GENETIC DIVERSITY OF CLINICAL ISOLATES OF *Mycobacterium tuberculosis* CIRCULATING WITHIN WESTERN KENYA BETWEEN 2013 AND 2014.

BY

CHARLES K. CHELIMO

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN CELL AND MOLECUAR BIOLOGY.

DEPARTMENT OF ZOOLOGY

MASENO UNIVERSITY

DECLARATION

I declare that this is my original work and has not been presented to any other institution for any award.

Signature…….Date……

Charles K. Chelimo
(PG/MSC/00127/2011)

Supervisors Declaration:

This thesis has been submitted for examination with our approval as the supervisors:

Signature…….....Date……

Dr. Paul O. Ang'ienda, Ph.D.
Department of Zoology,
School of Biological and Physical Sciences,
Maseno University, Kenya

Signature……Date……..

Prof. Henry Nyamogoba, Ph.D.
Department of Microbiology,
School of Medicine,
Moi University, Kenya

ACKNOWLEDGEMENT

All glory and honour goes to the almighty God for His grace and mercies over my life throughout my study period. I would like to offer my sincere gratitude to my supervisors, Dr. Paul. O. Ang'ienda of Maseno University and Prof. Henry Nyamogoba of Moi University for their guidance and mentorship throughout the project period starting from the proposal development till the completion of the thesis writing. I owe much thank you, to all staff of Maseno University, among them Professor David Onyango for his mentorship and guidance during the course work.

I am much indebted to the staff of Moi Teaching and Referral Hospital (MTRH), Tuberculosis (TB) Laboratory for supporting me during the sample collection and preparation stage

I am also very much indebted to my employer, the MTRH for granting me permission to pursue my Master of Science degree at Maseno University. Much gratitude to the Director and Head of Laboratory service at MTRH. I wish also in a special way to thank my family, wife Gladys, sons Shadrack, Ignatius, Alfonse and daughter Cynthia for their tremendous support and encouragement during my study period.

Lastly, all the people of good will who assisted me in one way or another during the preparation of this thesis.

DEDICATION

I dedicate this thesis to my entire family, and to all that brought out the best of me in this life. May God bless you abundantly.

ABSTRACT

Tuberculosis (TB) is a pulmonary disease caused by *Mycobacterium tuberculosis* complex. It is a major global public health problem with an estimated 8 million incidences and 3 million mortalities annually. In Kenya, the prevalence rate of TB was 233 per 100, 000 persons in the year 2017. Majority of the counties in Western Kenya, have been noted to have a higher than national prevalence rate of tuberculosis infections. Previous studies have identified various species and strains of *Mycobacteria*. However, in many clinical setups in Kenya and other developing countries, it is assumed that the most identified species is *M. tuberculosis* complex (MTBC). This increases the chance of misdiagnosis when other species of *Mycobacteria* causes TB infection in humans. Human migration has been associated with spatial dissemination of foreign strains from their local evolutionary regions to different geographical regions and this increases the risk of transfer of strains associated with drug resistance to different regions. Knowledge on the genotypes and genetic diversity is needed in ensuring adequate prevention, control and treatment strategies. This study aimed at determining the genetic diversity of the clinical isolates of *Mycobacterium tuberculosis* in Western Kenya. Specifically, it identified the MTBC species and strains and their proportions in Western Kenya. Furthermore, it determined the genetic diversity of *M. tuberculosis* complex in Western Kenya. This was a cross-sectional study on 40 archived clinical isolates(between 2013 to 2014) from patients with confirmed tuberculosis following Lowenstein-Jensen (LJ) medium cultures. The study was conducted at Moi Teaching and Referral Hospital (a national referral hospital serving Western Kenya Counties), in Eldoret –Kenya. Mycobacterial deoxyribonucleic acids (DNA) was extracted using molecular grade water in an ultrasonicator bath before 12-loci MIRU-VNTR genotyping was performed to determine the species and strains of MTBC. Oligonucleotide primer sequences for amplification used were on MIRU 02, 04, 10, 16, 20, 23, 24, 26, 27, 31, 39 and 40 loci. Descriptive statistics techniques were used to describe the study population. Genetic data analysis was conducted on the MIRU-VNTR *plus* web server to display NJ Tree Dendrogram (for cluster analysis) and conduct similarity searches to identify the species. Allelic diversity was calculated using Hunters-Gaston diversity index (HGDI). The species identified in this population were *M. tuberculosis*, *M. africanum* and *M. bovis* and the strains identified in the *M. tuberculosis* species were: Beijing, LAM, URAL, Uganda 1 and EAI strains. However, some isolates had a mixture of more than one species, while others had unknown species (new/unassigned). The strains were grouped into five clusters of: Beijing (n=10), LAM (n=5), Uganda 1 (n=3), EAI (n=1) and Ural (n=1) with clustering rate of 56.4%. The study revealed that there is more than a single species and strains of *Mycobacterium* that causes tuberculosis in Western Kenya. These strains were highly diverse as seen in the high allelic diversity index of 0.808. There is need for more awareness among healthcare and other stakeholders on the existence of foreign species and strains in Western Kenya. Policy makers should adopt molecular detection techniques of the strains circulating in Western Kenya to boost prevention, control and treatment strategies. Further studies using 24 loci MIRU-VNTR should be conducted to build on findings

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

bp	base pair
CDR	Case detection rate
DLTLD	Division of Leprosy, Tuberculosis and Lung Diseases
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DOTS	Directly observed treatment short
DR	Direct repeat
HGDI	Hunter-Gaston diversity index
MDR-TB	Multi-Drug Resistant Tuberculosis
MIRU	Mycobacterium interspersed repetitive unit
MTRH	Moi Teaching and Referral Hospital
PCR	Polymerase chain reaction
PGRS	Polymorphic GC rich repeat sequences
RD	Region of Deletion
RFLP	Repetitive fragment length polymorphism
TB	Tuberculosis
TR	Tandem repeats
VNTR	Variable number tandem repeat
WHO	World Health Organization

OPERATIONAL DEFINITION OF TERMS

***Mycobacterium tuberculosis* complex** (MTBC): Mycobacteria belong to the genus Mycobacterium, which is the single genus within the family of Mycobacteriaceae.The complex is a combination of mycobacterium species and strains.

***Mycobacterium tuberculosis* species:**After the genus, we have the species. These include: *Mycobacterium tuberculosis complex* (MTBC) namely: *M. tuberculosis, M. bovis, M. africanum, Bacillus Calmette-Guerin (BCG), M. microti, M. carprae, M. pinnipedii and M. cannettii*

Mycobacterium tuberculosis* strains:This is the subspecies (families) of *M. tuberculosis.Mycobacterial species are traditionally differentiated based on phenotypic characteristics. These include:*Beijing, East African Indian (EAI), Delhi/Central Asian Strain (CAS), Latin American Mediterranean (LAM), Cameroon, TUR, Haarlem, URAL,Ghana, Uganda I and Uganda II.*

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background

Tuberculosis is considered one of the world's deadliest infectious diseases (Lopez-Avalos *etal*., 2017).According to the World Health Organization (WHO), an estimated 10.4 million individuals contracted tuberculosis with a mortality of 1.4 million in the year 2017 (WHO, 2018). In European countries such as Belgium, the incidence of tuberculosis in the year 2015 was 8.8 cases per 100,000 inhabitants, making it a low tuberculosis incidence country (Drobniewski *etal*., 2015). In the year 2017, the WHO report stated that there were 10.4 million new cases (incidences) of tuberculosis (WHO, 2018). The estimated tuberculosis incidence in the horn of Africa ranges from 65 cases per 100, 000 persons per year in Eritrea; 192 cases per 100, 000 persons in Ethiopia; 274 per 100, 000 persons in Somalia to 378 per 100, 000 persons in Djibouti (WHO, 2018).

Kenya is among the 30 high burden tuberculosis countries in the world(WHO, 2018). Tuberculosis is the fourth leading cause of mortality in Kenya with the most affected age group being those aged between 15-34 years (Kosgei *etal*., 2015; WHO, 2017). The estimated Kenyan national tuberculosis prevalence in 2015 was 233 per 100,000 individuals with an incidence of 81,518 (WHO, 2015). The findings of the Kenya Tuberculosis Prevalence Survey 2016 however revealed that the burden of TB is higher than previous estimation by the WHO representing an underestimation of up to 40% (NTLD-P, 2017). Data from the Ministry of Health survey showed that the prevalence stands at 558 people out of 100, 000 persons, up from the WHO figures of 233 in every 100,000. The incidence of TB is approximately 138,000 cases per year according to the national prevalence survey (NTLD-P, 2017). The significance of this data is that

about 20,000 cases of the disease go undetected thus reversing the efforts and investments towards the control of the disease (NTLD-P, 2017).

Mycobacterium tuberculosis is the main causative agent for tuberculosis disease (Coscolla & Gagneux, 2014). There are eight species of *Mycobacterium tuberculosis complex* (MTBC) namely: *M. tuberculosis, M. bovis, M. africanum, Bacillus Calmette-Guerin (BCG), M. microti, M. carprae, M. pinnipedii and M. cannettii* (Pankhurst et al., 2016). Even though majority of pulmonary tuberculosis are caused by *M. tuberculosis,* there is need for species identification of MTBC to properly identify the correct treatment combination and limit on drug resistance (Hoza, Mfinanga, Rodloff, Moser, & König, 2016). For instance, *M. bovis* has been found to be resistant to first line Pyrazinamide, while disseminated *M. bovis* BCG has been demonstrated as a complication following vaccination (Louw *etal*., 2006; Ramirez-Busby & Valafar, 2015). Mycobacteria are obligate aerobes, non-motile, rod-shaped bacilli (Thacker *etal*., 2014). The members of the genus *Mycobacterium* have characteristically unique compared to other bacterial genera. This is largely due to structural differences in cell wall composition. Mycobacterial cell walls contain higher contents of complex lipids (> 60% as opposed to approximately 5% and 20% in gram-positive and gram-negative organisms respectively) with a long chain (C 60 –C 90) of fatty (mycolic) acids (Kumar *etal*., 2007). These mycolic acids make the cell walls extremely hydrophobic and enhance resistance to desiccation, killing by disinfectants, staining with basic aniline dyes and penetration by many of the drugs that are used to treat infections caused by other bacteria (Teramoto *etal*., 2015). These unique features of Mycobacterial cell wall structure provide the basis for special laboratory considerations when performing direct stains from specimens, growing organisms in culture and determining species identification by molecular techniques (Kumar *etal*., 2007). The emergence of multidrug resistant strains of tuberculosis has

limited the available treatment options of tuberculosis creating a risk of an untreatable and fatal disease (Muthaiah *etal*., 2017).

Mycobacterium tuberculosis complex (MTBC) is classified based on strains into a phylogenetic framework. This strain classification has a long history that Gagneux and colleagues used comparative genomics to identify large sequence polymorphisms (LSPs) defining major lineages in the MTBC (Gagneux *etal*., 2006). This was later supported by the sequencing of 89 genes from 108 strains (Hershberg et al., 2008). The most commonly identified strains include: *Beijing, East African Indian (EAI), Delhi/Central Asian Strain (CAS), Latin American Mediterranean (LAM), Cameroon, TUR, Haarlem, URAL, Uganda I and Uganda II, Ghana,* and many others (Supply *etal*., 2013). Majority of the studies in species and strains have been conducted in Asian Countries and a few African countries (Azé *etal*., 2015; Klopper *etal*., 2013; Saifodine *etal*., 2014). In Tanzania, the EAI strain was found to be more predominant in Bunda while the was more predominant in Dar-es-Salaam (Mbugi et al., 2016). In a Kenyan study conducted in Nairobi, the CAS and Beijing strains were found to be the most predominant (Githui *etal*., 2004). There are limited current studies on the specific species and strains in the Western Kenya region which is a tuberculosis endemic area. This study therefore aimed to determine the specific species and strains of *M. tuberculosis* circulating in Western Kenya.

Molecular epidemiology studies are used to identify factors of tuberculosis distribution and transmission. This is often done using molecular characterization techniques such as spoligotyping and Mycobacterium Interspersed Repetitive Units – Variable Number Tandem Repeats (MIRU-VNTR) of genetic elements (Supply *etal*., 2002). Genetic lineages of *M. tuberculosis* strains have previously been determined based on the spoligotyping. The combination of spoligotyping and 12-loci MIRU has been successfully used in resource limited

settings to decipher potential transmission chains (Jagielski *etal*., 2016; Mbugi *etal*., 2016). This combination of 12-loci MIRU and spoligotyping allowed for the assessment of ongoing transmission underlying a total of 12-clusters containing 35 isolates. In Belgium, tuberculosis strain typing is done using two main available techniques, namely: 24-loci MIRU-VNTR and spoligotyping. Both techniques are polymerase chain reaction (PCR)-based methods. Spoligotyping detects the presence or absence of spacer sequences located between repetitive elements by PCR and Reverse Dot-Blot analysis (Supply *etal*., 2001). Despite the fact that spoligotyping is an easy technique; used alone, it lacks discriminatory power (Kremer *etal*., 2005). Majority of the *M. tuberculosis* genetic diversity studies conducted in Kenya, have adopted the spoligotyping technique, with limited use of MIRU-VNTR (Githui *etal*., 2004; Ogaro *etal*., 2012). This study therefore aimed to use MIRU-VNTR technique - which has a higher resolution and discriminatory power - to determine the molecular characteristics of the *M. tuberculosis* clinical isolates obtained in Western Kenya.

1.2 Problem Statement

Pulmonary tuberculosis is a major cause of morbidity and mortality globally affecting mainly people within the economically productive age of between 15 to 49 years (WHO, 2017). In most clinical setups, it is assumed that the commonly diagnosed MTBC species is the *M. tuberculosis*(Hoza, Lupindu, Mfinanga, Moser, & König, 2016). This presents the risk of missing out the other seven species namely: *M. bovis, M. africanum, Bacillus Calmette-Guerin (BCG), M. microti, M. carprae, M. pinnipedii and M. cannettii* that may also cause tuberculosis (Glaziou, Floyd, & Raviglione, 2018).This misdiagnosis could lead to an increase in tuberculosis associated morbidities and mortalities(Bloomfield et al., 2012). However, the identity of *M. tuberculosis* species currently circulating within Western Kenya is not known(Nduba, Van't

Hoog, Mitchell, Borgdorff, & Laserson, 2018; Nyamogoba & Mbuthia, 2018; Pavlinac et al., 2015). The various strains of *M. tuberculosis* complex have been associated with drug resistance(Alyamani et al., 2019; Berrada et al., 2016; Zhang et al., 2005). Despite this, there is limited documented data on the various strains and their proportions in Western Kenya. This could increase the risk of the occurrence and spread of multidrug resistant tuberculosis. Human migration has been associated with the dissemination of various strains in non-habitat areas(Gagneux et al., 2006). Majority of the strains have often been confined to specific geographical regions, however, there is a change in this trend(Allix-Béguec, Fauville-Dufaux, & Supply, 2008; Jagielski et al., 2016; Supply et al., 2013). Therefore, because the molecular characteristics and allelic diversity of *M. tuberculosis* complex in Western Kenya is not known; there is need to carry out molecular characterization of various *Mycobacterium tuberculosis* complex isolates in Western Kenya to inform control and treatment strategies of tuberculosis.

