Aus dem Forschungszentrum Borstel
eingereicht über das Institut für Mikrobiologie und Tierseuchen des
Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Tuberculosis in animals and humans in Eastern Sudan and the genetic diversity among clinical strains of *Mycobacterium tuberculosis* complex lineage 3

Inaugural-Dissertation
zur Erlangung des Grades eines
PhD of Biomedical Sciences
an der
Freien Universität Berlin

vorgelegt von
Yassir Adam Shuaib
Tierarzt aus El-Obeid, Sudan

Berlin 2019
Journal-Nr.: 4138

mbvberlin | mensch und buch verlag

Aus dem Forschungszentrum Borstel

eingereicht über das Institut für Mikrobiologie und Tierseuchen des

Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Tuberculosis in animals and humans in Eastern Sudan and the genetic diversity among clinical strains of *Mycobacterium tuberculosis* complex lineage 3

Inaugural-Dissertation

zur Erlangung des Grades eines

PhD of Biomedical Sciences

an der

Freien Universität Berlin

vorgelegt von

Yassir Adam Shuaib

Tierarzt aus El-Obeid, Sudan

Berlin 2019

Journal-Nr.: 4138

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Dekan:	Univ.-Prof. Dr. Jürgen Zentek
Erster Gutachter:	Prof. Dr. Lothar H. Wieler
Zweiter Gutachter:	Prof. Dr. Stefan Niemann
Dritter Gutachter:	Univ.-Prof. Dr. Thomas Alter

Deskriptoren (nach CAB-Thesaurus):
ruminants, man, tuberculosis, carcasses, public health, meat hygiene, meat production, food safety, epidemiology, prevalence, polymerase chain reaction, microscopy, Sudan

Tag der Promotion: 18.10.2019

Bibliografische Information der *Deutschen Nationalbibliothek*
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://dnb.de> abrufbar.

ISBN: 978-3-96729-037-0
Zugl.: Berlin, Freie Univ., Diss., 2019
Dissertation, Freie Universität Berlin
D188

Choriner Str. 85 - 10119 Berlin
verlag@menschundbuch.de – www.menschundbuch.de

Table of content

List of tables

List of figures

1 Introduction

1.1 History of TB

Tuberculosis (TB), also known as the “white plague”, is a chronic infectious disease of both humans and animals and is characterized by a progressive development of granulomatous lesions in the affected organs[1,2]. TB is caused by bacteria of the *Mycobacterium tuberculosis* complex (MTBC) which mainly form granulomas in the lungs (pulmonary TB) but can also affect other parts of the body, e.g. lymph nodes, bones, and brain in case of extra-pulmonary TB[1,2].

The history of MTBC bacteria is shaped by ages of co-evolution and adaption to their hosts[3,4]. Historic specimens with indications of TB disease have been found in Egyptian mummies and the first ancestor of all MTBC strains is estimated to have evolved in East Africa 40,000–70,000 years ago[3,5,6]. Along with human migration waves out of Mesopotamia, MTBC strains were distributed to different parts of the world and diversified into different major human- and animal-adapted lineages[7-9]. Recent studies have shown that the animal species indeed evolved later than the ancient human pathogens[3,6].

In 1882, a German scientist named Robert Heinrich Hermann Koch identified the bacterium that causes TB[5]. Koch continued his revolutionary research on TB, in 1890, he announced the isolation of a substance from tubercle bacilli that he called tuberculin[5]. Koch indicated that tuberculin could render TB bacterium non-pathogenic and harmless. Injections of tuberculin rapidly came into vogue as a treatment for TB[5]. However, it was ineffective and unsuccessful[5]. After conducting specific experiments, Koch concluded that tuberculin might perhaps rather be useful in the diagnosis of TB because of its antigenic characteristics and ability to provoke an immunological reaction in individuals who were exposed to the bacterium[5].

