay: A Cross-Sectional Survey among Urban Adults in Mwanza, Tanzania. *PLoS ONE*. **8**(5). e64008. doi.org/10.1371/journal.pone.0064008

Johnson, M. M. and Odell, J. A. (2014). Nontuberculous mycobacterial pulmonary infections. *Journal of Thoracic Disease*. 6(3). 210–220. doi.org/10.3978/j.issn.2072-1439.2013.12.24.

Juréen, P., Werngren, J., Toro, J. C. and Hoffner, S. (2008). Pyrazinamide resistance and pncA gene mutations in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*. **52**(5). 1852–1854. doi.org/10.1128/AAC.00110-08

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., … van Embden, J. (1997). Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *Journal of Clinical Microbiology*. **35**(4). 907–914.

Katale, B. Z., Mbugi, E. V., Kendal, S., Fyumagwa, R. D., Kibiki, G. S., Godfrey-Faussett, P., … Matee, M. I. (2012). Bovine tuberculosis at the human-livestock-wildlife interface: Is it a public health problem in Tanzania? A review. *Onderstepoort Journal of Veterinary Research*. **79**(2). 8 pages. doi.org/10.4102/ojvr.v79i2.463.

Katale, B. Z., Mbugi, E. V., Siame, K. K., Keyyu, J. D., Kendall, S., Kazwala, R. R., … Matee, M. I. (2017). Isolation and Potential for Transmission of *Mycobacterium bovis* at Human-livestock-wildlife Interface of the Serengeti Ecosystem, Northern Tanzania. *Transboundary and Emerging Diseases*. **64**(3). 815–825. doi.org/10.1111/tbed.12445

Kendall, B. and Winthrop, K., (2013). Update on the Epidemiology of Pulmonary Nontuberculous Mycobacterial Infections. *Seminars in Respiratory and Critical Care Medicine*. **34**(01). 087–094. doi.org/10.1055/s-0033-1333567

Khan, J., Irfan, M., Zaki, A., Beg, M., Hussain, S. and Rizvi, N. (2006). Knowledge, attitude and misconceptions regarding tuberculosis in Pakistani patients. *Journal of Pakistan Medical Association*. **56**(5):211-4.

Kilale, A. M., Ngadaya, E., Muhumuza, J., Kagaruki, G. B., Lema, Y. L., Ngowi, B. J., … Hinderaker, S. G. (2016). Who Has Mycobacterial Disease? A Cross Sectional Study in Agropastoral Communities in Tanzania. *PLOS ONE*. **11**(5). e0153711. doi.org/10.1371/journal.pone.0153711

Kim, B. J., Hong, S. H., Kook, Y. H. and Kim, B. J. (2013a). Molecular Evidence of Lateral Gene Transfer in rpoB Gene of *Mycobacterium yongonense* Strains via Multilocus Sequence Analysis. *PLoS ONE*. **8**(1). e51846. doi.org/10.1371/journal.pone.0051846.

Kim, B. J., Kim, B. R., Kook, Y. H. and Kim, B. J. (2017). Role of the DNA Mismatch Repair GeneMutS4in Driving the Evolution of *Mycobacterium yongonense* Type I via Homologous Recombination. *Frontiers in Microbiology*. **8**. 2578. doi.org/10.3389/fmicb.2017.02578

Kim, B. J., Kim, B. R., Lee, S. Y., Kim, G. N., Kook, Y. H. and Kim, B. J. (2016). Molecular Taxonomic Evidence for Two Distinct Genotypes of *Mycobacterium yongonense* via Genome-Based Phylogenetic Analysis. *PLOS ONE*. **11**(3). e0152703. doi.org/10.1371/journal.pone.0152703

Kim, B. J., Kim, B. R., Lee, S. Y., Seok, S. H., Kook, Y. H. and Kim, B. J. (2013b). Whole-Genome Sequence of a Novel Species, *Mycobacterium yongonense* DSM 45126T. *Genome Announcements*. **1**(4). doi.org/10.1128/genomeA.00604-13.

Kim, B. J., Math, R. K., Jeon, C. O., Yu, H. K., Park, Y. G., Kook, Y. H. and Kim, B. J. (2013b). *Mycobacterium yongonense* sp. nov., a slow-growing non-chromogenic species closely related to *Mycobacterium intracellulare*. *International Journal of Systematic And Evolutionary Microbiology*. **63**(Pt 1). 192–199. doi.org/10.1099/ijs.0.037465-0

Klein, K. and Yang, Z. (2014). Comparison of ambient air survival of *Mycobacterium tuberculosis* clinical strains associated with different epidemiological phenotypes. *International Journal of Mycobacteriology*, ***3***(3), 211–213. doi.org/10.1016/J.IJMYCO.2014.04.002

Knechel, N. A. (2009). Tuberculosis: pathophysiology, clinical features, and diagnosis. *Critical Care Nurse*. **29**(2). 34–43. quiz 44. doi.org/10.4037/ccn2009968.

Koh, W. J., Yu, C. M., Suh, G. Y., Chung, M. P., Kim, H., Kwon, O. J., … Lee, K. S. (2006). Pulmonary TB and NTM lung disease: comparison of characteristics in patients with AFB smear-positive sputum. *The International Journal of Tuberculosis and Lung Disease.* **10**(9).1001-7.