1.3 Objectives

1.3.1 General objective

To determine the genetic diversity of clinical isolates of *M. tuberculosis* circulating in Western Kenya between 2013 and 2014.

1.3.2 Specific objectives

1. To identify *M. tuberculosis complex* species in clinical isolates obtained from patients in Western Kenya between 2013 to 2014.
2. To determine the proportions of M. tuberculosis complex strains isolated from patients in Western Kenya between 2013 to 2014.

3. To determine the molecular characteristics and allelic diversity of *Mycobacterium tuberculosis complex* isolated from tuberculosis patients in Western Kenya.

1.4 Research questions

1. What are the various *M. tuberculosis complex* species in clinical isolates obtained from patients seen at Moi Teaching and Referral Hospital, Eldoret-Kenya between 2013 and 2014?
2. What are the different strains of *Mycobacterium tuberculosis complex* and their proportions isolated from patients seen at Moi Teaching and Referral Hospital, Eldoret-Kenya between 2013 and 2014?
3. What is the allelic diversity of *Mycobacterium tuberculosis* isolates circulating in western Kenya in 2013 and 2014?

1.5 Justification

Pulmonary diseases such as Tuberculosis have remained a major cause of increased disease burden and death among the human population globally and more so in the developing countries. Western Kenya, being in a developing country is tuberculosis endemic with a prevalence at 329 people out of 100, 000 persons in Busia County and 379 people out of 100, 000 persons in Kisumu County. Because of the contagious nature of tuberculosis and its high cost of management, control strategies must be put in place. One of the ways of developing tuberculosis disease control strategies is by genotyping the available *Mycobacterium tuberculosis* complex (MTBC)species and strains to identify their prevalence and endemic zones. This information on identification may be necessary in manipulating their genetic makeup and ultimate eradication. The development of multidrug resistance tuberculosis following multiple strains infection among patients with *M. tuberculosis* further compounds

the problem, creating the need for its total eradication. This eradication of multidrug resistance tuberculosis can only be achieved by knowledge of the molecular and allelic diversity of both the species and strains circulating in the Western Kenya.

1.6 Significance of study

Genetic diversity studies are important in the control of tuberculosis infections, diagnosis, drug development, multi-drug resistance prevention and vaccine development. The findings of this study will inform prevention, treatment and care of patients with tuberculosis infections. It will further inform policy makers on what species and strains that needs to be targeted in the control of tuberculosis infection. It will further add to the existing body of knowledge on genetic diversity of *M. tuberculosis*.

Identification of the genetic diversity will enable the development of treatment methods that bypass strains that can cause resistance to specific anti-tuberculotics. This will ultimately reduce on the long-term development of either multi-drug (MDR) or extremely multi-drug resistance tuberculosis (XMDR). By preventing the occurrence of MDR and X-MDR, there will be a reduction in wastage as a result of drug losses caused by patients developing drug resistance. This will eventually reduce on tuberculosis management costs. The control of the occurrence of MDR and XMDR tuberculosis and proper treatment of those with the disease will reduce the occurrence of new infections in the population and further reduce the prevalence of tuberculosis.

The findings of this study will further inform academicians, researchers and the public at large on the circulation Mycobacterium tuberculosis complex species, strains and molecular diversity. This will improve control strategies, create new knowledge for targeted drug

development and form a basis for future research on the molecular epidemiology of tuberculosis.

1.7 Limitations of the study

This study adopted 12-loci MIRU-VNTR instead of 24-loci MIRU-VNTR. The 24-loci MIRU-VNTR has a higher resolution and discriminatory power than the 12-loci MIRU-VNTR.

CHAPTER TWO

LITERATURE REVIEW

2.1 Epidemiology of Tuberculosis

Tuberculosis (TB) is a contagious and airborne disease caused by the *Mycobacterium tuberculosis* (Lönnroth *etal*., 2015) . It is the ninth global cause of mortality and the leading cause of morbidity from a single infectious agent, with the disease burden higher in low and medium income countries especially in Africa, such as Kenya (WHO, 2017). Tuberculosis is the leading cause of mortality among adults in the most economically productive age groups globally (Glaziou, Raviglione, Falzon, & Floyd, 2015).In 2016, the World Health Organization (WHO) estimated that 1.3 million TB deaths occurred among HIV negative people representing a significant decrease when compared with 1.7 million deaths in 2000 with an additional 374,000 deaths among HIV-positive people (WHO, 2017). The WHO report further estimates 10.4 million cases of tuberculosis, 90% of whom were adults. More than half (65%) of all the case notifications were males and 10% had a HIV infection. Most of the incidences of tuberculosis occurred in the South-East Asia Region (45%), followed by African Region at 25%. The lowest incidence rates have been shown in high-income countries such as those in Western Europe, North America, Australia and New Zealand. Averagely, these high-income countries have incidence rates of less than 10 per 100,000 persons per year. For TB to be eliminated, an annual decline of 4-5% in TB incidence is required. The decline since 2010 has exceeded 4% per year in several high TB burden countries, including Ethiopia, Kenya, Lesotho, Namibia, the Russian Federation, the United Republic of Tanzania, Zambia and Zimbabwe.

Kenya is among the 30 high burden tuberculosis countries in the world. Tuberculosis is the fourth leading cause of mortality in Kenya with the most affected age group being those aged

between 15-34 years (WHO, 2017). Kenya's neighbors in the horn of Africa have tuberculosis prevalence as follows: 65 cases per 100, 000 persons per year in Eritrea; 192 cases per 100, 000 persons in Ethiopia; 274 per 100, 000 persons in Somalia and 378 per 100, 000 persons in Djibouti (WHO, 2017). The estimated Kenyan national tuberculosis prevalence in 2015 was 233 per 100,000 individuals with an incidence of 81,518 and about 80% of the cases detected (WHO, 2017). The findings of the Kenya Tuberculosis Prevalence Survey of 2016, however, revealed that the burden of TB is higher than previous estimations by the WHO. This represented an underestimation of up to 40% (NTLD-P, 2017). Data from Kenya's Ministry of Health survey showed that the prevalence stands at 558 people out of 100, 000 persons, up from the WHO figures of 233 in every 100,000. The incidence of TB is approximately 138,000 cases per year according to the national prevalence survey (NTLD-P, 2017). The significance of this data is that about 40% cases of the disease go undetected thus reversing the efforts and investments towards the control of the disease.

The Kenya National Tuberculosis, Leprosy and Lung Disease Program initiated the TB REACH project with the objective of increasing case detection, enhance early detection and ensure timely and complete treatment (MOH, 2015). The immediate short-term goal is to achieve the 70/85 targets – that is to detect 70% of infectious TB and cure 85% of the detected cases and then sustain this effort over a long time. NTLD-P aims to implement integrated, patient-centred care and prevention; bold policies and supportive systems to intensify research and innovation (Ministry of Health, 2015). Majority of the Western Kenya Counties (Siaya, Busia, Kisumu, Kakamega and Uasin-Gishu) are among the ten high burden counties (Nyamogoba & Mbuthia, 2018) targeted by the TB REACH programme with a prevalence of 322 per 100, 000 individuals (NTLD-P, 2017).The top six TB burden counties in Kenya in terms of prevalence are: Mombasa

(535 per 100,000); Nairobi (490 per 100, 000); Homabay (426 per 100, 000); Kisumu (379 per 100, 000); Isiolo (338 per 100, 000) and Busia (329 per 100, 000). According to the Center for Disease Control (CDC), the at-high risk population includes persons infected with human immunodeficiency virus (HIV), people in close contact with TB patients sharing the same house or other enclosed environment (CDC, 2012). Persons born in foreign land from countries with high prevalence of tuberculosis or medically underserved low-income community (Raviglione & Sulis, 2016); as well as alcoholics and intravenous drug injectors may also be at a risk of tuberculosis infection.

Effective tuberculosis controls require early detection of cases and complete treatment of diagnosed cases until they are cured. Tuberculosis case notification and treatment success rates are used to measure the tuberculosis control program performance. Monitoring the treatment outcome and understanding of reasons for unsuccessful treatment outcome is important in order to improve treatment outcomes. The world health organization (WHO) estimates that one third of the world population is MTB infected (WHO, 2012).

The Global Fund report of 2013; indicated that 1.1 billion smear positive TB cases were reported in 2012. Of these, 83% were from twenty-two (22) high burden countries in both Asia and the Sub-Saharan Africa region. These countries have experienced a rapid increase in the recent years particularly in Nigeria, Angola, Cameroon, DRC Congo and Mozambique (WHO, 2013). These five countries accounted for one third of the new smear positive TB cases in the Sub-Saharan region in the year 2012.The report also indicated that MDR-TB is still a threat with 69,000 cases treated in 2012 and an estimated of 250,000 MDR-TB cases were left untreated in the year.

The WHO indicates that the countries with the highest numbers of new tuberculosis cases are India at 2.3million, China at 1 million, South Africa at 490,000, Indonesia 450,000 and Pakistan reporting 400,000 new cases. These five countries alone accounted for half of the 8.8 million reported new tuberculosis patients worldwide for the year 2010. The report further indicated that 60% of the total Global disease burden resided in the continent of Asia, 25% in Kenya, it has been found that new TB cases in Kenya dropped from 116,000 in 2007 to 99,152 in 2012. According to the Division of Leprosy, Tuberculosis and Lung Diseases (DLTLD) of Kenya, HIV pandemic, poor socio-economic leading to overcrowding in slums coupled with poor nutrition and limited access to health services have been identified as contributing factors to the high disease burden (Sitienei, Nyambati, & Borus, 2013). The incidence of TB varies with age. In Africa, tuberculosis primarily affects adolescents and young adults more than the older population due to crowding, risky behaviour, poor work environment and housing conditions(Kumar *etal*., 2007). However, in countries where TB has gone from high to low incidence, such as the United States, TB is mainly a disease of older people, and the immuno-compromised patients (Louw *etal*., 2006).

The importance of TB control in social and economic development has been widely acknowledged, including in the Millennium Development Goals. In this context, the World Health Organization (WHO) STOP TB Partnership has set two targets: to reduce prevalence and deaths by 50% by 2015, relative to 1990 levels; and to eliminate TB as a public health problem by 2050. Because tuberculosis is spread through contact, there is need for investigation involving the systematic evaluation of the contacts of known TB patients to identify active disease or latent TB infection (LTBI). It is one of a few active case-finding strategies that have been proposed to

increase case detection. The identification of the types and spread of *MTB* complex genetic strains is important in development of targeted disease control strategies.

2.1.1 Tuberculosis Diagnostic Techniques

2.1.1.1 MTB Diagnosis by AFB Smear Microscopy

Despite much work to develop new diagnostic tools for MTB diagnosis, AFB smear microscopy is still the primary means of TB diagnosis due to its rapid nature. Most clinical staff still base their diagnosis on Acid Fast Bacilli (AFB) Ziel Nelsen (ZN) staining. One of the major challenge to this technique is lack of sensitivity and reproducibility, as it largely depends on the experience and accuracy of the microscopist. Even with concentrated samples, the sensitivity of microscopy is not great, as it is in the order of 105 AFBs per ml of sputum.

The second challenge is differentiating tuberculous mycobacterium from non-tuberculous mycobacteria (NTM). This is due to the fact that once the presence of mycobacteria has been established, additional biochemical testing is required to identify the species. This requires experienced personnel to accurately identify the isolates.

The third challenge is the high prevalence of smear-negative TB, and this diagnosis may take weeks resulting in loss to care for many patients while they await culture results.

2.1.1.2 MTB Diagnosis by Liquid Culture

Lowenstein-Jensen (LJ) culture is often required as a confirmation for tuberculosis as it is both specific and sensitive. The subsequent isolate is required for identification of the organism to the species level, to determine drug susceptibility and to obtain the molecular profile for epidemiologic purposes. The major obstacle however, is the substantial time delays that require up to 8 weeks for cultures to become positive.

.

Automated liquid culture systems such as BacT/ALERT MP (bioMerieux Inc, Durham, NC USA) and BD BACTEC MGIT (Becton Dickinson) are currently considered gold standard approaches for isolating mycobacteria. Meta-analyses and systematic reviews have gone further to show that liquid systems are more sensitive for detection of the mycobacteria and can increase the case yield by 10% compared to solid media. They also reduce turnaround time from weeks to days. However, these systems are quite costly in our resource limited set-up.

Despite these benefits, liquid culture systems are prone to contamination and thus require stringent quality assurance systems and training standards.

2.1.1.3 MTB Diagnosis by Molecular tests

Nucleic acid amplification tests (NAATs) have been in use for many years especially in the developing countries. NAATs according to several met-analyses and systematic reviews, have been demonstrated to have a high specificity and Positive predictive value (PPV) but modest and variable sensitivity, especially in smear negative and extra-pulmonary tuberculosis (EPTB),. These high PPV allows for the discrimination between MTB and NTM while the compromised sensitivity may be attributed to the presence of PCR inhibitors in clinical specimens and loss of nucleic acids during processing of clinical specimens.

The introductions of NAATs have improved diagnosis by giving results within the same day. The role of polymerase chain reaction (PCR) in MTB identification has been established as a useful tool in the detection of both pulmonary and extra-pulmonary TB. Among the available NAATs (*COBAS TaqMan MTB –Roche, Gene Probe, ProbeTec ET DTB – Becton Dickinson, Xpert MTB/RIF*), PCR is the most widely used and best studied and published technique. It has

been established as a supplementary test since it has good rates of PPV and a better turnaround time than both culture and smear examination. Alternatives to PCR include real-time PCR (RT-PCR) and ligase chain reaction (LCx) all of which have been developed by Abbott.

The poor result quality of conventional microbiological techniques in TB diagnosis have stimulated the use of PCR in the identification of TB(Hoza, Mfinanga, Rodloff, et al., 2016).The development of PCR and other NAATs has led to the introduction of Point of Care (POC) technologies and Line Probe Assays (LPAs)(Ritter et al., 2014). The ideal TB diagnostic test is supposed to be simple, low-technology point of care test that could be rapidly performed and yield accurate results(Drobniewski et al., 2015; Foster et al., 2015). No existing test meets all these specifications. Early reports on Xpert MTB/RIF technology suggest that it was designed to be placed at point of contact sites where nurses could provide treatment immediately upon diagnosis(Detjen et al., 2015; Dorman et al., 2018).

The challenge for South Africa for having Xpert MTB/RIF technology at POC was due to the cost and the required improvement to the peripheral health facilities, such as air conditioning and stable electricity supply. The current Xpert MTB/RIF catridge has a 2 hour processing time which translates to a small capacity of 3-256 tests per day with an error rate of about 3.4%.

2.1.2 Tuberculosis drug resistance

Multidrug-resistant tuberculosis (MDR-TB) is caused by bacteria that do not respond to, at least, isoniazid and rifampicin, the two most powerful anti-TB medicines(Balian et al., 2017). It is diagnosed by a molecular based method which is performed in a Gene Xpert analyzer leading to detection of Rifampicin resistant tuberculosis strains(Detjen et al., 2015; Hoza, Mfinanga, & König, 2015).