For control of TB, the *Mycobacterium bovis*-derived Bacillus Calmette-Guérin (BCG) vaccine is effective at preventing severe childhood forms of TB; however, it is largely ineffective in adults[10]. At present, many vaccine candidates are in clinical trials[10]. These vaccines are either designed to replace BCG (pre-exposure vaccine) and to be given in infancy or adolescence to augment BCG-mediated protection (prime–boost strategy), or to shorten or potentiate treatment (therapeutic vaccine)[10]. Categories of vaccines include live attenuated *M. tuberculosis*, re-engineered BCG, mycobacterial whole cells or extract, adjuvant-complexed single or fusion tuberculosis-specific proteins, and mycobacterial proteins expressed through a viral vector or plasmid DNA[10,11].

1.2 Global epidemiology of human TB

Today, TB is ranked as one of the top 10 causes of death worldwide and the top leading cause of death due to a single pathogen[12-14]. In 2017, an estimated 10 million (133 per 100,000 population) incident cases of TB were reported worldwide[14]. Most of these newly afflicted individuals were males (64%) and adults (90%). Furthermore, most of these incident cases of TB occurred in low- and middle-income countries and probably more than a third remained undiagnosed[12,14].

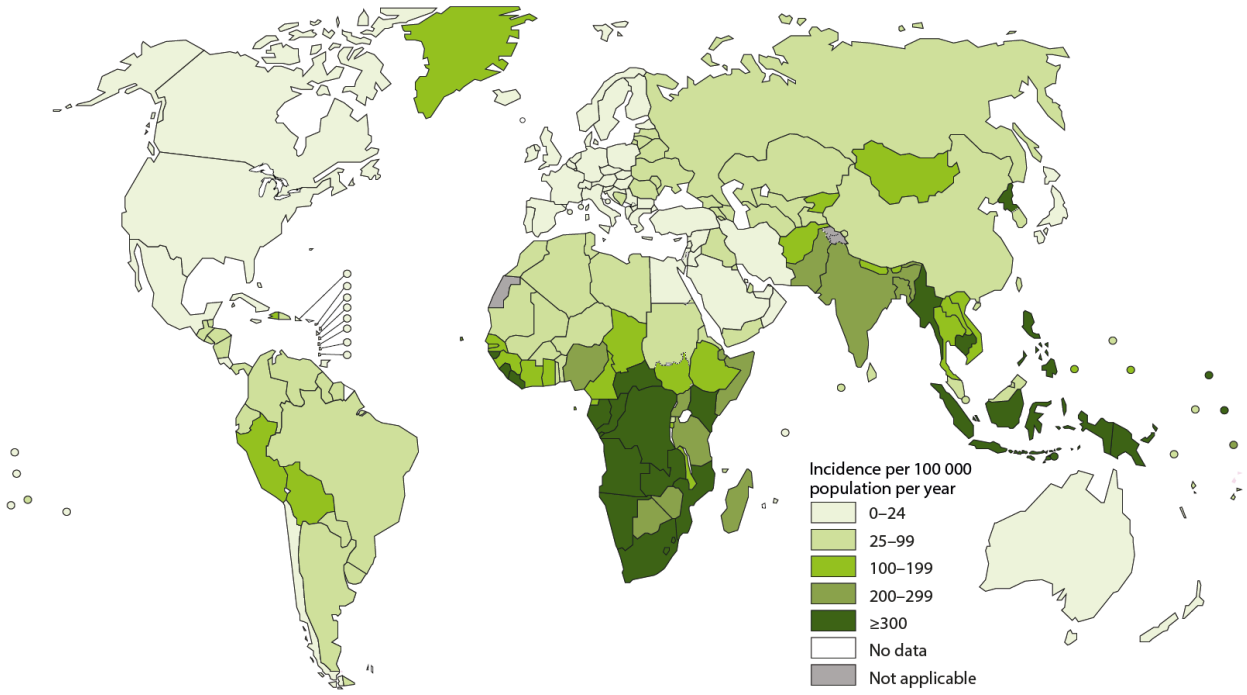


Figure 1.1: Estimated incident TB cases in 2017 by country. Most of the reported cases in 2017 occurred in Africa and Asia. The largest numbers of new cases of TB were reported from two African countries (i.e. Nigeria and South Africa) and six Asian countries (i.e. India, Indonesia, China, the Philippines, Bangladesh, and Pakistan). Source: WHO[14].