Köser, C. U., Feuerriegel, S., Summers, D. K., Archer, J. A. C. and Niemann, S. (2012). Importance of the genetic diversity within the *Mycobacterium tuberculosis* complex for the development of novel antibiotics and diagnostic tests of drug resistance. *Antimicrobial Agents and Chemotherapy*. **56**(12). 6080–6087. doi.org/10.1128/AAC.01641-12.

Kuan, C. S., Chan, C. L., Yew, S. M., Toh, Y. F., Khoo, J. S., Chong, J., … Levy, S. (2015). Genome Analysis of the First Extensively Drug-Resistant (XDR) *Mycobacterium tuberculosis* in Malaysia Provides Insights into the Genetic Basis of Its Biology and Drug Resistance. *PLOS ONE*. **10**(6). e0131694. doi.org/10.1371/journal.pone.0131694.

Kwak, N., Park, J., Kim, E. Lee, C. H., Han, S. K. and Yim, J. J. (2017). Treatment Outcomes of *Mycobacterium avium* Complex Lung Disease: A Systematic Review and Meta-analysis. *Clinical Infectious Diseases*. **65**(7). 1077–1084. doi.org/ 10.1093/cid /cix517.

Lacoma, A., Garcia-Sierra, N., Prat, C. Ruiz-Manzano, J., Haba, L., Rosés, S., … Domínguez, J. (2008). GenoType MTBDRplus assay for molecular detection of rifampin and isoniazid resistance in *Mycobacterium tuberculosis* strains and clinical samples. *Journal of Clinical Microbiology*. **46**(11). 3660–3667. doi.org/10.1128/JCM.00618-08.

Lange, C., Abubakar, I., Alffenaar, J. W. C., Bothamley, G., Caminero, J. A., … Cirillo, D.M. (2014). Management of patients with multidrug-resistant/extensively drug-resistant tuberculosis in Europe: a TBNET consensus statement. *The European Respiratory Journal*. **44**(1). 23–63. doi.org/10.1183/09031936.00188313.

Laurenzo, D. and Mousa, S. A. (2011). Mechanisms of drug resistance in *Mycobacterium tuberculosis* and current status of rapid molecular diagnostic testing. *Acta Tropica, 119*(1), 5–10. https://doi.org/10.1016/J.ACTATROPICA.2011.04.008

Lee, I., Ouk K. Y., Park, S. C. and Chun, J. (2016). OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology*. **66**(2). 1100–1103. doi.org/10.1099/ijsem.0.000760

Lee, J. S., Kim, E. C., Joo, S. I., Lee, S. M., Yoo, C. G., Kim, Y. W., … Yim, J. J. (2008). The Incidence and Clinical Implication of Sputum with Positive Acid-Fast Bacilli Smear But Negative in Mycobacterial Culture in a Tertiary Referral Hospital in South Korea. *Journal of Korean Medical Science*. **23**(5). 767. doi.org/10.3346/jkms.2008.23.5.767

Lee, R. E. B., Li, W., Chatterjee, D. and Lee, R. E. (2004). Rapid structural characterization of the arabinogalactan and lipoarabinomannan in live mycobacterial cells using 2D and 3D HR-MAS NMR: structural changes in the arabinan due to ethambutol treatment and gene mutation are observed. *Glycobiology*. **15**(2). 139–151. doi.org/10.1093/glycob/cwh150

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

Li, H., Handsaker, B., Wysoker, A. Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**(16). 2078–2079. doi.org/10.1093/bioinformatics/btp352.

Lingaraju, S., Rigouts, L., Gupta, A., Lee, J., Umubyeyi, A. N., Davidow, A. L., … Safi, H. (2016). Geographic Differences in the Contribution of ubiA Mutations to High-Level Ethambutol Resistance in *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*. **60**(7). 4101–4105. doi.org/10.1128/AAC.03002-15.

Lo, H. Y., Yang, S. L., Lin, H. H., Bai, K. J., Lee, J. J., Lee, T. I. and Chiang, C. Y. (2016). Does enhanced diabetes management reduce the risk and improve the outcome of tuberculosis? *The International Journal of Tuberculosis and Lung Disease.* **20**(3). 376–382. doi.org/10.5588/ijtld.15.0654

Lockman, S., Sheppard, J. D., Braden, C. R., Mwasekaga, M. J., Woodley, C. L., Kenyon, T. A., … Tappero, J. W. (2001). Molecular and conventional epidemiology of *Mycobacterium tuberculosis* in Botswana: a population-based prospective study of 301 pulmonary tuberculosis patients. *Journal of Clinical Microbiology*. **39**(3). 1042–1047. doi.org/10.1128/JCM.39.3.1042-1047.2001.

Lönnroth, K., Roglic, G. and Harries, A. D. (2014). Improving tuberculosis prevention and care through addressing the global diabetes epidemic: from evidence to policy and practice. *The Lancet. Diabetes and Endocrinology*. **2**(9). 730–739. doi.org/10.1016/S2213-8587(14)70109-3.

Lu, W., Feng, Y., Wang, J. and Zhu, L. (2016). Evaluation of MTBDR plus and MTBDRsl in Detecting Drug-Resistant Tuberculosis in a Chinese Population. *Disease Markers*. 9 pages. doi.org/10.1155/2016/2064765.

Maiga, M., Siddiqui, S., Diallo, S., Diarra, B., Traor, B., Shea, Y. R., … Tounkara, A. (2012). Failure to recognize nontuberculous mycobacteria leads to misdiagnosis of chronic pulmonary tuberculosis. *PLoS ONE*. **7**(5). doi.org/10.1371/journal.pone.0036902.