The latest global burden of rifampicin- or multidrug-resistant tuberculosis (MDR/RR-TB) is at 4.1% for new and 19% for previously treated tuberculosis (Rosales-Klintz et al., 2012). Multi Drug Resistant tuberculosis can develop through primary or secondary pathways(Rosales-Klintz et al., 2012). Primary pathway involves getting exposure and infection with an already resistant TB strain while secondary is a result of initially being infected with the non-resistant strain then resistant strains emerge due to inadequate, incomplete or poor treatment quality(Borrell, Gagneux, & Gagneux, 2009).

Similar to drug-susceptible tuberculosis, drug-resistant tuberculosis only progresses to active disease in a minority of those infected and can remain latent for long periods of time(Alyamani et al., 2019). A poorly functioning immune system increases the risk of progression, and therefore factors that can impair the immune system like HIV, treatment with immunosuppressant, systemic diseases under nutrition, diabetes, alcohol abuse, silicosis and smoking are also risk factors for developing drug-resistant TB disease(Kariuki & Dougan, 2014).

2.2 *M. tuberculosis complex* species in clinical isolates

Mycobacterium tuberculosis (MTB) is a mycolic acid-coated bacterium discovered in 1882 as a causative agent of Tuberculosis disease (Merker, Kohl, Niemann, & Supply, 2017). The disease presents as either a pulmonary (lung) or extra-pulmonary (outside the lung) disease and is a major public health problem globally (Glaziou et al., 2018). There are many species of the Mycobacterium tubercle bacilli that have been associated with tuberculosis infection in humans (Teramoto etal., 2015; Wayne, 1982). The main mammalian tubercle bacilli infections result from, *Mycobacterium tuberculosis* complex, which consists of closely related species. These

species are *M. tuberculosis, M. bovis, M. africanum, M. microti, M. canetti* and *M. pinnipedii* (Caulfield & Wengenack, 2016).

Although *M. tuberculosis* mainly causes tuberculosis infection in humans, it may also affect animals that have contact with infected humans(Mbugi *etal*., 2016). It can present as pulmonary type of tuberculosis or the extra-pulmonary type. *M. bovis* which causes bovine tuberculosis has been found to be a zoonotic infection in humans(Brosch *etal*., 2002; Glaziou *etal*., 2018). *M. Africanum* is an intermediate species between *M. tuberculosis* and *M. bovis,* being closer to the latter(Wayne, 1982; Wirth *etal*., 2008). The organism causes tuberculosis in humans and occasionally animals such as apes. *M. microti* which is attenuated in humans is a member of the complex and termed as the vole bacillus as it causes tuberculosis in voles *(Microtus agrestis)* and other small rodents(Brosch *etal*., 2002; Burthe *etal*., 2008). Because of the proximity of humans to rodents and birds, there is a great likelihood of *M. microti* being a zoonotic infection(Burthe *etal*., 2008).*M. canetti* is rarely seen but can cause tuberculosis in humans (Glaziou *etal*., 2018; Sola *etal*., 2003).

Mycobacterium tuberculosis species identification can be used as an additional tool in epidemiological investigations in order to gain a better understanding of factors that influence tuberculosis transmission, identification of risk factors of tuberculosis transmission in a community and for evaluation of regional control programmes permitting a rational design of more adequate control measures (DePristo *etal*., 2011; Pankhurst *etal*., 2016).

2.3 Proportions of *Mycobacterium* strains isolated from clinical isolates

The world health organization reported that of the 1.1billion smear positive tuberculosis cases were reported in 2016, 83% were from twenty two (22) high burden countries in both Asia and

the Sub-Saharan Africa region (WHO, 2017). These countries - Nigeria, Angola, Cameroon, DRC Congo and Mozambique - have experienced a rapid increase in the recent years (Glaziou *etal*., 2018). The proportion of *M. tuberculosis complex* strains causing tuberculosis disease varies from region to region, though human migratory behavior disseminates these strains to different geographical locations globally (Ködmön, Zucs, & Van Der Werf, 2016).

The various *M. tuberculosis* complex strains include: Beijing, Haarlem, Uganda I and II, Bovis, Latin American Mediterranean (LAM), West African I and II, East African Indian (EAI), Central Asian Strain (CAS), Delhi, URAL, Cameroon and TUR. The Beijing strain has been found to be very predominant in Asian Countries such as China, Vietnam, Thailand and Indonesia.

The proportion of tuberculosis due to the Beijing strain has been found to be >50% in many Asian studies(Glynn*et al,* 2002). However, other studies have reported the Beijing strain in countries outside Asia, such as those in Middle East (Iran), Africa (Senegal, Tanzania, Mozambique, Kenya), Europe (Russia, Azerbaijan, Estonia, Netherlands, Spain), North America (USA, Canada), South America (French Guiana) and the Caribbean (Cuba) (Azé *etal*., 2015; Brosch *etal*., 2002; Glynn *etal*., 2002).

In a study carried out in Addis Ababa, Ethiopia, a total of 53 different patterns were identified among 192 isolates examined. 169 of the isolates were classified into one of the 33 shared SITs, whereas the remaining 23 corresponded to 20 orphan patterns. Other prominent strains identified were: CAS, Haarlem, LAM, Beijing, and Unknown comprising 26%, 13%, 2.6%, 0.5%, and 2.1%, respectively. Among HIV-positive patients, 10 patterns were observed among 25 isolates. The T (38.5%), H (26.9%), and CAS (23.1%) families were the most common among HIV-positive individuals. The diversity of the *M. tuberculosis* strains found in this study is very high, and there was no difference in the distribution of families in HIV-positive and HIV-negative TB

patients except the H family. Tuberculosis transmission in Addis Ababa is due to only the modern *M. tuberculosis* strains, namely: CAS, LAM, T, Beijing, Haarlem, and U(Mihret *etal*., 2012).Distribution of the predominant clusters of *M. tuberculosis* strains have been shown to vary among different populations with the majority of the isolates (83.2%) belonging to three major families: T (54.2%), CAS (26%) and H (13.2%) (Mihret *etal*., 2012).

There are forty seven different strain families of *M. tuberculosis* that have been identified in Nairobi, Kenya (Ogaro *etal*., 2012). The principal groups were: CAS1_KILI (17%), T1 (12%), Beijing (12%), LAM (9%), LAM-3 & S/Conversant(7%), LAM-11_ZWE (5%), CAS1_DELHI (4%), T2 (4%) and a possible new *M. tuberculosis* strain family designated as the Nairobi sub-type (4%) (Ogaro *etal*., 2012).

2.4 Molecular characteristics of *Mycobacterium tuberculosis*

The *M. tuberculosis* genome size is made up of 4 million base pairs, with 3959 genes; 40% of which have had their function characterized, with possible function postulated for another 44%(Brosch *etal*., 2002; Merker *etal*., 2017; Wirth *etal*., 2008). Within the genome there are also six pseudogenes(Azé *etal*., 2015). The genome contains 250 genes involved in fatty acid metabolism, with 39 of these involved in the polyketide metabolism generating the waxy coat(Allix-Béguec*et al*,2008). Nine have been characterized in *M. tuberculosis* with a further 56 predicted in a bioinformatics screen(Pankhurst *etal*., 2016).

Studies have demonstrated that the genome of *M. tuberculosis* is highly conserved(Arnvig& Young, 2009). Comparative sequence analysis of the 275 bp internal transcribed spacer (ITS) region that is highly polymorphic and separates the 16SrRNA and the 23SrRNA revealed

complete conservation between members of the *M. tuberculosis complex*. Sequence analysis of 56 structural genes in several hundreds of phylogenetically and geographically diverse *M. tuberculosis complex* isolates indicated that allelic polymorphism is extremely rare(Shamputa *etal*., 2010). While the *M. tuberculosis* genome is highly restricted in relation to other pathogenic bacteria, it does have regions which are characteristically punctuated by monomeric sequences repeated periodically(Ei*et al*, 2016). This could be due to evolutionary changes and adaptation over time.

There are two types of repetitive units commonly referred to as interspersed repeats, that is direct repeats (insertion repeats) and tandem repeats (TR) with the head to tail direct uninterrupted(Allix-Béguec *etal*., 2008). Prokaryotic microsatellites (1-10) and mini satellites (10-100 bp repeats) commonly referred to as tandem repeats are located in the intergenic regions, regulatory regions or within the open reading frame and are abundant throughout the bacterial genome. Tandem repeats commonly observed in *M. tuberculosis* include insertion sequences (IS), Polymorphic GC-rich repetitive sequences (PGRS), Spacer oligonucleotide sequences (Spoligotype), Variable number tandem repeats (VNTR), Mycobacterium interspersed repetitive units (MIRU) and Single nucleotide polymorphism (SNP)(Azé *etal*., 2015). Insertion sequences (IS) are small mobile genetic elements usually less than 2.5 kb in size, widely distributed in bacteria genome and function to carry only the genetic information relating to their transposition and regulation.

Thierry *et al* (1990) first described IS6110, as a 1,355-bp member of IS_3 family that when intact is unique for *M. tuberculosis complex*(Supply *etal*., 2006). IS6110 has an imperfect 28-bp inverted repeats at its ends and generates 3-4 bp target duplications or insertion and are randomly distributed throughout the genome of tuberculosis, however the hot spots (regions IS prefers to

insert) are noted. The IS6110genotyping was used to demonstrate endogenous reactivation of tuberculosis after over 30 years of latency(Azé *etal*., 2015).

M. tuberculosis complex strains contain distinct chromosomal regions consisting of multiple 36-bp direct repeats (DR) interspersed by direct unique DNA sequence 35-41 base pa (bp)(Brosch *etal*., 2002; Rosales-Klintz *etal*., 2012).Spacer oligonucleotide typing (Spoligotype) is based on the detection of 43 interspersed spacer sequences in the genomic DR regions of *M. tuberculosis complex* strains(Bazira, Asiimwe, & Joloba, 2010).

Mycobacterium interspersed repetitive units (MIRUs) are 40–100 bp DNA elements often found as tandem repeats and dispersed in intergenic regions of the *M. tuberculosis complex* genomes(Savine *etal*., 2002). The *M. tuberculosis* H37Rv chromosome contains 41 MIRU loci. After polymerase chain reaction (PCR) and sequence analyses of these loci in 31 *M. tuberculosis complex* strains, 12 of them were found to display variations in tandem repeat copy numbers and, in most cases, sequence variations between repeat units as well(Supply *etal*., 2006). These features are reminiscent of those of certain human variable mini-satellites. Tandemly repeated sequences are dispersed by thousands of copies in virtually all higher eukaryote genomes (Merker *etal*., 2015). Many of these loci show hyper variability in their repeat numbers in humans and in animals, and are therefore also called variable number tandem repeat (VNTR) loci(Perdigão *etal*., 2016). The VNTRs are also invaluable tools for studying various aspects of evolution, ranging from pedigree to evolutionary distant phylogenetic relationships(Alyamani *etal*., 2019).

Genetic diversity refers to both the vast numbers of different species as well as the diversity within a species(Getahun*et al*, 2016). A diversity index is a quantitative measure that reflects how many different types (such as species, strains) there are in a dataset, and simultaneously

takes into account how evenly the basic entities (such as individuals) are distributed among those types(Ahmed *etal*., 2014; Sola *etal*., 2003). The value of a diversity index increases both when the number of types increases and when evenness increases. For a given number of types, the value of a diversity index is maximized when all types are equally abundant.

Genetic polymorphisms among bacterial pathogens are exploited by epidemiologists to monitor disease outbreaks and transmission of specific strains. *M. tuberculosis* was widely regarded as a homogeneous population with limited genetic variability(Pankhurst *etal*., 2016). However, data from recent studies suggest that the functional diversity of this pathogen has been underestimated (Hershberg *etal*., 2008).

Tuberculosis is an ancient disease, with evidence that the causative agent *M. tuberculosis* has been co-evolving with its human host since mankind emerged from Africa 50,000 years ago (Merker *etal*., 2017). Members of the *M. tuberculosis complex* are considered genetically monomorphic with a high level of genomic sequence similarity (>99.95%), limited horizontal gene transfer, and a clonal population structure(Carey *etal*., 2018). This apparent homogeneity led to the assumption that genetic diversity among *M. tuberculosis complex* strains would not be of clinical significance.

In different parts of the world many studies conducted demonstrated the presence of wide genetic diversity among *M. tuberculosis complex* (Githui *etal*., 2004). Using microarrays, genetic diversity of *M. tuberculosis* strains has been found to find significant impact on global gene expression(Hoza*et al*, 2016). There is mounting evidence that the genetic variability among clinical isolates may have dramatic consequences on the outcome of infections(Wirth *etal*., 2008). It is also believed that the genetic diversity of the *M. tuberculosis complex* contributes to

the wide spectrum of tuberculosis clinical presentations, including acute primary tuberculosis (localized or disseminated), latent disease and reactivation (Bazira *etal*., 2010).

According to surveyed sequence diversity within a global collection of strains belonging to *M. tuberculosis complex* using seven megabase pairs of DNA sequence data; it was shown that the members of *M. tuberculosis complex* affecting humans are more genetically diverse than generally assumed. These findings suggest that the current increases in human population, urbanization, and global travel, combined with the population genetic characteristics of *M. tuberculosis*, could contribute to the emergence and spread of drug-resistant tuberculosis(Fonseca, Knight, & McHugh, 2015). Many previous studies have demonstrated and revealed wide genetic diversity using different molecular techniques(Getahun *etal*., 2016).

In Khuzestan province (Iran), a study on genetic diversity involving 80 *M. tuberculosis* isolates identified 48 distinct patterns comprising 11 clusters and 37 unique patterns(Khosravi *etal*., 2017). The discriminatory power of MIRU-VNTR was high (HGDI=0.991) for these samples. Distribution of the predominant clades (clusters) of *M. tuberculosis* strains showed variation among different populations and the majority of the isolates belonged to three major clusters: T cluster, CAS and H cluster (A Mihret *e tal*., 2012).

In a study carried out in Mbarara, Uganda to determine the genetic diversity of *M. tuberculosis* complex isolated from 125 patients using regions of difference (RD) and spoligotyping analysis; it was shown that spoligotyping revealed wide diversity of 79 patterns with an overall diversity of 63.2% and sixty isolates formed 16 clusters of 2-15 isolates(Kiwanuka *etal*., 2013). Two studies conducted in Kenya demonstrated a high diversity of *M. tuberculosis* in Nairobi city based on spoligotypes, however there are no recent molecular studies conducted in western Kenya to demonstrate the genetic diversity of tubercle bacilli circulating in the region(Githui

etal., 2004; Ogaro *etal*., 2012). Over recent years, there has been an increasing acknowledgment of the diversity that exists among *Mycobacterium tuberculosis* clinical isolates. To facilitate comparative studies aimed at deciphering the relevance of this diversity to human disease, an unambiguous and easily interpretable method of strain classification and assigning isolates into a series of unambiguous lineages is achieved by the method of Gagneux (2006). It was proposed that *M. tuberculosis* has a clonal genetic population structure that is geographically constrained(Gagneux *etal*., 2006)and major *M. tuberculosis* lineages are associated with the country of origin. The introduction of molecular methods for typing and identification of strains/lineages of *M. tuberculosis* has gained increased acceptance as a powerful tool for epidemiological and phylogeny studies of tuberculosis. These molecular methods include: Mycobacterium interspersed repeat unit-Variable number tandem repeat (MIRU-VNTR), Spacer Oligonucleotide typing (Spoligotype), Insertion Sequence 6110(IS6110); Restriction fragment Length Polymorphism (RFLP).Regions of Deletion(RD), Pulsed field gel electrophoresis(PGE), and Polymorphic GC-rich repetitive sequence PGCRS). Methods based on mini-satellites that contain variable numbers of tandem repeats (VNTRs) have been demonstrated to be effective and portable methods for typing *M. tuberculosis* due to the numeric nature of the data generated.