Africa accounted for 25% of all worldwide reported incident cases of TB in 2017[14]. This demonstrates the huge magnitude of TB in the continent. The rest of global TB burden in 2017 was distributed between South East Asia (44%), the Western Pacific region (18%), the Eastern Mediterranean region (7.7%), the European region (2.7%), and the region of the Americas (2.8%) (Figure 1.1)[14]. Two African countries in particular, Nigeria and South Africa, were among the seven countries that had the largest number of reported incident cases[14]. Incident cases of TB ranged from under 10 up to 300 per 100,000 population in low- and high-TB–burden countries. Lesotho, Mozambique, and South Africa had more than 500 incident cases of TB per 100,000 population[14]. Nevertheless, the current World Health Organization (WHO) country estimates may contain uncertainties mediated by limitations in laboratory and diagnostics infrastructures, poor case

detection, and insufficient recording and reporting systems[15]. A good example of such uncertainties is the underestimation of TB burden in Nigeria[15]. A recent national survey of TB in the country showed that TB incidence, prevalence, and mortality were two to four times higher than the estimates declared by the WHO[15].

Worldwide, there has been a noticeable ominous increase in the number of incident cases of multi-drug resistant (MDR) and extensively drug resistant (XDR) TB[11]. MDR TB is due to infection with MTBC strains that are resistant to at least rifampicin (RMP) and isoniazid (INH), but XDR TB is caused by MTBC strains that are resistant to RMP (RR), INH, plus any fluoroquinolone (FQ) and at least one of the three injectable second-line TB drugs—amikacin (AMK), kanamycin (KAN), and capreomycin (CPR)[10,11,16,17]. Currently, there is a growing concern about the emergence of totally drug resistant (TDR) TB (i.e. TB that is caused by MTBC strains that are resistant to all first and second line drugs including INH, RMP, streptomycin [SM], ethambutol [EMB], pyrazinamide [PZA], ethionamide [ETH], para-aminosalicylic acid [PAS], cycloserine [CLS], ofloxacin [OFX], ciprofloxacin [CFX], and AMK, KAN, and CPR)[10,11,18]. In 2017, an estimated 3.5% of incident cases and 18% of previously treated cases of MDR/RR TB were reported from around the globe (Figure 1.2)[14]. China, India, and Russia accounted for the largest numbers of MDR/RR cases of TB (47% of the global total) (Figure 1.2)[14].

In 2017, XDR TB was reported in 127 different countries[14]. Existing data suggest that 8.5% of MDR TB cases also had XDR TB[14]. In Africa, an XDR TB epidemic was recently detected in South Africa[19,20]. Concurrently, there have also been increasing reports of XDR TB in Europe, particularly in Belarus and Lithuania, while TDR TB cases have been found in South Africa, Italy, Iran, and India[10-12,18].

Human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS) is recognized as a disease that enhances the risk of MTBC infection and the reactivation of latent TB[10,11]. It is considered one of the important risk factors that increase the number of TB infections and stand as an obstacle towards TB control[10,11,14]. In 2017, 9% of the incident cases of TB were HIV-positive patients[14]. The proportion of TB cases living with HIV was highest in the WHO African region and exceeded 50% in the southern part of the continent[14].

It has been hypothesized that TB kills one person every 16 seconds or so[4,21]. Thus, TB has killed more than one billion people during the last 200 years[4,21]. In 2017, about 1.3 million case fatalities amongst HIV-negative individuals were caused by TB, in addition to 0.3 million deaths among HIV-positive patients[14]. Generally, as high as 95% of these case fatalities occurred in low- and

middle-income countries, besides; 82% of all of the reported TB-related deaths were among HIV-negative individuals in Africa and South East Asia[14].