Manalebh, A., Demissie, M., Mekonnen, D., Abera, B. and Kumar, P. (2015). The Quality of Sputum Smear Microscopy in Public-Private Mix Directly Observed Treatment Laboratories in West Amhara Region, Ethiopia. *PLOS ONE*. **10**(4). e0123749. doi.org/10.1371/journal.pone.0123749.

Martínez, J. L., Baquero, F. and Andersson, D. I. (2011). Beyond serial passages: new methods for predicting the emergence of resistance to novel antibiotics. *Current Opinion in Pharmacology*. **11**(5). 439–445. doi.org/10.1016/j.coph.2011.07.005.

McEvoy, C. R. E., Cloete, R., Müller, B., Schürch, A. C., van Helden, P. D., Gagneux, S., … Katoch, V. (2012). Comparative Analysis of *Mycobacterium tuberculosis* pe and ppe Genes Reveals High Sequence Variation and an Apparent Absence of Selective Constraints. *PLoS ONE*. **7**(4). e30593. doi.org/10.1371/journal.pone.0030593.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*. **20**(9). 1297–1303. doi.org/10.1101/gr.107524.110

McMurry, L. M., McDermott, P. F. and Levy, S. B. (1999). Genetic evidence that InhA of *Mycobacterium smegmatis* is a target for triclosan. *Antimicrobial Agents and Chemotherapy*. **43**(3). 711–713. DOI: 10.1128/AAC.43.3.711.

McNerney, R. and Daley, P. (2011). Towards a point-of-care test for active tuberculosis: obstacles and opportunities. *Nature Reviews Microbiology*. **9**(3). 204–213. doi.org/10.1038/nrmicro2521

Mdluli, K., Slayden, R. A., Zhu, Y., Ramaswamy, S., Pan, X., Mead, D., … Barry, C. E. (1998). Inhibition of a *Mycobacterium tuberculosis* β-Ketoacyl ACP Synthase by Isoniazid. *Science*. **280**(5369).

Merker, M., Blin, C., Mona, S. and Duforet-Frebourg, N. (2015). Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature*. **47**(3). 242-9. doi: 10.1038/ng.3195.

Mfinanga, G. S., Ngadaya, E., Mtandu, R., Mutayoba, B., Basra, D., Kimaro, G., … Kitua, A. Y. (2008). The quality of sputum smear microscopy diagnosis of pulmonarytuberculosis in Dar es Salaam, Tanzania. *Tanzania Health Research*. **9**(3):164-8.

Mfinanga, S. G., Warren, R. M., Kazwala, R., Ngadaya, E., Kahwa, A., Kazimoto, T., … Cleaveland, S. (2014). Genetic profile of *Mycobacterium tuberculosis* and treatment outcomes in human pulmonary tuberculosis in Tanzania. *Tanzania Journal of Health Research*. **16**(2). doi.org/10.4314/thrb.v16i2.1

Ministry of Health and Social Welfare (MoHSW). (2012). National Tuberculosis and leprosy Program. Operational Guidelines for the Management of Drug Resistant TB in Tanzania. First Edition 2012a.

Ministry of Health and Social Welfare (MoHSW). (2013). The First National Tuberculosis Prevalence Survey in the United Republic of Tanzania Final Report.

Ministry of Health and Social Welfare (MoHSW). The United Republic of Tanzania. (2009). Health Sector Strategic Plan III July 2009 – June 2015: "Partnership for Delivering the MDGs.

Ministry of Health and Social Welfare (MoHSW). The United Republic of Tanzania. (2016). *National Policy Guidelines for Collaborative TB /HIV Activities*.

Ministry of Health and Social Welfare (MoHSW). The United Republic of Tanzania. (2009) Health Sector Strategic Plan III July 2009–June 2015: "Partnership for Delivering the MDGs; 2009

Ministry Of Health Community Development Gender Elderly and Children (MoHCDGEC), The United Republic of Tanzania (2016). The National Tuberculosis and Leprosy Programme Annual report for 2016.

Mnyambwa N.P., Kim, D. J., Ngadaya, E. S., Kazwala, R., Petrucka, P. and Mfinanga, S. G. (2017a). Clinical implication of novel drug resistance-conferring mutations in resistant tuberculosis. *European Journal of Clinical Microbiology and Infectious Diseases*. **36**(11). 2021–2028. doi.org/10.1007/s10096-017-3027-3

Mnyambwa, P. N., Kim, D. J., Ngadaya, E. S., Kazwala, R., Petrucka, P. and Mfinanga, S. G. (2017b). Assessment of sputum smear-positive but culture-negative results among newly diagnosed pulmonary tuberculosis patients in Tanzania. *International Journal Of General Medicine.* 199–205.

Mnyambwa, N. P., Lekule, I., Ngadaya, E. S., Kimaro, G., Petrucka, P., Kim, D. J., … Mfinanga, S. G. (2018). Assessment of GeneXpert GxAlert platform for multi-drug resistant tuberculosis diagnosis and patients' linkage to care in Tanzania. *BMC Research Notes*. **11**(1). 121. doi.org/10.1186/s13104-018-3235-7.