CHAPTER THREE

MATERIALS AND METHODS

3.1. Study area

The study was carried out at the Moi Teaching and Referral Hospital (MTRH), Eldoret town (0.5143° N, 35.2698° E) in Western Kenya. MTRH is one of the three national teaching and referral hospitals in Kenya. It houses a number of universities, centers of excellence for both infectious and non-communicable diseases. MTRH serves several counties in western Kenya region, such as Uasin-Gishu, Nandi, Kakamega, Bungoma, Trans Nzoia, Elgeyo-Marakwet, West Pokotand Turkana(Appendix 1). Western Kenya region has a high tuberculosis disease burden(NTLD-P, 2017)making MTRH a suitable site for studying *M. tuberculosis*-related disease and its mechanisms.

3.2 Study design

3.2.1. Research design

This was a retrospective study in which archived samples between 2013 and 2014 were used. Only samples from patients with confirmed tuberculosis based on Lowenstein-Jensen (LJ) medium cultures were used.

3.2.2. Sample size

The desired sample size was determined using Cochran formula(Cochran, 1977). The 2013 average tuberculosis prevalence in Kenya was 23.3%(NTLD-P, 2017; WHO, 2015).

This was used as a reference for a single proportion with a margin of error of 10% at 90% power.

$$n_0 = \frac{t^2 \times p(1-p)}{d^2}$$

Where:

t =alpha level (1.65)

p = proportion with tuberculosis (0.233)

d = acceptable margin of error (0.1)

$$n_0 = \frac{(1.65)^2(0.233)(0.767)}{(0.1)^2} = 48.65$$

Since the calculated sample size exceeded 10% (34.6) of the archived samples which was 346 during the study period, Cochran's correction formula was used to obtain the final samples size as shown below.

$$n_i = \frac{n_0}{1 + \frac{n_0}{346}} = 42.3$$

Therefore, 42 archived samples were selected in the present study. However, only 40 samples were adequate for PCR analysis. This reduced the desired sample size by two. However, even with the reduction, the power was still within the acceptable limits as determined by the resource equation method previously described(Charan & Biswas, 2013).

3.2.3.Inclusion criteria

i. Archived samples positive for tuberculosis.
ii. Archived samples with identifiable socio-demographic records.

3.2.4. Exclusion criteria

i. Archived samples with mixed growth (such as molds) and contamination.
ii. Archived samples not positive for tuberculosis from the Lowenstein-Jensen cultures.

3.2.5. Ethical consideration

Use of sputum samples from human participants and all experimental protocols were reviewed and approved by Institutional of Research and Ethics Committee (IREC) of Moi Teaching and Referral Hospital (MTRH) and Moi University (Approval number: IREC: 0001214) (AppendixII). Permission to use archived samples was also granted by the MTRH management. An informed consent waiver was also obtained since this study only used archived samples. Confidentiality was maintained throughout the study by de-identifying clinical isolates and storing the results in a password protected database that was only accessible to the principal investigator and the research assistants.

3.2.6. Sampling technique

Simple random sampling was used as previously described (Degu & Tessema, 2005). Briefly, at the period of study there were 346 samples at MTRH. The samples were numbered from 1-346. The desired sample size (n=42)were selected using alottery method in which slip of papers was used to represent each of the samples. The papers were put in a box and mixed, and a sample of the required size was drawn from the box.

3.3. Sample processing

Tuberculosis-positive samples archived between 2013 and 2014 at -80°C were retrieved after confirming patient identification numbers from the laboratory records. The samples were placed on in a laminar flow at 25°C to thaw.

3.3.1 Identification of species, strains, clusters and allelic diversity of *Mycobacteria* using MIRU-VNTR

Samples were assayed using PCR and analyzed using MIRU-VINTRplus platform to identify species and strains, perform clusters analysis and calculate allelic diversity in each of the 12

MIRU-VNTR loci of the *Mycobacteria* circulating in western Kenya between 2013 and 2014 as described in the subsequent procedures.

3.3.2 DNA extraction

Mycobacterial DNA was extracted as previously described(Shamputa *etal*., 2010). Briefly, 1 mL of mycobacterial sample was put into an Eppendorftube and centrifuged for 15 minutes at 10,000 rpm to pelletize the mycobacterial cells. The supernatant was discarded, and the pellet re-suspended in 300µL of molecular grade water. The suspension was incubated at 95°C for 20 minutes followed by ultrasonic bath incubation for 15 minutes. Then, the product was centrifuged for 5 minutes at 10,000 rpm. The supernatant was transferred into a new Eppendorftube for further analysis.

3.3.3 PCR amplification

The *M. tuberculosis* isolates were genotyped by performing PCR amplification on 12 loci MIRU-VNTR using an automated bio-sequencer method as described by Supply*et al*, 2006. Briefly, 2µLof DNA samples were transferred into a 96 well-plate that had 25µL mixture of 2 µL PCR buffer, 0.2 mMdNTPs, 1.5-3 mM $MgCl_2$, 0.4 µM of each fluorescent primer, 0.4 U of Taq DNA polymerase. PCR amplification of each sample was done on 12 different MIRU-VNTR regions using six quadruplex PCRs with fluorescent primers (Table 3-1) specific for each flanking region (Supply *et al*., 2001).The cycling program was carried out in a thermal cycler (Applied Biosystems), with the following conditions: 15 min initial denaturation at 95°C followed by 40 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min depending on MIRU loci, and extension at 72°C for 90 seconds. The reaction was terminated by a final extension at 72°C for 10 min(appendix III).

3.3.4 Sequencing

The PCR fragments labeled with the three different fluorescent dyes from each multiplex PCR were combined with an internal size standard, labeled by a fourth dye, and then analyzed in single lanes on an automated DNA sequencer. The electrophoresis and analysis parameters were adapted in order to size amplicons from 200 to 1,000 bp, which covers the size ranges of all MIRU-VNTR alleles obtained for M. tuberculosis isolates studied(Kremer et al., 2005; Supply et al., 2001). The accuracy of the size determination was controlled by analyzing labeled PCR fragments amplified in multiplex PCR from M. tuberculosis H37Rv and other isolates, for which the exact sizes were determined by sequence analysis (Allix et al., 2006). Sizing was found to be reproducible, with within-run and between-run average precisions of ±0.5 and ±0.6 bp, respectively. Mean errors ± standard deviations for the sizing accuracy were 1.1 ± 1.0 bp for fragments below 500 bp and 0.8 ± 4.6 bp for fragments from 500 to up to 971 bp. This ensures the unambiguous identification of MIRU-VNTR alleles, which differ in size by 51 to 77 bp. Stutter peaks are common during the genotyping of shorttandem repeat sequences and mostly reflect artifactual strand slippage of the polymerase during PCR. Such stutter peaks were also quite frequently observed for the PCRs of the various MIRU-VNTR loci. However, they could be easily diagnosed because, as for short tandem repeat sequences, they appeared as a ladder of much-lower-intensity peaks, corresponding to sizes of PCR fragments that lack one or more repeats. In these cases, the stutter peaks from the multiplex reaction had heights comparable to those of the fragments corresponding to these alleles. When these samples were reanalyzed by PCR with only the primer pair corresponding to this locus, single high-intensity peaks were obtained, allowing us to correctly assign the alleles. Genotyping of a blinded set of M. tuberculosis complex isolates from 38 countries. A synthetic interlaboratory study recently.

Table3-1: Oligonucleotide primer sequences used for amplification(Allix et al., 2006)

Locus	Repeat unit length, bp	Set	Primer sequence (5' to 3') with labels
MIRU 4	77	FW	GCGCGAGAGCCCGAACTGC (**FAM**)
		RV	GCGCAGCAGAAACGCCAGC
MIRU 26	51	FW	TAGGTCTACCGTCGAAATCTGTGAC
		RV	CATAGGCGACCAGGCGAATAG (**VIC**)
MIRU 40	54	FW	GGGTTGCTGGATGACAACGTGT(**NED**)
		RV	GGGTGATCTCGGCGAAATCAGATA
MIRU 10	53	FW	GTTCTTGACCAACTGCAGTCGTCC
		RV	GCCACCTTGGTGATCAGCTACCT (**FAM**)
MIRU 16	53	FW	TCGGTGATCGGGTCCAGTCCAAGTA
		RV	CCCGTCGTGCAGCCCTGGTAC (**VIC**)
MIRU 31	53	FW	ACTGATTGGCTTCATACGGCTTTA
		RV	GTGCCGACGTGGTCTTGAT (**NED**)
MIRU 02	53	FW	TGGACTTGCAGCAATGGACCAACT
		RV	TACTCGGACGCCGGCTCAAAAT (**FAM**)
MIRU 23	53	FW	CTGTCGATGGCCGCAACAAAACG (**VIC**)
		RV	AGCTCAACGGGTTCGCCCTTTTGTC
MIRU 39	53	FW	CGCATCGACAAACTGGAGCCAAAC
		RV	CGGAAACGTCTACGCCCCACACAT (**NED**)
MIRU 20	77	FW	TCGGAGAGATGCCCTTCGAGTTAG
		RV	(**FAM**) GGAGACCGCGACCAGGTACTTGTA
MIRU 24	54	FW	CGACCAAGATGTGCAGGAATACAT
		RV	GGGCGAGTTGAGCTCACAGAA (**VIC**)
MIRU27	53	FW	TCGAAAGCCTCTGCGTGCCAGTAA
		RV	GCGATGTGAGCGTGCCACTCAA (**NED**)

Note: FW: Forward

RV: Reverse

Boldface letters depict various fluorescent dyes used

3.4 Socio demographic Data Collection

Socio demographic data of participants age and sex were obtained from the medical records available in the study setting.

3.5Data analysis

3.5.1 MIRU-VNTR typing

To search for the concordance between 12 MIRU-VNTR typing, the 12 MIRU-VNTR profiles obtained from the isolates tested were identified in the MIRU-VNTR *plus* database accessible at http://www.miru-vntrplus.org. A phylogenetic analysis using parsimony (PAUP) was further done.

An N-J dendrogram was constructed based on the MIRU-VNTR data, as appropriated, using the MIRU-VNTR *plus* international database. A cluster was defined as group of two or more strains with identical profile. A minimum spanning tree was also constructed using the MIRU-VNTR *plus* database to investigate phylogenetic relationships within the sample and identify clonal complexes. A clonal complex was defined as groups of isolates that are within dual-*locus* variants of each other.

3.5.2. Evaluation of allelic diversity of MIRU-VNTR loci using Hunter-Gaston discriminatory index (HGDI) calculation

The calculation of the discriminatory index was based on Simpson's index of diversity as described by Hunter and Gaston(Azé *et al*., 2015). This value is commonly referred to as the HGDI and was calculated by using the Web tool http://insilico.ehu.es/minitools/discriminatory power. HGDI was used to quantify the discriminatory index and to evaluate the allelic diversity of the different MIRU-VNTR loci. Based on their respective HGDI score, the discriminatory powers were classified as high (HGDI $\geq$0.6), moderate (0.3 < HGDI < 0.6) or low (HGDI $\leq$0.3) (Cowan *etal*., 2005).

CHAPTER FOUR

RESULTS

4.1 Socio-demographic characteristics of study participants

The study was done on 40archived samples from participants of both gender of which 57.5% and 41.7% were males and females respectively. The participants' age ranged between 18 and 65 years with median age being 30years.

Table 4-1: Participants' Socio-demographic Characteristics

Parameter	**% (n/IQR)**
Male	57.5 (23)
Female	42.5 (17)
Age	30 (IQR 25.0 – 38.3) years

Note: IQR: Interquartile range

4.2: Identification of *M. tuberculosis* species in clinical isolates

From the 40 isolates, three *Mycobacterium* species namely: *M. tuberculosis (60%; n=24),M. bovis (2.5%; n=1)*and *M. africanum (2.5%; n=1)*were identified, with *M. tuberculosis* being the most identified and predominant species. Further, there were unknown species (32.5%; n=13) of *Mycobacterium* identified as well as those that had mixed (multiple) matches(2.5%; n=1) of the *Mycobacterium* species (Figures4.1 and 4.3).

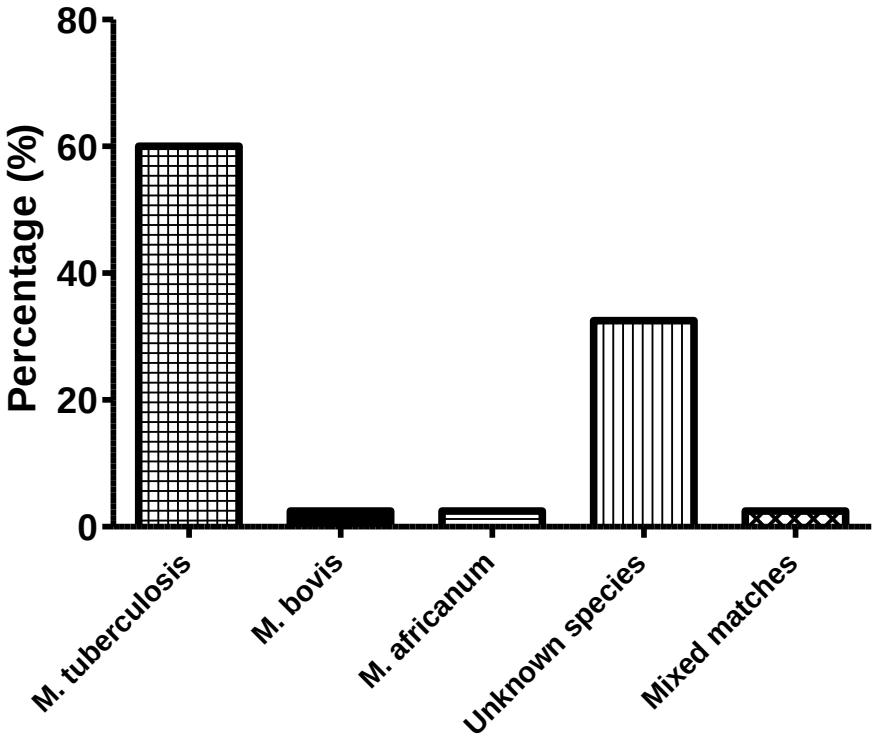


Figure 4.1: Relative abundance of Mycobacterium species circulating in Western Kenya between 2013 and 2014 (*M. tuberculosis, 60%; M. bovis*, 2.5%; *M. africanum*, 2.5%; Unknown species, 32.5%; Mixed matches, 2.5%).