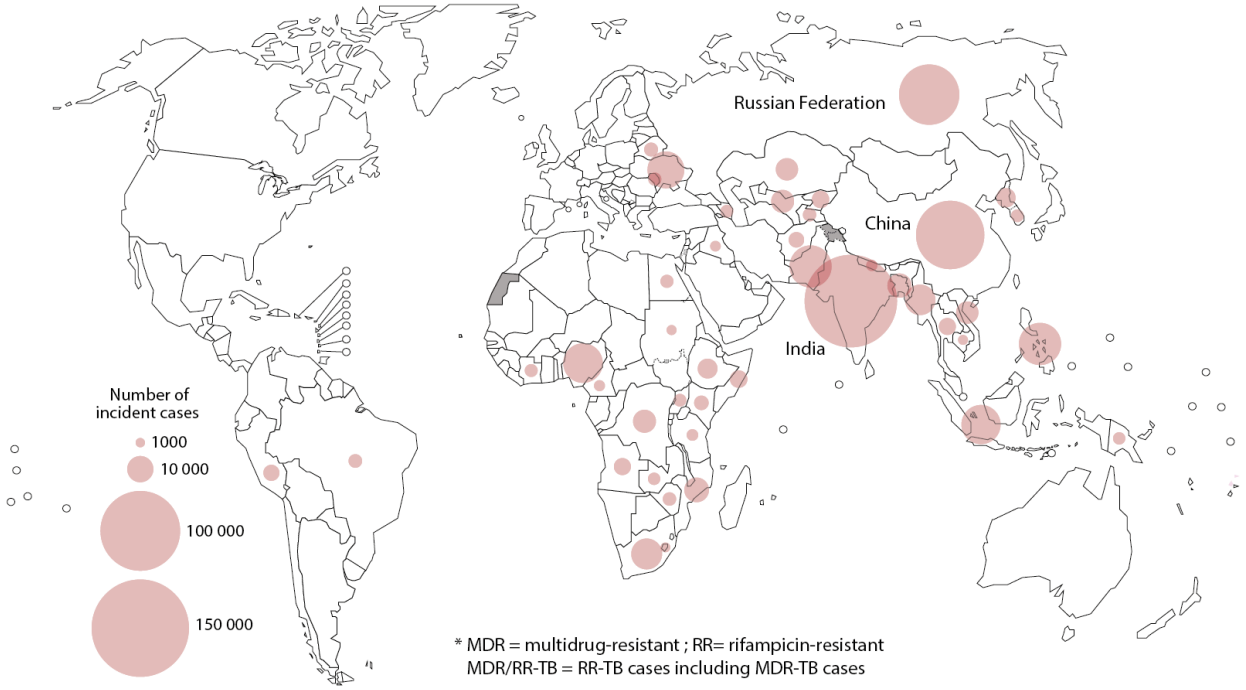


Figure 1.2: Estimated incident cases of MDR/RR TB by country. Globally, 3.5% of incident and 18% of previously treated cases had MDR/RR TB in 2017. Source: WHO[14].

1.3 Causative agents of TB

TB is caused by members of the MTBC that contains a number of highly related sub-species of mycobacteria as defined by several typing methods[8,22-24]. Members of MTBC are characterized by 99.9% similarity at the nucleotide level and identical 16S rRNA sequences and are divided into human- and animal-adapted MTBC and *M. canettii*[3,4,25].

Human-adapted MTBC species include *M. tuberculosis*, which is the primary and the predominant causative agent of human TB, and *M. africanum*[4,8,24]. In general, MTBC strains are characterized by a slow growth rate (i.e. doubles once per day) and by being chemo-organotrophic, non-motile, non-spore-forming, aerobic bacillus, and ability to produce niacin and reduce nitrate[26]. Acid-fast stains, such as the Ziehl–Neelsen (ZN) stain, should be used for smear microcopy as the cell wall of *M. tuberculosis* and *M. africanum* retains carbol fuschin stain in the face of acid-alcohol washes[26].