Möller, M. and Hoal, E. G. (2010). Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. *Tuberculosis*. **90**(2). 71–83. https://doi.org/10.1016/j.tube.2010.02.002

Mpagama, S. G., Heysell, S. K., Ndusilo, N. D., Kumburu, H. H., Lekule, I. A., Kisonga, R. M., … Kibiki, G. S. (2013). Diagnosis and Interim Treatment Outcomes from the First Cohort of Multidrug-Resistant Tuberculosis Patients in Tanzania. *PLoS ONE*. **8**(5). e62034. doi.org/10.1371/journal.pone.0062034.

Müller, B., Borrell, S., Rose, G. and Gagneux, S. (2013). The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Trends in Genetics*. **29**(3). 160–169. doi.org/10.1016/j.tig.2012.11.005.

Murray, C. J., Styblo, K. and Rouillon, A. (1990). Tuberculosis in developing countries: burden, intervention and cost. *Bulletin of the International Union against Tuberculosis and Lung Disease*. **65**(1). 6–24.

Musser, J. M., Amin, A. and Ramaswamy, S. (2000). Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: Evidence of limited selective pressure. *Genetics*. **155**(1). 7–16.

National AIDS Control Program (NACP). (2013). *The United Republic of Tanzania. National Comprehensive Guidelines for HIV Testing and Counselling*.

Nagu, T. J., Aboud, S., Mwiru, R., Matee, M. I., Rao, M., Fawzi, W. W., … Mugusi, F. (2017). Tuberculosis associated mortality in a prospective cohort in Sub Saharan Africa: Association with HIV and antiretroviral therapy. *International Journal of Infectious Diseases : IJID : Official Publication of the International Society for Infectious Diseases*. **56**. 39–44. doi.org/10.1016/j.ijid.2017.01.023

National Tuberculosis and Leprosy Programme (NTLP). (2015). United Republic of Tanzania Ministry of Health and Social Welfare National Tuberculosis and Leprosy Programme Manual for the Management of Tuberculosis and Leprosy.

Pai, M. and Schito, M. (2015). Tuberculosis diagnostics in 2015: Landscape, priorities, needs, and prospects. *Journal of Infectious Diseases.* **211**(Suppl 2). S21–S28. doi.org/10.1093/infdis/jiu803.

Park, Y. K., Bai G. H. and Kim, S. J. (2000). Restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolated from countries in the western pacific region. *Journal of Clinical Microbiology*. **38**(1). 191–197.

Parsons, L. M., Somoskövi, A., Gutierrez, C., Lee, E., Paramasivan, C. N., Abimiku, A.,… Nkengasong, J. (2011). Laboratory diagnosis of tuberculosis in resource-poor Countries: Challenges and opportunities. *Clinical Microbiology Reviews*. **24**(2). 314–350. doi.org/10.1128/CMR.00059-10.

Perdigão, J., Macedo, R., Machado, D., Silva, C., Jordão, L., Couto, I., … Portugal, I. (2014). GidB mutation as a phylogenetic marker for Q1 cluster *Mycobacterium tuberculosis* isolates and intermediate-level streptomycin resistance determinant in Lisbon, Portugal. *Clinical Microbiology and Infection*. **20**(5). O278–O284. doi.org/10.1111/1469-0691.12392.

Potdar, P. and Thakur, P. (2013). Development of Sequence Based Molecular Diagnostic Test to Evaluate MDR and XDR in *M. tuberculosis* Patients from Western India. *American Journal of Infectious Diseases and Microbiology*. **1**(3). 50–58. doi.org/10.12691/ajidm-1-3-3.

Rahman, S. A., Singh, Y., Kohli, S., Ahmad, J., Ehtesham, N. Z., Tyagi, A. K. and Hasnain, S. E. (2014). Comparative analyses of nonpathogenic, opportunistic, and totally pathogenic mycobacteria reveal genomic and biochemical variabilities and highlight the survival attributes of *Mycobacterium tuberculosis*. *MBio*. **5**(6), e02020. doi.org/10.1128/mBio.02020-14.

Raju, R. M., Raju, S. M., Zhao, Y. and Rubin, E. J. (2016). Leveraging advances in tuberculosis diagnosis and treatment to address nontuberculous mycobacterial disease. *Emerging Infectious Diseases*. **22**(3). 365–369. doi.org/10.3201/eid2203.151643

Ramaswamy, S. and Musser, J. M. (1998). Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tubercle and Lung Disease*. **79**(1). 3–29. doi.org/10.1054/tuld.1998.0002

Ribeiro, S., Gomes, L. and Amaral, E. (2014). *Mycobacterium tuberculosis* strains of the modern sublineage of the Beijing family are more likely to display increased virulence than strains of the ancient sublineage. *Journal of Clinical Microbilogy*. **52**(7):2615-24. doi: 10.1128/JCM.00498-14.

Richter, E., Rüsch-Gerdes, S. and Hillemann, D. (2006). Evaluation of the GenoType Mycobacterium assay for identification of mycobacterial species from cultures. *Journal of Clinical Microbiology*. **44**(5). 1769–1775. doi.org/10.1128/JCM.44.5.1769-1775.2006.

Riojas, M. A., McGough, K. J., Rider-Riojas, C. J., Rastogi, N. and Hazbón, M. H. (2018). Phylogenomic analysis of the species of the *Mycobacterium tuberculosis* complex demonstrates that *Mycobacterium africanum*, *Mycobacterium bovis*, *Mycobacterium caprae*, *Mycobacterium microti* and *Mycobacterium pinnipedii* are later heterotypic synonyms of *Mycobacterium tuberculosis*. *International Journal of Systematic and Evolutionary Microbiology*. **68**(1). 324–332. doi.org/10.1099/ijsem.0.002507.