4.3. Proportions of *M. tuberculosis complex* strains

The study revealed that there were a number of *Mycobacterium tuberculosis* complex strains circulating in western Kenya between 2013 and 2014. These strains included Beijing, Latin American Mediterranean (LAM), Uganda I, East African Indian (EAI) and the multiple *M. tuberculosis* strains. The multiple *M. tuberculosis* strains consisted of a mixture of combinations, such as URAL/Harleem/Cameroon, Camerron/ Harleem, Ghana/Delhi/CAS, Beijing/Uganda I and II. The proportion of Mycobacterium strains identified is shown in figure 4.2.

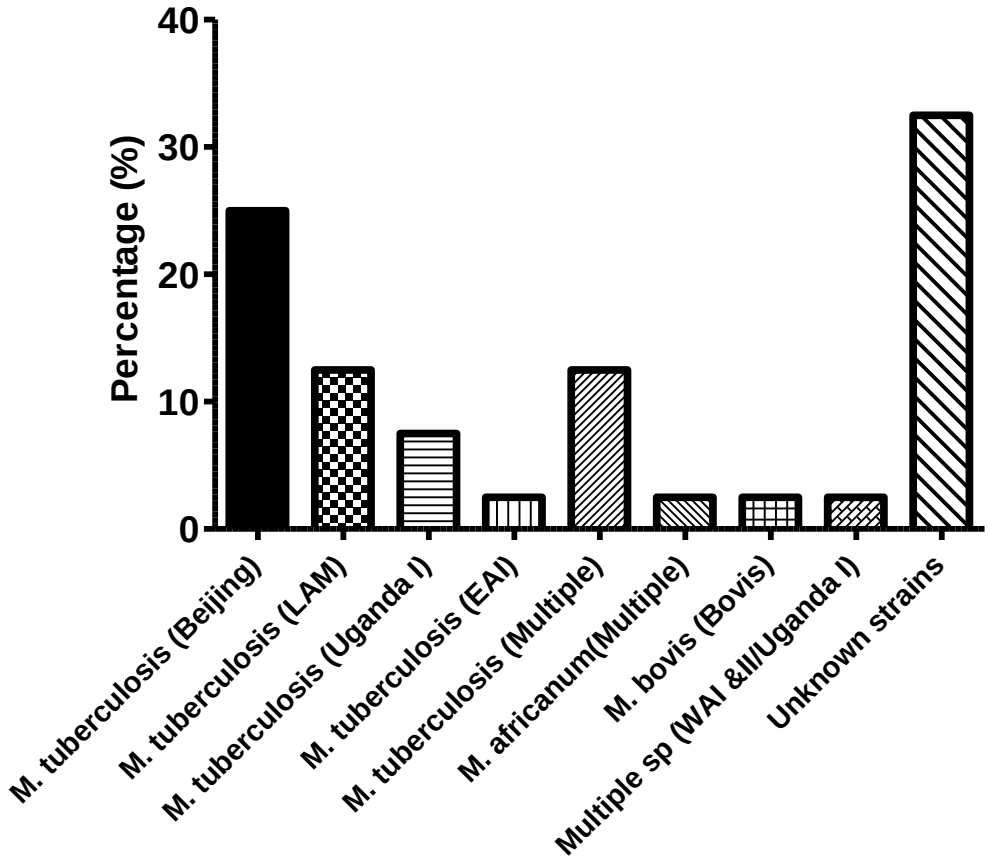


Figure 4.2: Relative abundance of Mycobacterium strains (Beijing = 25%; LAM = 12.5%, Uganda I = 7.5%; EAI = 2.5%; Multiple strains = 12.5%, M. Africanum =2.5%; M. bovis =2.5%, Multiple species = 2.5%, Unknown strains = 32.5%**).**

4.4. Genetic diversity of *Mycobacterium* isolates

4.4.1. Cluster analysis

Using a neighbor-joining algorithm, a dendrogram and a minimum spanning tree were generated for cluster analysis. A clustering rate of 56.4% was noted in the 12 MIRU loci. Using similarity search, five clusters, namely Beijing(n=10), URAL (n=1), EAI(n=1), UgandaI(n=3) and LAM(n=5) were identified from the 40 isolates. The five clusters had varying number of strains: MTB Beijing, LAM, Ugandan I, URAL and EAI with ten, five, three, one and one strains

respectively. Furthermore, a minimum spanning tree showed that the known *Mycobacterial* isolates circulating in western Kenya belong to six clonal complexes with most of the isolates falling in complex 1 (Figure 4.3).

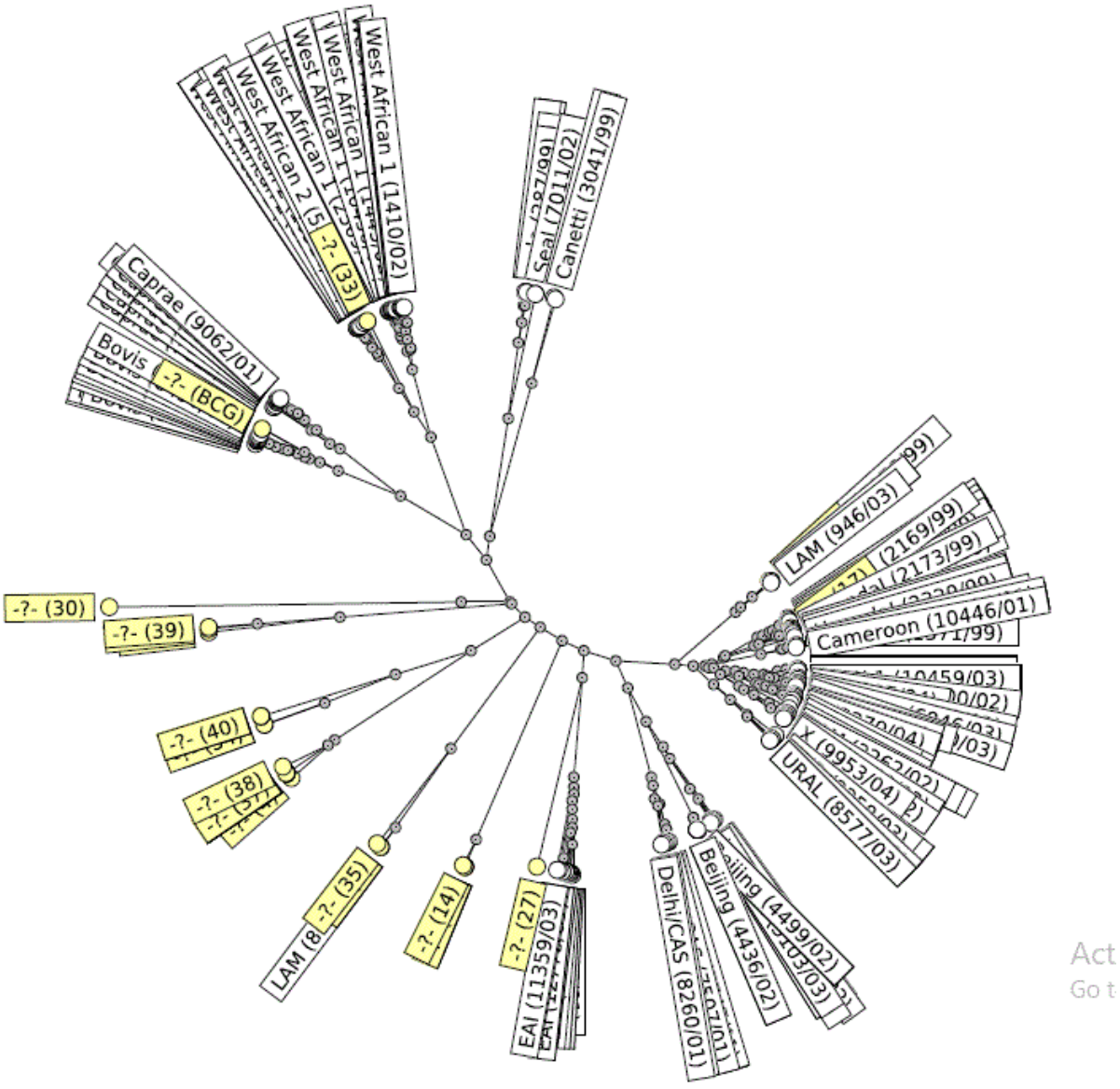


Figure 4.3:MIRU-VNTR dendrogram (12 *loci*) of the 40 *M. tuberculosis* clinical isolates circulating in western Kenya between 2013 and 2014.

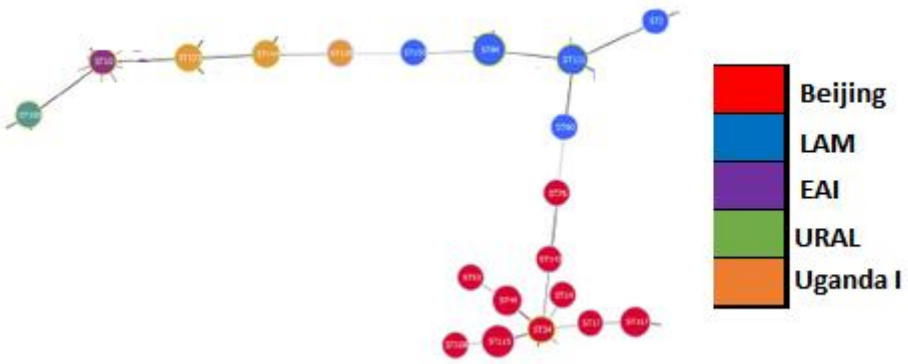


Figure 4.4:Minimum spanning tree of *M. tuberculosis* clinical isolates circulating in western Kenya between 2013 and 2014.

The minimum spanning tree (Figure 4.4) was constructed based on 12-*loci* MIRU-VNTR genotypic data and clonal complexes defined as MIRU-VNTR profiles within *locus* variants of each other.

4.4.2. Allelic diversity

The allelic diversity was calculated for each MIRU-VNTR locus. The allelic diversity index of *Mycobacterium* strains circulating in western Kenya ranged from 0.31 to 0.84 with a mean HGDI of 0.808 (Table 4-2).

Table 4-2: Allelic diversity of M. tuberculosis complex isolates circulating in western Kenya between 2013 and 2014.

Locus	HGDI
MIRU39	0.74
MIRU26	0.84
MIRU40	0.72
MIRU10	0.83
MIRU16	0.81
MIRU27	0.71
MIRU23	0.63
MIRU31	0.83
MIRU20	0.31
MIRU02	0.45
MIRU24	0.51
MIRU04	0.56

Note: High (HGDI $\geq$ 0.6), moderate (0.3 $\leq$ HGDI $\leq$ 0.6) or low (HGDI $\leq$ 0.3)

CHAPTER FIVE

DISCUSSION

5.1. *Mycobacterium tuberculosis* species from clinical isolates circulating in western Kenya

Tuberculosis remains a major source of morbidity and mortality globally(WHO, 2017). Thus, better ways of treating and controlling the spread of TB are imperative. This requires close monitoring of the species and strains, clusters and allelic diversity of strains circulating in a given region. It is against this backdrop that the current study was designed. The present study determined the specific *M. tuberculosis* complex species circulating in a high TB prevalent area, Western Kenya region in 2013 and 2014. The main species found circulating in the region were: *M. tuberculosis M. Bovis* and *M. africanum.* Further, there were unknown *Mycobacterium* species and multiple matches of *Mycobacterium* species drawn from *M. tuberculosis* and *M. Africanum.* The proportion of the *Mycobacterium* species ranged from 2.5-60% with *M. tuberculosis* being the predominant species. The finding of the current study is in agreement with several previous studies in Uganda (Wampande *etal*., 2013), Tanzania (Mbugi *etal*., 2016), Mozambique (Saifodine *etal*., 2014) in which *M. tuberculosis* species was reported to be the most common circulating species. However, the findings is not consistent with a previous study in which *M. Africanum* was reported to be the most prevalent *M. tuberculosis* complex (MTBC) species in Ghana(Otchere *etal*., 2016). The discrepancy between the present and previous report with reference to *M. africanum* prevalence could be attributed to limited migrations and spread of the *M. africanum*, which is predominantly located in West African countries (de Jong, Antonio, & Gagneux, 2010).*M.bovis*is the main causative agent of tuberculosis in cattle (bovines). However, *M. bovis*has a wide host range i.e.goats, cats, dogs, pigs, lions, non-human primates and humans. The susceptibility of humans to *M. bovis* is due to their putative spill over

host nature, however the infection is not confined to them(Allix *etal*., 2006).Majority of the families in Western Kenya rear cattle and keep dogs. This could provide a risk of zoonotic infection to the farmers as demonstrated in this study where clinical isolates were identified as *M. bovis*.

The observation of different *Mycobacterium* species in circulation and many unknown species (32.5%) is a worrying trend since western Kenya is gateway to most of the east African countries hence could enhance the spread and development of genetic variants of *Mycobacterium* which might render treatment difficult.

5.2. *Mycobacterium tuberculosis* strains from clinical isolates circulating in western Kenya

In the present study, several *Mycobacterium* strains were identified among the clinical isolates in circulation in western Kenya. These strains included Beijing, LAM, Uganda I, Ural, EAI and the multiple *M. tuberculosis* strains. Beijing and unknown were the predominant strains in this study. The present study is consistent with previous studies which have identified Beijing family as predominant strain in most countries(Chuang, Lee, Liu, Lee, & Ideker, 2007; Jagielski *etal*., 2016; Mbugi *etal*., 2015). It is well established that Beijing strain is most prevalent in East Asian region, accounting for over 50% of all the strains(Sekizuka *etal*., 2015; Shamputa *etal*., 2010). The high prevalence of Beijing in western Kenya and in many other continents could be related to the high human migratory activities from and to East Asia. This view is supported by a study in Tanzania in which the Beijing strain was high in Serengeti ecosystem, a popular tourist attraction region, where individuals from other regions congregate(Mbugi *etal*., 2015). The LAM and EAI strains are commonly found in African countries that share a border with Mozambique, such as Tanzania, Malawi, Zambia and Zimbabwe(Saifodine *etal*., 2014).The high economic activities between Kenya and many other countries in eastern and central Africa might be

enhancing the transmission of LAM and EAI strains. The same explanation could be responsible for the presence of Uganda I strain found in this study.

In agreement with a study in the East African region (Mbugi *etal*., 2015),the present study revealed a considerable prevalence of unknown *Mycobacterium* strains. Although the explanation for the high prevalence of unknown *Mycobacterium* strains in western Kenya remains unclear, a possible explanation could be related to epidemiological or evolutionary dynamics of the *Mycobacterium* circulating in western Kenya.The high prevalence of unknown strains could probably be due to high burden of HIV in western Kenya resulting in high number of immune-compromised persons(Hoshi *etal*., 2016; Pavlinac *etal*., 2015), who facilitate evolution of bacterial pathogens(Kariuki & Dougan, 2014), including *Mycobacterium* organisms.