Animal-adapted MTBC (such as *M. bovis*, *M. caprae*, *M. microti*, and *M. pinnipedii*) as well as the newly detected MTBC (including *M. suricattae*, *M. orygis*, *M. mungi*, Dassie bacillus, and

Chimpanzee bacillus) are the causative agents of TB in domestic and wild animals[1,4,9,25,27]. *M. bovis* has a wide host range[1]. It can survive for up to 2 years in covered dungs and for up to 5 months under diffuse sunlight[1]. Moreover, it can be isolated after 8 weeks from contaminated water, carcasses of slaughtered animals, and feces of infected animals[1]. *M. caprae* was first isolated from carcasses of goats, *M. microti* was reported from small rodents, and *M. pinnipedii* infections were diagnosed in marine animals[1,9]. The newly detected animal-adapted MTBC were isolated from meerkats (*Suricata suricatta*), banded mongooses (*Mungos mungo*), rock hyraxes (dassies, *Procavia capensis*), and chimpanzees (*Pan troglodytes*)[1,4,9,25,27].

M. canettii is characterized by unusually smooth and glossy colonies[4]. All known cases of TB caused by *M. canettii* have been contracted in the Horn of Africa[4]. However, the prevalence of *M. canettii* may be underestimated in the routine microbiology laboratory due to their close similarity to *M. tuberculosis*[4].

1.4 TB transmission and pathogenesis in humans and animals

TB is an airborne disease[28]. Its causative agents are excreted into the environment by individuals with active TB when coughing or sneezing (i.e. index cases), in fine airborne particles (1-5 µm) called droplet nuclei. These droplets can remain suspended in the air for several hours[28,29]. Transmission occurs when a susceptible individual inhales the droplet nuclei[28,29].

Inhaled droplet nuclei navigate through the air passages to reach the terminal alveoli, where the MTBC bacteria are taken up by phagocytic cells such as macrophages, dendritic cells, and non-phagocytic cells in the alveolar space[28-30]. Macrophages respond to the infection and produce pro-inflammatory cytokines and chemokines to drive the recruitment of different types of leukocytes[28-30]. Neutrophils and monocytes migrate to the site of the infection and participate in the ongoing process of phagocytosis[28-30]. However, the MTBC bacteria are able to escape the killing mechanisms by preventing phagosome-lysosome fusion, resulting in the death of the macrophages instead[28-30]. Meanwhile, MTBC-loaded dendritic cells migrate to the nearest lymph node and stimulate T cells against mycobacterial antigens. The primed T cells migrate to the site of the MTBC invasion, where the accumulation of cells of different types results in the development of a granulomatous lesion (Figure 1.3). Phagocytic cells full of viable MTBC bacteria may cross the alveolar barrier to cause extrapulmonary or miliary TB before the development of the adaptive immune responses[28-30].