Ryu, Y. J. (2015). Diagnosis of pulmonary tuberculosis: Recent advances and diagnostic algorithms. *Tuberculosis and Respiratory Diseases*. **78**(2), 64–71. doi.org/10.4046/trd.2015.78.2.64

Safi, H., Lingaraju, S., Amin, A. Kim, S., Jones, M. Holmes, M., … Alland, D. (2013). Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-β-D-arabinose biosynthetic and utilization pathway genes. *Nature Genetics*. **45**(10). 1190–1197. doi.org/10.1038/ng.2743

Sanz, J., Navarro, J., Arbués, A., Martín, C., Marijuán, P. C. and Moreno, Y. (2011). The Transcriptional Regulatory Network of *Mycobacterium tuberculosis*. *PLoS ONE*. **6**(7), e22178. doi.org/10.1371/journal.pone.0022178.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. **30**(14), 2068–2069. doi.org/10.1093/bioinformatics/btu153

Sekiguchi, J., Miyoshi-Akiyama, T., Augustynowicz-Kopeć, E., Zwolska, Z., Kirikae, F., Toyota, E., … Kirikae, T. (2007). Detection of multidrug resistance in *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*. **45**(1). 179–192. doi.org/10.1128/JCM.00750-06.

Shahraki, A. H., Heidarieh, P., Bostanabad, S. Z., Khosravi, A. D., Hashemzadeh, M., Khandan, S., … Mirsaeidi, M. (2015). "Multidrug-resistant tuberculosis" may be nontuberculous mycobacteria. *European Journal of Internal Medicine*. **26**(4). 279–284. doi.org/10.1016/j.ejim.2015.03.001.

Shaw, J. E., Sicree, R. A. and Zimmet, P. Z. (2010). Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Research and Clinical Practice*. **87**(1). 4–14. doi.org/10.1016/j.diabres.2009.10.007.

Silva Miranda, M., Breiman, A., Allain, S., Deknuydt, F. and Altare, F. (2012). The Tuberculous Granuloma: An Unsuccessful Host Defence Mechanism Providing a Safety Shelter for the Bacteria? *Clinical and Developmental Immunology*. **2012**. 1–14. doi.org/10.1155/2012/139127.

Sims, D., Sudbery, I., Ilott, N. E., Heger, A. and Ponting, C. P. (2014). Genomics is extending its reach into diverse fields of biomedical research from agriculture to clinical diag- nostics. Despite sharp falls in recent years. *Nature Reviews.* https://doi.org/10.1038/nrg3642

Singh, J., Sankar, M., Kumar, P., Couvin, D. and Rastogi, N. (2015). Genetic diversity and drug susceptibility profile of *Mycobacterium tuberculosis* isolated from different regions of India. *Journal of Infection*. **71**(2):207-19. doi: 10.1016/j.jinf.2015.04.028.

Singh, P., Jain, A., Dixit, P., Prakash, S., Jaiswal, I., Venkatesh, V. and Singh, M. (2014). A novel gyrB gene mutation in fluoroquinolone resistant clinical isolates of *Mycobacterium tuberculosis*. *BMC Infectious Diseases*. **14**(Suppl 3). O14. doi.org/10.1186/1471-2334-14-S3-O14

Small, P. M., Hopewell P. C., Singh, S. P., Paz, A., Parsonnet, J., Ruston, D. C., … Schoolnik, G. K. (1994a). The Epidemiology of Tuberculosis in San Francisco. A Population-Based Study Using Conventional and Molecular Methods. *New England Journal of Medicine*. **330**(24). 1703–1709. doi.org/10.1056/NEJM199406163302402.

Small, P. M., Hopewell, P. C., Singh, S. P., Paz, A., Parsonnet, J. Ruston, D. C., … Schoolnik, G. K. (1994b). The Epidemiology of Tuberculosis in San Francisco. A Population-Based Study Using Conventional and Molecular Methods. *New England Journal of Medicine*. **330**(24). 1703–1709. doi.org/10.1056/NEJM199406163302402

Sola, C., Filliol, I., Legrand, E., Lesjean, S., Locht, C., Supply, P. and Rastogi, N. (2003). Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*. **3**(2). 125–133.

Sonnenberg, P., Murray, J., Glynn, J. R., Shearer, S., Kambashi, B. and Godfrey, P. (2001). HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers. *The Lancet*. **358**(9294). 1687–1693. doi.org/10.1016/S0140-6736(01)06712-5.

Spies, F. S., Ribeiro, A. W., Ramos, D. F., Ribeiro, M. O., Martin, A., Palomino, J. C., … Zaha, A. (2011a). Streptomycin resistance and lineage-specific polymorphisms in *Mycobacterium tuberculosis* gidB gene. *Journal of Clinical Microbiology*. **49**(7). 2625–2630. doi.org/10.1128/JCM.00168-11.

Spies, F. S., Ribeiro, A., W. Ramos, D. F., Ribeiro, M. O., Martin, A., Palomino, J. C., … Zaha, A. (2011b). Streptomycin resistance and lineage-specific polymorphisms in *Mycobacterium tuberculosis* gidB gene. *Journal of Clinical Microbiology*. **49**(7). 2625–2630. doi.org/10.1128/JCM.00168-11.

Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., Whittam, T. S. and Musser, J. M. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Microbiology Communicated by Barry R. Bloom, Albert Einstein College of Medicine*. **94**. 9869–9874.

Sreevatsan, S., Pan, X. Zhang, Y., Kreiswirth, B. N. and Musser, J. M. (1997). Mutations associated with pyrazinamide resistance in pncA of *Mycobacterium tuberculosis* complex organisms. *Antimicrobial Agents and Chemotherapy*. **41**(3). 636–640.

Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D. and Locht, C. (2001). Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *Journal of Clinical Microbiology*. **39**(10). 3563–3571. doi.org/10.1128/JCM.39.10.3563-3571.2001

Supply, P., Warren, R. M., Bañuls, A. L. Lesjean, S., Van Der Spuy, G. D., Lewis, L. A., … Locht, C. (2003). Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. *Molecular Microbiology*. **47**(2). 529–538. doi.org/10.1046/j.1365-2958.2003.03315.x.

Tortoli, E., Mariottini, A., Pierotti, P., Simonetti, T. M. and Rossolini, G. M. (2013). *Mycobacterium yongonense* in pulmonary disease, Italy. *Emerging Infectious Diseases*. **19**(11). 1902–1904. doi.org/10.3201/eid1911.130911

Tortoli, E., Rindi, L., Garcia, M. J. Chiaradonna, P., Dei, R. Garzelli, C., … Scarparo, C. (2004). Proposal to elevate the genetic variant MAC-A, included in the *Mycobacterium avium* complex, to species rank as *Mycobacterium chimaera* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. **54**(4). 1277–1285. doi.org/10.1099/ijs.0.02777-0

Trébucq, A., Enarson, D., A. Chiang, C. Y., Deun, A., Van, Harries, A. D., Boillot, F., … Rieder, H. L. (2011). Xpert ® MTB/RIF for national tuberculosis programmes in low-income countries: when, where and how? *International Journal Tuberc Lung Diseases*. **15**(12). 1567–1571. doi.org/10.5588/ijtld.11.0392.

Tyrrell, F. C., Budnick, G. E., Elliott, T., Gillim-Ross, L., Hildred, M. V., Mahlmeister, P., … Starks, A. M. (2012). Probability of negative *Mycobacterium tuberculosis* complex cultures based on time to detection of positive cultures: a multicenter evaluation of commercial-broth-based culture systems. *Journal of Clinical Microbiology*. **50**(10). 3275–3282. doi.org/10.1128/JCM.01225-12

Uplekar, M., Weil, D., Lonnroth, K., Jaramillo, E., Lienhardt, C., Dias, H. M., … for WHO's Global TB Programme. (2015). WHO's new end TB strategy. *Lancet*. **385**(9979). 1799–1801. doi.org/10.1016/S0140-6736(15)60570-0.

van Ingen, J., Boeree, M. J., Kosters, K., Wieland, A., Tortoli, E., Dekhuijzen, P. N. R. and van Soolingen, D. (2009). Proposal to elevate *Mycobacterium avium* complex ITS sequevar MAC-Q to *Mycobacterium vulneris* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. **59**(9). 2277–2282. doi.org/10.1099/ijs.0.008854-0.

van Soolingen, D., Hermans, P. W., de Haas, P. E., Soll, D. R. and van Embden, J. D. (1991). Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *Journal of Clinical Microbiology*. **29**(11). 2578–2586.

Velayati, A. A., Masjedi, M. R., Farnia, P., Tabarsi, P., Ghanavi, J., ZiaZarifi, A. H. and Hoffner, S. E. (2009). Emergence of New Forms of Totally Drug-Resistant Tuberculosis Bacilli. *Chest*. **136**(2). 420–425. doi.org/10.1378/chest.08-2427

Vilchèze, C., Weisbrod, T. R., Chen, B., Kremer, L., Hazbón, M., H., Wang, F., … Jacobs, W. R. (2005). Altered NADH/NAD+ ratio mediates coresistance to isoniazid and ethionamide in mycobacteria. *Antimicrobial Agents and Chemotherapy*. **49**(2). 708–720. doi.org/10.1128/AAC.49.2.708-720.2005

Weyer, K., Mirzayev, F., Migliori, G. B., Van Gemert, W. D., Ambrosio, L., Zignol, M., … Raviglione, M. (2013). Rapid molecular TB diagnosis: evidence, policy making and global implementation of Xpert MTB/RIF. *The European Respiratory Journal*. **42**(1). 252–271. doi.org/10.1183/09031936.00157212

World Health Organization (WHO). (2009a). *Guidelines for surveillance of drug resistance in tuberculosis 5 th Edition*.

World Health Organization (WHO). (2009b). *Treatment of tuberculosis*.

World Health Organization (WHO). (2013). *Global tuberculosis report 2013*.

World Health Organization (WHO). (2014a). *Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis.*

World Health Organization (WHO). (2014b). *Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis*.

World Health Organization (WHO). (2015). *Global tuberculosis report 2015*.

World Health Organization (WHO). (2016). *Global Tuberculosis Report 2016*

World Health Organization (WHO). (2016). *Global report on diabetes*. *ISBN*. Geneva, Switzerland.

World Health Organization (WHO). (2016). Treatment guidelines for drug-resistant tuberculosis.

World Health Organization (WHO) (2009). *Guidelines for surveillance of drug resistance in tuberculosis, Fourth Edition*.