5.3.Clusters and allelic diversity of *Mycobacterium tuberculosis* complex circulating in western Kenya

MIRU-VNTR cluster analysis grouped the strains into five clusters resulting in a moderate clustering rate of 56.4%. The clusters obtained in the current study were: Beijing, LAM, UgandanI, Ural and EAI with ten, five, three, one and one strains respectively. The clustering rate of 56.4% observed in this study is comparable to those reported in study conducted in Singapore(Chen, 2015). However, this study's findings differed from other clustering rates reported in Canada (75.69%), Hebei-China (18.4%), Iraq (18.03%), Tanga-Tanzania (21.3%) and Taiwan (40.85%)(Ahmed et al., 2014; Chen et al., 2017; Christianson et al., 2010; Hoza, Lupindu, et al., 2016; Li et al., 2016). This difference could be attributed to different assay methods e.g. spoligotyping, 24 and 15-loci MIRU-VNTR and different sample sizes. Being a prospective study, there were no data available to assess possible epidemiologic links

between/among the patients in the above clusters. The moderate clustering rate observed suggest that there is a considerable degree of transmission of *Mycobacterium* strains in western Kenya.

The MIRU-VNTR loci were found to be highly discriminate (mean HGDI was 0.808)in western Kenya. This was comparable to other previous studies in Canada at 0.895 (Christianson, 2010); China at 0.999 (Li, 2016), in Iraq at 0.992 (Ahmed, 2014), in Tanzania at 0.9889 and (Hoza, 2016), Taiwan at 0.9925 (Chen, 2015) and Angola at 0.973(Perdigão *etal*., 2016). The high discriminatory power found in this study suggests MIRU-VNTR alone achieved an acceptable level of resolution for classifying *M. tuberculosis* complex strains circulating in western Kenya and other neighbouring countries.

CHAPTER SIX

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

6.1. Summary

In summary, the present study showed that there were three *Mycobacterium* species namely: *M. tuberculosis M. bovis* and *M. africanum* with *M. tuberculosis* being predominant species were identified. Further, there were unknown species of *Mycobacterium* and those that had mixed matches (multiple matches) of the *Mycobacterium* species. Further, the study revealed that there were a number of *Mycobacterium* complex strains i.e. Beijing, LAM, Uganda I, EAI and the multiple *M. tuberculosis* strains. Moreover, five clusters, namely Beijing, URAL, EAI, Uganda I and LAM were identified from the 40 isolates. The allelic diversity index of *Mycobacterium* strains circulating in western Kenya is high.

6.2. Conclusions and implications

1. There is more than a single species of *M. tuberculosis complex* circulating in Western Kenya. There is need for close monitoring of transmission of species between humans and livestock to minimize zoonotic infection.
2. Western Kenya has a number of *M. tuberculosis* strains in circulation with Beijing strain being predominant. This could predispose the residents to multi-drug resistant strains of tuberculosis.
3. The strains isolated in western Kenya belong to different clusters and clonal complex as well as high allelic HGDI suggesting that there is high genetic diversity among MTC circulating in the region. The HGDI of MIRU-VNTR was high suggesting that MIRU-VINTR alone could be suitable for monitoring Mycobacteria transmission in this region.

6.3. Recommendations

6.3.1 Recommendations from the present study

1. The findings of this study could be applied by the National Tuberculosis and Lung Diseases Program in describing the molecular epidemiology of *M. tuberculosis complex* species circulating in Western Kenya.
2. Tuberculosis disease control should be targeted to specific *M. tuberculosis* strains in both the community and clinical settings.
3. The genetic diversity of *M. tuberculosis* findings from this study could inform policy makers on tuberculosis disease control strategies.

6.3.2. Suggestions for further studies

1. Future studies that use both 24-loci MIRU-VNTR and spoligotyping techniques should be conducted to identify and assign the lineages of the unknown *M. tuberculosis complex* species and strains and further build on this study's findings.
2. Further molecular epidemiology studies should be conducted to trace transmission dynamics and drug resistance of the identified species and strains.
3. A prospective study using a wider population should be conducted to further characterize and determine the genetic diversity of *M. tuberculosis complex.*

REFERENCES

(CDC, C. for D. C. and P. (2012). Integrated prevention services for HIV infection, viral hepatitis, sexually transmitted diseases, and tuberculosis for persons who use drugs illicitly: summary guidance from CDC and the US Department of Health and Human Services. [JOUR]. *MMWR. Recommendations and Reports: Morbidity and Mortality Weekly Report. Recommendations and Reports*, *61*(RR-5), 1.

Ahmed, M. M., Mohammed, S. H., Nasurallah, H. A. A., Ali, M. M., Couvin, D., & Rastogi, N. (2014). Snapshot of the genetic diversity of Mycobacterium tuberculosis isolates in Iraq. *International Journal of Mycobacteriology*, *3*(3), 184–196. https://doi.org/10.1016/j.ijmyco.2014.07.006

Allix-Béguec, C., Fauville-Dufaux, M., & Supply, P. (2008). Three-Year Population-Based Evaluation of Standardized Mycobacterial Interspersed Repetitive-Unit-Variable-Number Tandem-Repeat Typing of Mycobacterium tuberculosis. *JOURNAL OF CLINICAL MICROBIOLOGY*, *46*(4), 1398–1406. https://doi.org/10.1128/JCM.02089-07

Allix-Béguec, C., Harmsen, D., Weniger, T., Supply, P., & Niemann, S. (2008). Evaluation and Strategy for Use of MIRU-VNTRplus, a Multifunctional Database for Online Analysis of Genotyping Data and Phylogenetic Identification of Mycobacterium tuberculosis Complex Isolates. *JOURNAL OF CLINICAL MICROBIOLOGY*, *46*(8), 2692–2699. https://doi.org/10.1128/JCM.00540-08

Allix, C., Walravens, K., Saegerman, C., Godfroid, J., Supply, P., & Fauville-Dufaux, M. (2006). Evaluation of the Epidemiological Relevance of Variable-Number Tandem-Repeat Genotyping of Mycobacterium bovis and Comparison of the Method with IS6110 Restriction Fragment Length Polymorphism Analysis and Spoligotyping †. *JOURNAL OF CLINICAL MICROBIOLOGY*, *44*(6), 1951–1962. https://doi.org/10.1128/JCM.01775-05

Alyamani, E. J., Marcus, S. A., Ramirez-Busby, S. M., Hansen, C., Rashid, J., El-kholy, A., … Talaat, A. M. (2019). Genomic analysis of the emergence of drug-resistant strains of Mycobacterium tuberculosis in the Middle East. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-41162-9

Arnvig, K. B., & Young, D. B. (2009). Identification of small RNAs in Mycobacterium tuberculosis. *Molecular Microbiology*, *73*(3), 397–408.

Azé, J., Sola, C., Zhang, J., Lafosse-Marin, F., Yasmin, M., Siddiqui, R., … Refrégier, G. (2015). Genomics and machine learning for taxonomy consensus: The mycobacterium tuberculosis complex paradigm. *PLoS ONE*, *10*(7). https://doi.org/10.1371/journal.pone.0130912

Balian, D. R., Davtyan, K., Balian, A., Grigoryan, A., Hayrapetyan, A., & Davtyan, H. (2017). Tuberculosis treatment and Smoking, Armenia, 2014–2016. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, *8*, 1–5. https://doi.org/10.1016/J.JCTUBE.2017.04.001

Bazira, J., Asiimwe, B. B., & Joloba, M. L. (2010). Genetic diversity of Mycobacterium tuberculosis in Mbarara, South Western Uganda. In *African Health Sciences* (Vol. 10). Retrieved from https://www.ajol.info/index.php/ahs/article/viewFile/63833/51650

Berrada, Z. L., Lin, S. Y. G., Rodwell, T. C., Nguyen, D., Schecter, G. F., Pham, L., … Desmond, E. (2016). Rifabutin and rifampin resistance levels and associated rpoB mutations in clinical isolates of Mycobacterium tuberculosis complex. *Diagnostic Microbiology and Infectious Disease*, *85*(2), 177–181. https://doi.org/10.1016/j.diagmicrobio.2016.01.019

Bloomfield, G. S., Lagat, D. K., Akwanalo, O. C., Carter, E. J., Lugogo, N., Vedanthan, R., … Sherman, C. B. (2012). Waiting to inhale: An exploratory review of conditions that may predispose to pulmonary hypertension and right heart failure in persons exposed to household air pollution in low- and middle-income countries. *Global Heart*, *7*(3), 249–259. https://doi.org/10.1016/j.gheart.2012.06.015

Borrell, S., Gagneux, S., & Gagneux, S. (2009). STATE OF THE ART STATE OF THE ART SERIES Drug-resistant tuberculosis, Edited by C-Y. Chiang NUMBER 2 IN THE SERIES Infectiousness, reproductive fi tness and evolution of drug-resistant Mycobacterium tuberculosis S U M M A R Y. In *INT J TUBERC LUNG DIS* (Vol. 13). Retrieved from www.theunion.org]

Brosch, R., Gordon, S. V, Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., … Cole, S. T. (2002). *A new evolutionary scenario for the Mycobacterium tuberculosis complex* (Vol. 99). Retrieved from www.pnas.orgcgidoi10.1073pnas.052548299

Burthe, S., Bennett, M., Kipar, A., Lambin, X., SMITH, A., TELFER, S., & BEGON, M. (2008). Tuberculosis (Mycobacterium microti) in wild field vole populations. *Parasitology*, *135*(3), 309–317. https://doi.org/10.1017/S0031182007003940

Carey, A. F., Rock, J. M., Krieger, I. V., Chase, M. R., Fernandez-Suarez, M., Gagneux, S., … Fortune, S. M. (2018). TnSeq of Mycobacterium tuberculosis clinical isolates reveals strain-specific antibiotic liabilities. *PLoS Pathogens*, *14*(3). https://doi.org/10.1371/journal.ppat.1006939

Caulfield, A. J., & Wengenack, N. L. (2016). Diagnosis of active tuberculosis disease: From microscopy to molecular techniques. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, *4*, 33–43. https://doi.org/10.1016/J.JCTUBE.2016.05.005

Charan, J., & Biswas, T. (2013). How to calculate sample size for different study designs in medical research? *Indian Journal of Psychological Medicine*, *35*(2), 121–126. https://doi.org/10.4103/0253-7176.116232

Chen, Y. Y., Chang, J. R., Huang, W. F., Hsu, C. H., Cheng, H. Y., Sun, J. R., … Dou, H. Y. (2017). Genetic diversity of the Mycobacterium tuberculosis East African–Indian family in three tropical Asian countries. *Journal of Microbiology, Immunology and Infection*, *50*(6), 886–892. https://doi.org/10.1016/j.jmii.2015.10.012

Christianson, S., Wolfe, J., Orr, P., Karlowsky, J., Levett, P. N., Horsman, G. B., … Sharma, M. K. (2010). *Evaluation of 24 locus MIRU-VNTR genotyping of Mycobacterium tuberculosis isolates in Canada*. *90*, 31–38. https://doi.org/10.1016/j.tube.2009.12.003

Chuang, H., Lee, E., Liu, Y., Lee, D., & Ideker, T. (2007). Network- based classification of breast cancer metastasis. *Molecular Systems Biology*, *3*(1), 140.

Cochran, W. F. (1977). SamplingTechniques: Chapter 5. *Sampling Techniques*, 10. Retrieved from https://scholar.google.com.tr/scholar?q=sampling+techniques&btnG=&hl=en&as_sdt=0,5#0

Coscolla, M., & Gagneux, S. (2014, December 1). Consequences of genomic diversity in mycobacterium tuberculosis. *Seminars in Immunology*, Vol. 26, pp. 431–444. https://doi.org/10.1016/j.smim.2014.09.012

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the Capacity of Attention: Its Estimation and Its Role in Working Memory and Cognitive Aptitudes. *Cogn Psychol*, *51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

de Jong, B. C., Antonio, M., & Gagneux, S. (2010). Mycobacterium africanum—Review of an Important Cause of Human Tuberculosis in West Africa. *PLoS Neglected Tropical Diseases*, *4*(9), e744. https://doi.org/10.1371/journal.pntd.0000744

Degu, G., & Tessema, F. (2005). Biostatistics for Health Science Students: lecture note series. *The Carter Center 9EPHTI), Addis Ababa*.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V, Maguire, J. R., Hartl, C., … Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*(5), 491–498. https://doi.org/10.1038/ng.806

Detjen, A. K., DiNardo, A. R., Leyden, J., Steingart, K. R., Menzies, D., Schiller, I., … Mandalakas, A. M. (2015). Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: A systematic review and meta-analysis. *The Lancet Respiratory Medicine*, *3*(6), 451–461. https://doi.org/10.1016/S2213-2600(15)00095-8

Dorman, S. E., Schumacher, S. G., Alland, D., Nabeta, P., Armstrong, D. T., King, B., … Xie, Y. (2018). Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance: a prospective multicentre diagnostic accuracy study. *The Lancet Infectious Diseases*, *18*(1), 76–84. https://doi.org/10.1016/S1473-3099(17)30691-6

Drobniewski, F., Cooke, M., Jordan, J., Casali, N., Mugwagwa, T., Broda, A., … Abubakar, I. (2015). Systematic review, meta-analysis and economic modelling of molecular diagnostic tests for antibiotic resistance in tuberculosis. *Health Technology Assessment (Winchester, England)*, *19*(34), 1–188, vii–viii. https://doi.org/10.3310/hta19340

Ei, P. W., Aung, W. W., Lee, J. S., Choi, G. E., & Chang, C. L. (2016). Molecular strain typing of Mycobacterium tuberculosis: A review of frequently used methods. *Journal of Korean Medical Science*, Vol. 31, pp. 1673–1683. https://doi.org/10.3346/jkms.2016.31.11.1673

Fonseca, J. D., Knight, G. M., & McHugh, T. D. The complex evolution of antibiotic resistance in Mycobacterium tuberculosis. , 32 International Journal of Infectious Diseases § (2015).

Foster, N., Vassall, A., Cleary, S., Cunnama, L., Churchyard, G., & Sinanovic, E. (2015). The economic burden of TB diagnosis and treatment in South Africa. *Social Science & Medicine*, *130*, 42–50. https://doi.org/10.1016/J.SOCSCIMED.2015.01.046

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B. C., Narayanan, S., … Small, P. M. (2006). Variable host-pathogen compatibility in Mycobacterium tuberculosis. *PNAS2*, *103*(8), 2869–2873. https://doi.org/10.1073/pnas.0511240103

Getahun, A., Adane, W., Aboma, Z., Ketema, T., & Gobena, A. (2016). Molecular characterization of Mycobacterium tuberculosis complex in Gambella region, Southwest Ethiopia. *African Journal of Microbiology Research*, *10*(29), 1086–1093. https://doi.org/10.5897/ajmr2015.7415

Githui, W. A., Jordaan, A. M., Juma, E. S., Kinyanjui, P., Karimi, F. G., Kimwomi, J., … Victor, T. C. (2004). Identification of MDR-TB Beijing/W and other Mycobacterium tuberculosis genotypes in Nairobi, Kenya. In *INT J TUBERC LUNG DIS* (Vol. 8). Retrieved from http://docserver.ingentaconnect.com/deliver/connect/iuatld/10273719/v8n3/s13.pdf?expires=1561109646&id=0000&titleid=3764&checksum=B880F35DEC4DCE43DD3DDEC38EA9F933

Glaziou, P., Floyd, K., & Raviglione, M. C. (2018). *Global Epidemiology of Tuberculosis*. https://doi.org/10.1055/s-0038-1651492

Glaziou, P., Raviglione, M., Falzon, D., & Floyd, K. (2015). Global Epidemiology of Tuberculosis. *Semin Respir Crit Care Med*, *34*(1), 3–16.