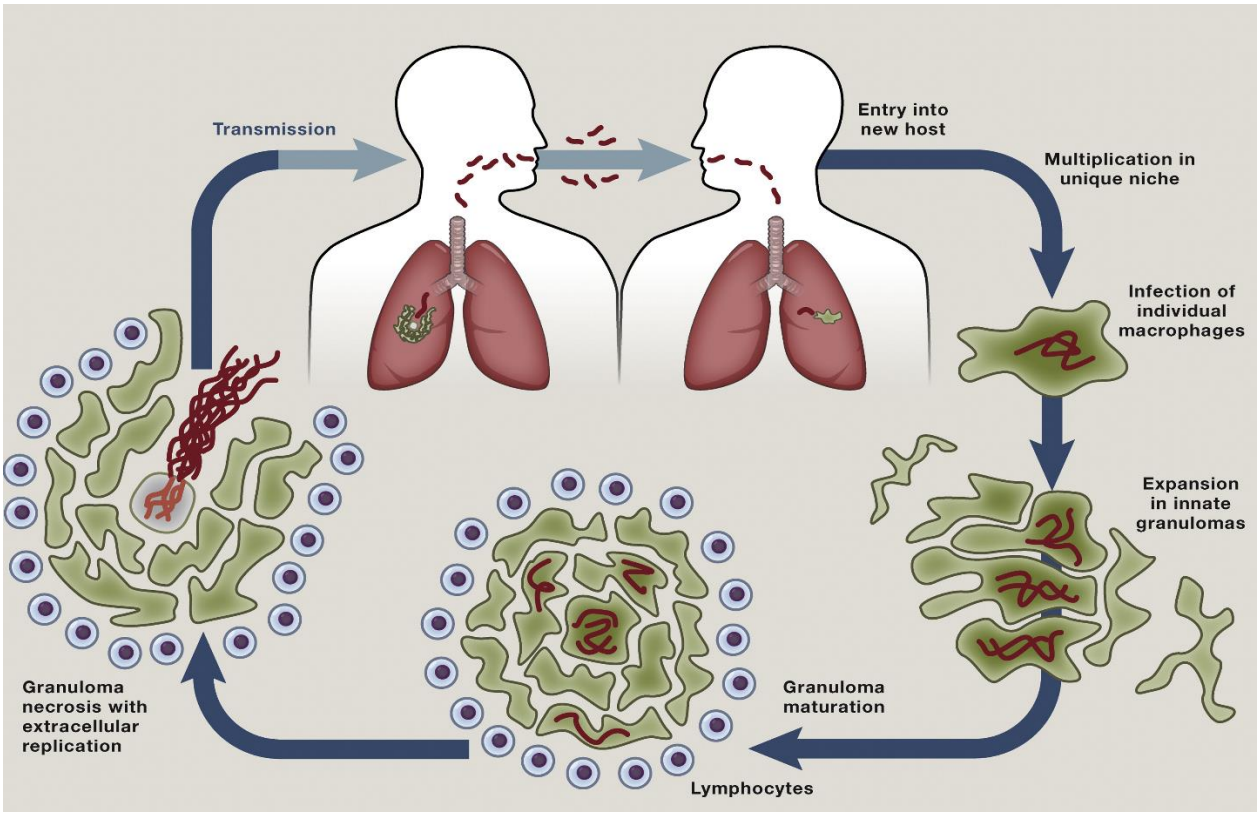


Figure 1.3: MTBC transmission and formation of the granulomatous lesions. After successful transmission of MTBC pathogens to a new host, infection starts by recruiting immune cells to the site of the invasion. Elimination of MTBC pathogens fails and more immune cells are attracted. Aggregation of immune and other cells leads to the formation of the granuloma. As adaptive immunity develops, the granuloma can restrict bacterial growth. However, under many circumstances, the infected granuloma macrophages can undergo necrosis, forming a necrotic core that supports bacterial growth and transmission to the next host. Source: Cambier et al.[29].

Of all individuals who inhale MTBC bacteria, only 5% develop active infectious TB and the vast majority (i.e. over 90%) of them recover upon treatment; however, half would die if no medical care and treatment were provided to them (Figure 1.4)[31]. The remaining 95% become asymptomatic but latently infected with TB as they are able to develop an effective cell-mediated immune response that controls further multiplication of the bacteria[31]. Any subsequent defect in cell-mediated immunity may result in reactivation of dormant MTBC bacilli, causing active disease. Several risk factors (e.g. HIV, diabetes, old age, and malnutrition) increase the risk of reactivation and progression to active TB, which affects the lungs most frequently but may also occur in any organ[31].