World Health Organization (WHO). (2013). *The use of molecular line probe assay for detection of resistance to second-line anti-tuberculosis drugs: Epert group meeting report Geneva 2013*.

World Health Organization (WHO). (2017). *Global tuberculosis report 2017*.

World Health Organization (WHO). (2014). (2014). *The End TB Strategy: Global strategy and targets for tuberculosis prevention, care and control after 2015*.

Wong, S., Lee, J., Kwak, H. and Via, L. (2011). Mutations in gidB confer low-level streptomycin resistance in *Mycobacterium tuberculosis*. *Antimicrobial Agents*. DOI: 10.1128/AAC.01814-10.

Xu, Z., Xiao, T., Li, Y., Yang, K., Tang, Y. and Bai, L. (2017). Reasons for Non-Enrollment in Treatment among Multi-Drug Resistant Tuberculosis Patients in Hunan Province, China. *PLOS ONE*. **12**(1), e0170718. doi.org/10.1371/journal.pone.0170718.

Yeung, M. W., Khoo, E., Brode, S. K., Jamieson, F. B., Kamiya, H., Kwong, J. C., … Sander, B. (2016). Health-related quality of life, comorbidities and mortality in pulmonary nontuberculous mycobacterial infections: A systematic review. *Respirology*. **21**(6). 1015–1025. doi.org/10.1111/resp.12767

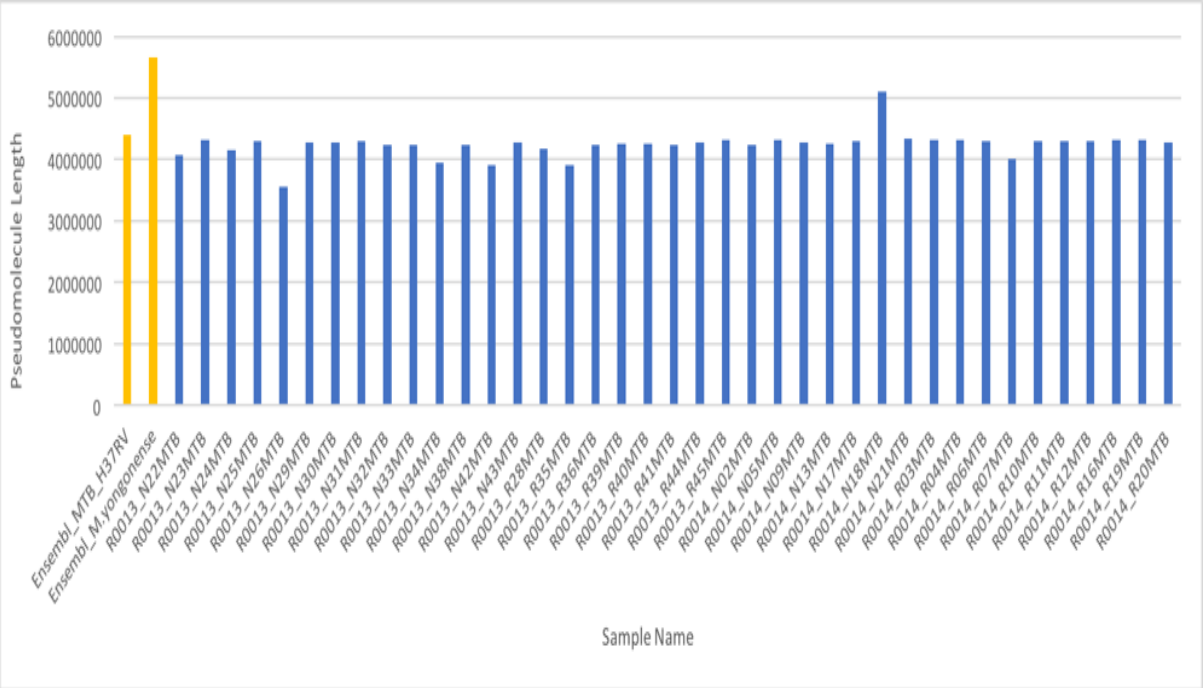
Zhang, H., Li, D., Zhao, L., Fleming, J., Lin, N., Wang, T., … Bi, L. (2013). Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. *Nature Genetics*. **45**(10). 1255–1260. doi.org/10.1038/ng.2735

Zhang, S., Chen, J., Shi, W., Liu, W., Zhang, W. and Zhang, Y. (2013). Mutations in panD encoding aspartate decarboxylase are associated with pyrazinamide resistance in *Mycobacterium tuberculosis*. *Emerging Microbes and Infections*. **2**(6). e34. doi.org/10.1038/emi.2013.38

Zink, A. R., Sola, C., Reischl, U., Grabner, W., Rastogi, N., Wolf, H. and Nerlich, A. G. (2003). Characterization of *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *Journal of Clinical Microbiology*. **41**(1), 359–367. doi.org/10.1128/JCM.41.1.359-367.2003.