Glynn, J. R., Whiteley, J., Bifani, P. J., Kremer, K., & Van Soolingen, D. (2002). Worldwide Occurrence of Beijing/W Strains of Mycobacterium tuberculosis: A Systematic Review. In *Emerging Infectious Diseases* • (Vol. 8).

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), 2658–2671. https://doi.org/10.1371/journal.pbio.0060311

Hoshi, T., Fuji, Y., Nzou, S. M., Tanigawa, C., Kiche, I., Mwau, M., … Kaneko, S. (2016). Spatial Distributions of HIV Infection in an Endemic Area of Western Kenya: Guiding Information for Localized HIV Control and Prevention. *PloS One*, *11*(2), e0148636. https://doi.org/10.1371/journal.pone.0148636

Hoza, A. S., Lupindu, A. M., Mfinanga, S. G. M., Moser, I., & König, B. (2016). The role of nontuberculous mycobacteria in the diagnosis, management and quantifying risks of tuberculosis in Tanga, Tanzania. *Tanzania Journal of Health Research*, *18*(2). https://doi.org/10.4314/thrb.v18i2.5

Hoza, A. S., Mfinanga, S. G. M., & König, B. (2015). Anti-TB drug resistance in Tanga, Tanzania: A cross sectional facility-base prevalence among pulmonary TB patients. *Asian Pacific Journal of Tropical Medicine*, *8*(11), 907–913. https://doi.org/10.1016/j.apjtm.2015.10.014

Hoza, A. S., Mfinanga, S. G. M., Rodloff, A. C., Moser, I., & 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). https://doi.org/10.1186/s13104-016-1928-3

Hoza, A. S., Mfinanga, S. G., Moser, I., & König, B. (2016). Molecular characterization of Mycobacterium tuberculosis isolates from Tanga, Tanzania: First insight of MIRU-VNTR and microarray-based spoligotyping in a high burden country. *Tuberculosis*, *98*, 116–124. https://doi.org/10.1016/j.tube.2016.02.002

Jagielski, T., Minias, A., van Ingen, J., Rastogi, N., Brzostek, A., Żaczek, A., & Dziadek, J. (2016). Methodological and Clinical Aspects of the Molecular Epidemiology of Mycobacterium tuberculosis and Other Mycobacteria. *Clinical Microbiology Reviews*, *29*(2), 239–290. https://doi.org/10.1128/CMR.00055-15

Kariuki, S., & Dougan, G. (2014). Antibacterial resistance in sub-Saharan Africa: an underestimated emergency. *Ann N Y Acad Sci*, *1323*(1), 43–55. https://doi.org/10.1038/jid.2014.371

Khosravi, A. D., Shahraki, A. H., Dezfuli, S. K., Hashemzadeh, M., Goodarzi, H., & Mohajeri, P. (2017). Genetic diversity of multidrug-resistant Mycobacterium tuberculosis strains isolated from tuberculosis patients in Iran using MIRU-VNTR technique. *Kaohsiung Journal of Medical Sciences*, *33*(11), 550–557. https://doi.org/10.1016/j.kjms.2017.06.011

Kiwanuka, J., Bazira, J., Mwanga, J., Tumusiime, D., Nyesigire, E., Lwanga, N., … Schiff, S. J. (2013). The Microbial Spectrum of Neonatal Sepsis in Uganda: Recovery of Culturable Bacteria in Mother-Infant Pairs. *PLoS ONE*, *8*(8), e72775. https://doi.org/10.1371/journal.pone.0072775

Klopper, M., Warren, R. M., Hayes, C., van Pittius, N. C. G., Streicher, E. M., Müller, B., … Trollip, A. P. (2013). Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa. *Emerging Infectious Diseases*, *19*(3), 449–455. https://doi.org/10.3201/eid1903.120246

Ködmön, C., Zucs, P., & Van Der Werf, M. J. (2016). Migration-related tuberculosis: epidemiology and characteristics of tuberculosis cases originating outside the European Union and European Economic Area, 2007 to 2013. *Euro Surveill*, 1. https://doi.org/10.2807/1560-7917.ES.2016.21.12.30164

Kosgei, R. J., Sitienei, J. K., Kipruto, H., Kimenye, K., Gathara, D., Odawa, F. X., … Carter, E. J. (2015). Gender differences in treatment outcomes among 15-49 year olds with smear-positive pulmonary tuberculosis in Kenya. *International Journal of Tuberculosis and Lung Disease, 19*(10), 1176–1181. https://doi.org/10.5588/ijtld.15.0070

Kremer, K., Arnold, C., Cataldi, A., Cristina Gutiérrez, M., Haas, W. H., Panaiotov, S., … Van Soolingen, D. (2005). Discriminatory Power and Reproducibility of Novel DNA Typing Methods for Mycobacterium tuberculosis Complex Strains. *JOURNAL OF CLINICAL MICROBIOLOGY, 43*(11), 5628–5638. https://doi.org/10.1128/JCM.43.11.5628-5638.2005

Kumar, A., Toledo, J. C., Patel, R. P., Lancaster, J. R., Steyn, A. J. C., Designed, A. J. C. S., & Performed, J. C. T. (2007). *Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor* (Vol. 104). Retrieved from https://www.pnas.org/content/pnas/104/28/11568.full.pdf

Li, Y., Cao, X., Li, S., Wang, H., Wei, J., Liu, P., … Dai, E. (2016). Characterization of Mycobacterium tuberculosis isolates from Hebei, China: Genotypes and drug susceptibility phenotypes. *BMC Infectious Diseases*, *16*(1), 1–11. https://doi.org/10.1186/s12879-016-1441-2

Lönnroth, K., Migliori, G. B., Abubakar, I., D'Ambrosio, L., De Vries, G., Diel, R., … Raviglione, M. C. (2015). Towards tuberculosis elimination: An action framework for low-incidence countries. *European Respiratory Journal*, *45*(4), 928–952. https://doi.org/10.1183/09031936.00214014

Lopez-Avalos, G., Gonzalez-Palomar, G., Lopez-Rodriguez, M., Vazquez-Chacon, C. A., Mora-Aguilera, G., Gonzalez-Barrios, J. A., … Alvarez-Maya, I. (2017). Genetic diversity of Mycobacterium tuberculosis and transmission associated with first-line drug resistance: a first analysis in Jalisco, Mexico. *Journal of Global Antimicrobial Resistance*, *11*, 90–97. https://doi.org/10.1016/j.jgar.2017.07.004

Louw, G. E., Warren, R. M., Donald, P. R., Murray, M. B., Bosman, M., Van Helden, P. D., … Victor, T. C. (2006). Frequency and implications of pyrazinamide resistance in managing previously treated tuberculosis patients. *International Journal of Tuberculosis and Lung Disease, 10*(7), 802–807.

Mbugi, E. V., Katale, B. Z., Siame, K. K., Keyyu, J. D., Kendall, S. L., Dockrell, H. M., … Van Helden, P. D. (2015). Genetic diversity of Mycobacterium tuberculosis isolated from tuberculosis patients in the Serengeti ecosystem in Tanzania. *Tuberculosis*, *95*(2), 170–178. https://doi.org/10.1016/j.tube.2014.11.006

Mbugi, E. V., Katale, B. Z., Streicher, E. M., Keyyu, J. D., Kendall, S. L., Dockrell, H. M., … Rastogi, N. (2016). Mapping of Mycobacterium tuberculosis Complex Genetic Diversity Profiles in Tanzania and Other African Countries. *PloS One*, *11*(5), e0154571. https://doi.org/10.1371/journal.pone.0154571

Merker, M. ;, Blin, C. ;, Mona, S., Duforet-Frebourg, N., Lecher, S., Willery, E., … Wirth, T. (2015). Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. *Nature Genetics*, *47*(3), 242–249. https://doi.org/10.1038/ng.3195

Merker, M., Kohl, T. A., Niemann, S., & Supply, P. (2017). *The Evolution of Strain Typing in the Mycobacterium tuberculosis Complex*. https://doi.org/10.1007/978-3-319-64371-7

Mihret, A, Bekele, Y., Aytenew, M., Assefa, Y., Abebe, M., Wassie, L., … Howe, R. (2012). Modern lineages of Mycobacterium tuberculosis in Addis Ababa, Ethiopia: Implications for the tuberculosis control programe. *African Health Sciences*, *12*(3), 339–344. https://doi.org/10.4314/ahs.v12i3.15

Mihret, Adane, Bekele, Y., Loxton, A. G., Jordan, A. M., Yamuah, L., Aseffa, A., … Walzl, G. (2012). Diversity of Mycobacterium tuberculosis Isolates from New Pulmonary Tuberculosis Cases in Addis Ababa, Ethiopia . *Tuberculosis Research and Treatment*, *2012*, 1–7. https://doi.org/10.1155/2012/892079

MOH. (2015). Kenya national strategic plan for tuberculosis, leprosy and lung health 2015-2018. *Ministry of Health*.

Muthaiah, M., Shivekar, S. S., Cuppusamy Kapalamurthy, V. R., Alagappan, C., Sakkaravarthy, A., & Brammachary, U. (2017). Prevalence of mutations in genes associated with rifampicin and isoniazid resistance in Mycobacterium tuberculosis clinical isolates. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*, *8*, 19–25. https://doi.org/10.1016/j.jctube.2017.06.001

Nduba, V., Van't Hoog, A. H., Mitchell, E. M. H., Borgdorff, M., & Laserson, K. F. (2018). Incidence of Active Tuberculosis and Cohort Retention among Adolescents in Western Kenya. *Pediatric Infectious Disease Journal*, *37*(1), 10–15. https://doi.org/10.1097/INF.0000000000001685

NTLD-P. (2017). *Kenya National Tuberculosis, Leprosy and Lung Disease Program Annual Report*.

Nyamogoba, H. D., & Mbuthia, G. (2018). Gender-age distribution of tuberculosis among suspected tuberculosis cases in western Kenya. *Medical Science*. https://doi.org/10.5455/medscience.2017.06.8735

Ogaro, T. D., Githui, W., Kikuvi, G., Okari, J., Asiko, V., Wangui, E., … Victor, T. C. (2012). Diversity of Mycobacterium tuberculosis strains in Nairobi , Kenya . *African Journal of Health Sciences*, *20*, 82–90.

Ortblad, K. F., Salomon, J. A., Bärnighausen, T., & Atun, R. (2015). Stopping tuberculosis: A biosocial model for sustainable development. *The Lancet*, *386*(10010), 2354–2362. https://doi.org/10.1016/S0140-6736(15)00324-4

Otchere, I. D., Asante-Poku, A., Osei-Wusu, S., Baddoo, A., Sarpong, E., Ganiyu, A. H., … Yeboah-Manu, D. (2016). Detection and characterization of drug-resistant conferring genes in Mycobacterium tuberculosis complex strains: A prospective study in two distant regions of Ghana. *Tuberculosis*, *99*, 147–154. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S1472979216301329

Pankhurst, L. J., del Ojo Elias, C., Votintseva, A. A., Walker, T. M., Cole, K., Davies, J., … Crook, D. W. (2016). Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *The Lancet Respiratory Medicine*, *4*(1), 49–58. https://doi.org/10.1016/S2213-2600(15)00466-X

Pavlinac, P. B., Naulikha, J. M., John-Stewart, G. C., Onchiri, F. M., Okumu, A. O., Sitati, R. R., … Walson, J. L. (2015). Mycobacterium tuberculosis Bacteremia Among Acutely Febrile Children in Western Kenya. *The American Journal of Tropical Medicine and Hygiene*, *93*(5), 1087–1091. https://doi.org/10.4269/ajtmh.15-0365

Perdigão, J., Clemente, S., Ramos, J., Masakidi, P., Machado, D., Silva, C., … Portugal, I. (2016). Genetic diversity, transmission dynamics, and drug resistance of Mycobacterium tuberculosis in Luanda, Angola. *International Journal of Mycobacteriology*, *5*(October 2016), S38–S39. https://doi.org/10.1016/j.ijmyco.2016.09.050

Ramirez-Busby, S. M., & Valafar, F. (2015). Systematic Review of Mutations in Pyrazinamidase Associated with Pyrazinamide Resistance in Mycobacterium tuberculosis Clinical Isolates. *Antimicrobial Agents and Chemotherapy*, *59*(9), 5267–5277. https://doi.org/10.1128/aac.00204-15

Raviglione, M., & Sulis, G. (2016). Tuberculosis 2015: Burden, Challenges and Strategy for Control and Elimination [JOUR]. *Infectious Disease Reports*, *8*(2), 6570. https://doi.org/10.4081/idr.2016.6570

Ritter, C., Lucke, K., Sirgel, F. A., Warren, R. W., van Helden, P. D., Böttger, E. C., & Bloemberg, G. V. (2014). Evaluation of the AID TB resistance line probe assay for rapid detection of genetic alterations associated with drug resistance in Mycobacterium tuberculosis strains. *Journal of Clinical Microbiology*, *52*(3), 940–946.

Rosales-Klintz, S., Jureen, P., Zalutskayae, A., Skrahina, A., Xu, B., Hu, Y., … Hoffner, S. E. (2012). Drug resistance-related mutations in multidrug-resistant Mycobacterium tuberculosis isolates from diverse geographical regions. *International Journal of Mycobacteriology*, *1*(3), 124–130. https://doi.org/10.1016/j.ijmyco.2012.08.001

Saifodine, A., Fyfe, J., Sievers, A., Coelho, E., Azam, K., & Black, J. (2014). Genetic diversity of mycobacterium tuberculosis isolates obtained from patients with pulmonary tuberculosis in beira city, mozambique. *International Journal of Mycobacteriology*, *3*(2), 94–100. https://doi.org/10.1016/j.ijmyco.2014.03.004

Savine, E., Warren, R. M., Van Der Spuy, G. D., Beyers, N., Van Helden, P. D., Locht, C., & Supply, P. (2002). Stability of Variable-Number Tandem Repeats of Mycobacterial Interspersed Repetitive Units from 12 Loci in Serial Isolates of Mycobacterium tuberculosis. *JOURNAL OF CLINICAL MICROBIOLOGY*, *40*(12), 4561–4566. https://doi.org/10.1128/JCM.40.12.4561-4566.2002

Sekizuka, T., Yamashita, A., Murase, Y., Iwamoto, T., Mitarai, S., Kato, S., & Kuroda, M. (2015). TGS-TB: total genotyping solution for Mycobacterium tuberculosis using short-read whole-genome sequencing. *PLoS One*, *10*(11), e0142951.