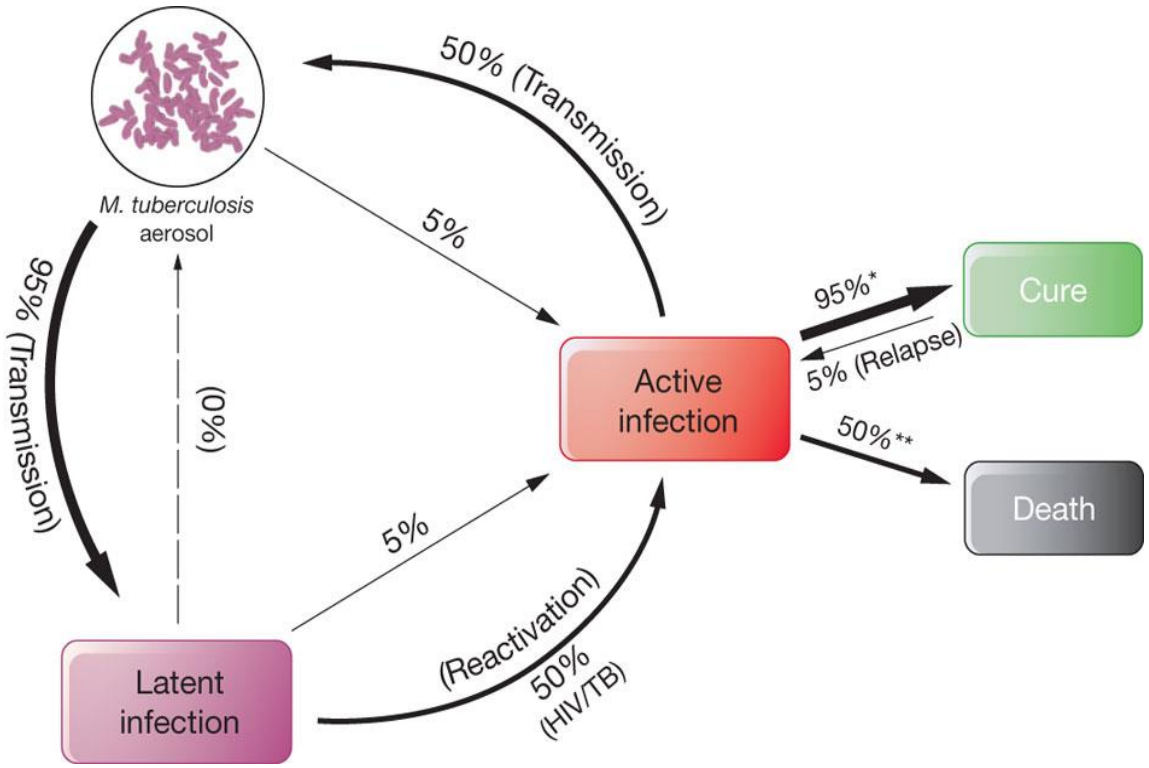


Figure 1.4: MTBC infection and the scenarios of progression to active or latent TB. Following inhalation of MTBC bacteria, 5% of the individuals develop active infectious TB and the rest become latently infected but remain with a high risk of relapse to active TB in case of immune suppression. Ninety five percent of patients with drug susceptible active TB recover upon treatment whereas 5% relapse. Nonetheless, if active TB patients are untreated, high mortality should be expected. Source: Koul et al.[31].

TB pathogenesis in animals is as complex as in humans involving a multiplicity of interactions between the host and the MTBC pathogens. In addition to inhalation, transmission of the pathogens can also be by ingestion[1]. After successful entry and tissue invasion, the MTBC bacteria form a primary complex, and a visible primary focus or lesion develops within eight days of entry at the point of entry[1]. Calcification of the lesions commences about two weeks later. The developing necrotic focus is soon surrounded by granulation tissue, monocytes, and plasma cells and the pathognomonic tubercle lesion is established. Some of the MTBC bacteria pass from this primary focus to a regional lymph node and lead to the development of a similar lesion there. Post-primary dissemination from the primary complex may take the form of acute miliary TB and discrete nodular lesions in various organs[1].