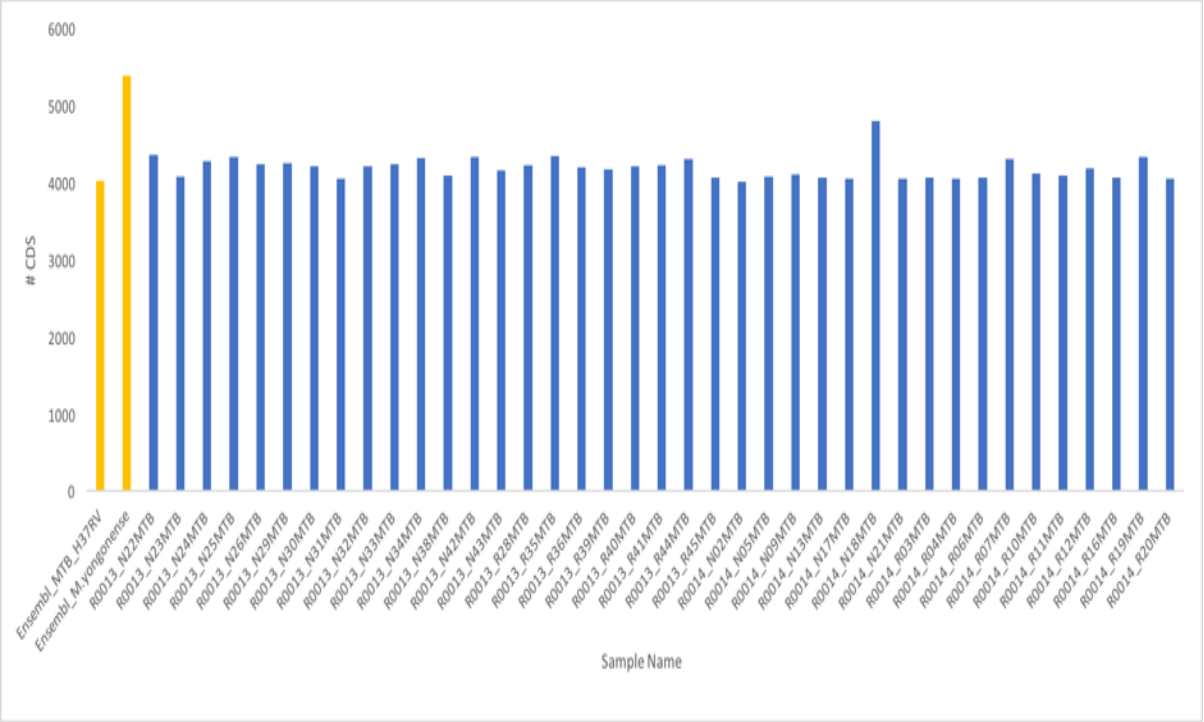
APPENDICES

Appendix 1: Pseudomolecule length of each individual genome analysed.



Blue colour bars shows pseudomolecule length of 39 samples, orange colour bar indicates Ensembl genome (*M. tuberculosis* H37RV and *M. yongonense* 05-1390^{T}) length

Appendix 2: Number of predicted coding sequence in each individual sample.



Blue colour indicates the number of predicted CDS sequence in pseudomolecule and orange indicates CDS sequences in Ensembl genome (*M. tuberculosis* H37Rv and *M. yongonense* 05-1390^{T}).

Appendix 3: Demographic information of MDR-TB patients and DST to first-line anti-TBs. CTB mean a test was contaminated

ID	Date	Sex F/M	INH 0.2/1.0 µg/ml	SM 5 µg/ml	RIF 4 µg/ml	EMB 2 µg/ml	Patient address
N1	2016	M	R	R	R	S	Zanzibar
N2	2016	M	R	R	R	R	Dar es salaam
N3	2016	M	R	S	R	R	Dar es salaam
N4	2016	M	R	R	R	R	Dar es salaam
N5	2016	M	R	R	R	R	Mtwara
N6	2016	M	R	S	R	S	Manyara
N7	2016	M	R	R	R	S	Dar es salaam
N9	2016	M	R	R	R	S	Dar es salaam
N10	2016	F	R	R	R	S	Dar es salaam
N11	2016	M	R	R	R	S	Kilimanjaro
N12	2016	F	R	S	R	S	Mbeya
R13	2016	M	R	S	R	R	Tanga
R16	2015	M	R	S	S	R	Dar es salaam
R17	2015	M	R	S	R	S	Iringa
R18	2015	M	S	S	S	S	Dar es salaam
N19	2016	M	R	R	R	S	Dar es salaam
N20	2016	F	R	R	R	R	Arusha
N21	2016	M	R	R	R	S	Dar es salaam
N22	2016	M	S	R	R	S	Arusha
N23	2016	M	R	S	R	S	Gaita
N24	2016	F	R	S	R	S	Lindi
R25	2015	M	R	R	S	R	Mtwara
R26	2015	M	R	R	R	R	Mwanza
R28	2015	M	R	R	R	R	Mwanza
R29	2015	F	R	R	R	R	Dar es salaam
R30	2015	F	R	R	R	R	Kilimanjaro
N31	2016	F	S	R	R	R	Dar es salaam
R32	2015	M	R	R	R	R	Kilimanjaro
R33	2015	M	R	R	R	S	Dar es salaam
R34	2015	M	R	S	R	S	Mtwara
R35	2015	F	R	R	S	S	Dar es salaam
R36	2016	M	S	S	R	S	Dar es salaam
R38	2015	M	R	R	R	R	Dar es salaam
R39	2016	F	R	R	R	R	Dar es salaam
N40	2015	M	R	R	R	R	Dar es salaam
N41	2016	M	R	S	R	CTB	Singida

N42	2016	M	S	S	R	S	Geita
N43	2016	M	R	R	S	R	Geita
R44	2015	M	R	S	R	S	Geita
R45	2015	M	S	S	R	R	Arusha