Shamputa, I. C., Lee, J., Allix-Béguec, C., Cho, E. J., Lee, J. I., Rajan, V., … Barry, C. E. (2010). Genetic diversity of Mycobacterium tuberculosis isolates from a tertiary care tuberculosis hospital in south Korea. *Journal of Clinical Microbiology*, *48*(2), 387–394. https://doi.org/10.1128/JCM.02167-09

Sitienei, J., Nyambati, V., & Borus, P. (2013). The Epidemiology of Smear Positive Tuberculosis in Three TB/HIV High Burden Provinces of Kenya (2003–2009). *Epidemiology Research International*, *2013*, 1–7. https://doi.org/10.1155/2013/417038

Sola, C., Filliol, I., Legrand, E., Lesjean, S., Locht, C., Supply, P., & 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*, *3*, 125–133. https://doi.org/10.1016/S1567-1348(03)00011-X

Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rüsch-Gerdes, S., Willery, E., … van Soolingen, D. (2006). *Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit–Variable-Number Tandem Repeat Typing of Mycobacterium tuberculosis*. *44*(12). https://doi.org/10.1128/JCM.01392-06

Supply, P., Lesjean, S., Savine, E., Kremer, K., Van Soolingen, D., & 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. https://doi.org/10.1128/JCM.39.10.3563-3571.2001

Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., Khanna, V., … Brosch, R. (2013). Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. *Nature Genetics*, *45*(2), 172–179. https://doi.org/10.1038/ng.2517

Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., & Locht, C. (2002). Variable human minisatellite-like regions in the Mycobacterium tuberculosis genome. *Molecular Microbiology*, *36*(3), 762–771. https://doi.org/10.1046/j.1365-2958.2000.01905.x

Teramoto, K., Suga, M., Sato, T., Wada, T., Yamamoto, A., & Fujiwara, N. (2015). Characterization of Mycolic Acids in Total Fatty Acid Methyl Ester Fractions from *Mycobacterium* Species by High Resolution MALDI-TOFMS. *Mass Spectrometry*, *4*(1), A0035–A0035. https://doi.org/10.5702/massspectrometry.a0035

Thacker, N., Pereira, N., Banavali, S., Narula, G., Vora, T., Chinnaswamy, G., … Arora, B. (2014). Epidemiology of blood stream infections in pediatric patients at a Tertiary Care Cancer Centre. *Indian Journal of Cancer*, *51*(4), 438. https://doi.org/10.4103/0019-509X.175311

Wampande, E. M., Mupere, E., Debanne, S. M., Asiimwe, B. B., Nsereko, M., Mayanja, H., … Joloba, M. L. (2013). Long-term dominance of Mycobacterium tuberculosisUganda family in peri-urban Kampala-Uganda is not associated with cavitary disease. *BMC Infectious Diseases*, *13*(1), 484. https://doi.org/10.1186/1471-2334-13-484

Wayne, L. G. (1982). Microbiology of tubercle bacilli. *The American Review of Respiratory Disease*, *125*(3 Pt 2), 31–41. https://doi.org/10.1164/arrd.1982.125.3P2.31

WHO. (2015). Implementing the End Tb Strategy. *Implementing the End Tb Strategy*, 1–141.

WHO. (2017). *Global tuberculosis report 2017*. https://doi.org/ISBN 978-92-4-156551-6

WHO. (2018). Global Tuberculosis Report. *WHO Reports*, (November 2017), 1–4. Retrieved from http://www.who.int/tb/publications/global_report/Exec_Summary_13Nov2017.pdf

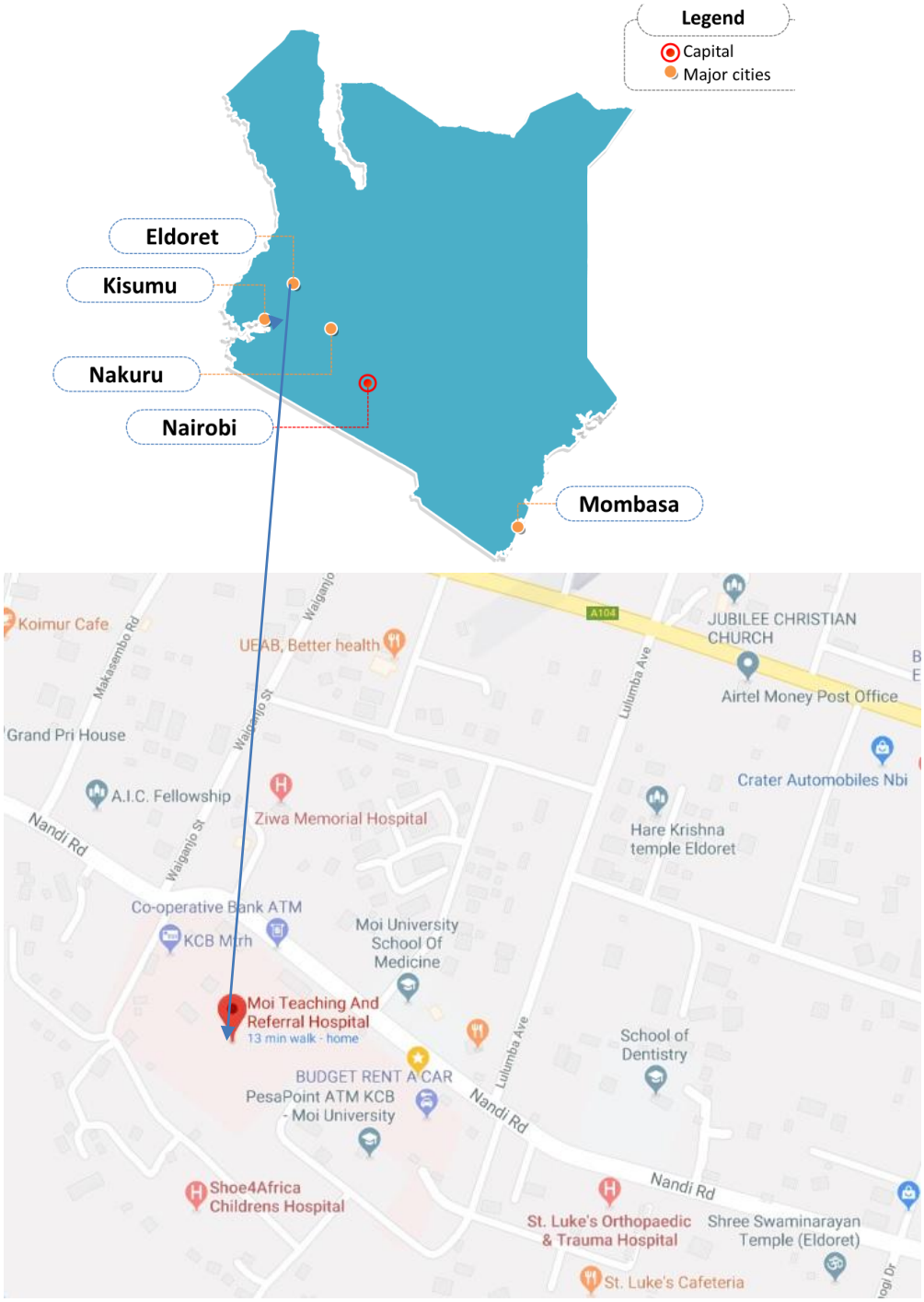
Wirth, T., Hildebrand, F., Allix-Béguec, C., Wölbeling, F., Kubica, T., Kremer, K., … Niemann, S. (2008). Origin, Spread and Demography of the Mycobacterium tuberculosis Complex. *PLoS Pathogens*, *4*(9), e1000160. https://doi.org/10.1371/journal.ppat.1000160

World Health Organization. (2017). Global Tuberculosis Report 2017. In *Who*. https://doi.org/WHO/HTM/TB/2017.23

Zhang, M., Yue, J., Yang, Y. P., Zhang, H. M., Lei, J. Q., Jin, R. L., … Wang, H. H. (2005). Detection of mutations associated with isoniazid resistance in Mycobacterium tuberculosis isolates from China. *Journal of Clinical Microbiology*, *43*(11), 5477–5482. https://doi.org/10.1128/JCM.43.11.5477-5482.2005

APPENDICES

Appendix I: MAP OF STUDY SETTING



Appendix II: Ethical approval

INSTITUTIONAL RESEARCH AND ETHICS COMMITTEE (IREC)

MOI TEACHING AND REFERRAL HOSPITAL
P.O. BOX 3
ELDORET
Tel: 33471/2/3

MOI UNIVERSITY
SCHOOL OF MEDICINE
P.O. BOX 4606
ELDORET

Reference: IREC/2014/110
Approval Number: 0001214

10th July, 2014

INSTITUTIONAL RESEARCH & ETHICS COMMITTEE
10 JUL 2014
APPROVED
P.O. Box 4606-30100 ELDORET

Charles K. Chelimo,
Maseno University,
Department of Zoology,
P.O. Box 333,
MASENO-KENYA.

Dear Mr. Chelimo,

RE: FORMAL APPROVAL

The Institutional Research and Ethics Committee has reviewed your research proposal titled:-

"Genetic Diversity of Clinical Isolates of Mycobacterium Tuberculosis in Western Kenya".

Your proposal has been granted a Formal Approval Number: ***FAN: IREC 1214*** on 10th July, 2014. You are therefore permitted to begin your investigations.

Note that this approval is for 1 year; it will thus expire on 9th July, 2015. If it is necessary to continue with this research beyond the expiry date, a request for continuation should be made in writing to IREC Secretariat two months prior to the expiry date.

You are required to submit progress report(s) regularly as dictated by your proposal. Furthermore, you must notify the Committee of any proposal change (s) or amendment (s), serious or unexpected outcomes related to the conduct of the study, or study termination for any reason. The Committee expects to receive a final report at the end of the study.

Sincerely,

PROF. E. WERE
CHAIRMAN
INSTITUTIONAL RESEARCH AND ETHICS COMMITTEE

cc Director - MTRH Dean - SOP Dean - SOM
Principal - CHS Dean - SON Dean - SOO

Appendix III. MIRU-VNTR Results

ID	Well	0580	0802	2996	0960	3192	1644	0154	4348	2531	2687	2059	3007
1	A01	2	2	1		3		2	2	5	1	2	3
2	B01	3	3					2		5			
3	C01	3	3	5	4	3	3	2	2	5	1	2	3
4	D01	2	1	6		5		2	3	5	1	2	1
5	E01	2	1			4		2	3	5	1	2	3
6	F01	2	1	5	3	3		2	2	8	1	2	2
7	G01	2	1	7	3	5		2	3	5	1	2	1
8	H01	2	4	5	2	2	3	1	2	6	1	2	3
9	A02	2	4	5		2		1	2	6	1	2	3
10	B02	2	1	5	4	3	1	2	2	6	1	2	2
11	C02											2	3
12	D02	3	3									2	
13	E02	2	3	7	3	4	4	2	2	5	1	2	3
14	F02	3	3	4	4	3	3	2	2	5	1	2	3
15	G02												
16	H02	2	1					2					
17	A03	2	2					2	2	5	1	2	3
18	B03	3	3	5	4	3		2	2	5	1	2	3
19	C03												
20	D03	2	3	7		4		2		5	1	2	3
21	E03	2	1	7	3	5	3	2	3	5	1	2	1
22	F03	2	1	7	3	5	3	2	3	5	1	2	1
23	G03	2	1	7	3	5	3	2	3	5	1	2	1
24	H03	2	1	7	3	5	3	2	3	5	1	2	1
25	A04	2	1	7	3	5	3	2	3	5	1	2	1
26	B04	2	1	7	3	5	3	2	3	5	1	2	1
27	C04	5	3	2	4	5	3	2	1	4	2	2	3
28	D04	2	1	7	3	5	3	2	3	5	1	2	1
29	E04	2	1	7	3	5	3	2	3	5	1	2	1
30	F04	5s	5	2	4	6	3	2	2	7	2	2	3
31	G04	2	1	7	3	5	3	2	3	5	1	2	1
32	H04	2	1	7	3	5		2	3	5	1	2	1
33	A05	2	1									2	3
34	B05	2	4	1	7	4	4	2	3	5	1	2	3
35	C05	2	1	4	4	2	3	1	2	6	1	2	3
36	D05	3	3	5	4	3	3	2	2	5	1	2	3
37	E05	2	1	7	3	5	3	2	3	5	1	2	1
38	F05	2	1	8	3	5	3	2	3	5	1	2	1
39	G05	2	1	4	4	3	1	2	2	6	1	2	2
40	H05	2	3	7	3	4	4	2	2	5	1	2	3

Appendix IV: PCR Fluorescent Primers

PCR amplification of each sample was done on 12 different MIRU-VNTR regions using six quardruplex PCRs with fluorescent primers (Table 1) specific for each flanking region (Supply ,., 2001)

Table 2: Fluorescent primers for 12 loci MIRU-VNTR

Locus	Alias	Repeat unit length, bp	PCR primer pair(5' to 3'), with labeling indicated
580	MIRU 4	77	GCGCGAGAGCCCGAACTGC (FAM) GCGCAGCAGAAACGCCAGC
2996	MIRU 26	51	TAGGTCTACCGTCGAAATCTGTGAC CATAGGCGACCAGGCGAATAG (VIC)
802	MIRU 40	54	GGGTTGCTGGATGACAACGTGT (NED) GGGTGATCTCGGCGAAATCAGATA
960	MIRU 10	53	GTTCTTGACCAACTGCAGTCGTCC GCCACCTTGGTGATCAGCTACCT (FAM)
1644	MIRU 16	53	TCGGTGATCGGGTCCAGTCCAAGTA CCCGTCGTGCAGCCCTGGTAC (VIC)
3192	MIRU 31	53	ACTGATTGGCTTCATACGGCTTTA GTGCCGACGTGGTCTTGAT (NED)
154	MIRU 02	53	TGGACTTGCAGCAATGGACCAACT TACTCGGACGCCGGCTCAAAAT (FAM)
2531	MIRU 23	53	CTGTCGATGGCCGCAACAAAACG (VIC) AGCTCAACGGGTTCGCCCTTTTGTC
4348	MIRU 39	53	CGCATCGACAAACTGGAGCCAAAC CGGAAACGTCTACGCCCCACACAT (NED)
2059	MIRU 20	77	TCGGAGAGATGCCCTTCGAGTTAG (FAM) GGAGACCGCGACCAGGTACTTGTA
2687	MIRU 24	54	CGACCAAGATGTGCAGGAATACAT GGGCGAGTTGAGCTCACAGAA (VIC)
3007	MIRU27	53	TCGAAAGCCTCTGCGTGCCAGTAA GCGATGTGAGCGTGCCACTCAA (NED)

Appendix V: M. tuberculosis complex strains isolated from all the samples.

Sample Number	*M. tuberculosis* strain
1	Mixed Match (URAL, Harleem, TUR)
2	Unknown
3	Mixed Match (Cameroon, Harleem)
4	Unknown
5	Mixed Match (Ghana, Delhi/CAS)
6	Unknown
7	Unknown
8	LAM
9	LAM
10	LAM
11	Bovis
12	Unknown
13	Uganda I
14	Unknown
15	Unknown
16	Mixed Match (Uganda II, West African I & II)
17	Uganda I
18	Unknown
19	Unknown
20	Beijing
21	Beijing
22	Beijing
23	Beijing
24	Beijing
25	Beijing
26	Beijing
27	EAI
28	Beijing
29	Beijing
30	Unknown
31	Beijing
32	Unknown
33	Mixed Match (West African I & II)
34	Unknown
35	LAM
36	Mixed Match (Cameroon, Harleem)
37	Beijing
38	Unknown
39	LAM
40	Uganda I