1.5 Diagnosis of TB

Rapid and accurate diagnosis of TB, TB/HIV, and drug resistant TB is critical for timely initiation of treatment and, ultimately, control of the disease[12,32]. TB should be suspected if a patient is

presenting with chronic cough that has lasted for more than two weeks, hemoptysis (or coughing up blood), night sweats, fatigue, loss of appetite and unexplained weight loss, high fever, and chest pains[33]. In case of extrapulmonary TB, symptoms are related to the affected body system[33]. Chest radiography (CXR) and computerized tomography (CT)-scan are very helpful tools for TB diagnosis in clinics[33]. For confirming suspected TB cases in resource-limited countries, ZN or fluorescent stained smears are routinely used[34]. Screening of two consecutive sputum specimens (spot-morning) is able to detect 95%–98% of smear-positive TB patients[34,35]. The use of cultures such as mycobacterium growth indicator tube (MGIT), Löwenstein–Jensen (LJ), Stonebrink, Ogawa, and Middlebrook 7H10 is the gold standard for the detection of TB[34]. However, cultures take up to eight weeks before the results are available[36]. For initial isolation and drug susceptibility testing (DST), liquid media is probably more useful than solid media, as it is faster in producing results and has more sensitivity[36]. Furthermore, rapid identification of TB bacteria can be done by a number of molecular methods, including polymerase chain reaction (PCR), Xpert MTB/RIF (Cepheid), line probe assays (LPAs), and sequencing of specific DNA targets[34,36].

1.6 Treatment of TB

Treatment of drug susceptible TB can be achieved by administration of INH, RMP, PZA, and EMB for two months (intensive phase) followed by four months of INH and RMP (continuation phase)[10,11,16,37-39]. For treatment of MDR TB, the WHO-recommended empirical regimen includes administration of a later-generation FQ such as moxifloxacin (MXF), gatifloxacin (GTF), or levofloxacin (LVF), an injectable aminoglycoside like AMK, CPR, or KAN, plus any first line drug to which the strain is susceptible in addition to either CLS, PAS, terizidone (TRZ), protionamide (PTN), or ETH[10,11,38]. The duration of the intensive phase should not be less than eight months whilst the continuation phase is 12–18 months. Monitoring of the effectiveness of the treatment should be done by culture conversion for 18 months after the date of the first negative culture[16,38]. Bedaquiline (BDQ) and delamanid (DLM) are new drugs that have been recently approved for treatment of MDR TB cases[40]. Before prescription of any anti-TB treatment, DST is strongly recommended to differentiate drug susceptible from drug resistant TB[16]. Regimens for treatment of XDR TB are tailored to individual patients based on DST results[16]. In the intensive phase, at least six active drugs should be used, whereas for the continuation phase the usage of four drugs is recommended[10,11,16,38]. Treatment of XDR TB can be also achieved by using linezolid (LIN), high-dose INH, and clofazimine (CLF) [10,11,16,38]. Cases with quinolone-sensitive strains of MTBC have improved cure rates. However, outcomes are poor, especially in TB/HIV cases that should receive antiretroviral therapy as well[10,11,16,38].

1.7 Drug resistance mechanisms among MTBC strains

Anti-TB drug resistance can be either primary (intrinsic) or secondary[19,41]. Primary drug resistance occurs when MTBC strains have never encountered the drug of interest in a particular host and are already resistant (e.g. *M. bovis* is naturally resistant to PZA) or in patients that have been infected with an already resistant MTBC strain[19]. On the other hand, secondary resistance (non-intrinsic) describes the resistance that arises in MTBC strains after being exposed to the drug (i.e. development of resistance during treatment)[19]. It occurs due to mutations that arise under selective pressure of antibiotic use, mainly as a result of inadequate or interrupted treatment[41,42]. Accumulation of different mutations leads to clinical poly- or multi-drug resistances[41,42]. Resistances to anti-TB drugs have been linked to mutations in a number of genes as summarized in Table 1.1[42,43].

Table 1.1: Resistance genes and their associated antibiotics

Drug	Gene	Rv No.	Gene start	Gene stop	Type
SM	*rpsL*	Rv0682	781560	781934	coding
SM	*gidB*	Rv3919c	4408202	4407528	coding
SM, AMK, KAN, CPR	*rrs*	Rvnr01	1471846	1473382	rRNA
SM, AMK, KAN, CPR	*rrs* upstream		1471743	1471845	intergenic
INH	*katG*	Rv1908c	2153889	2156111	coding
INH	*furA*	Rv1909c	2156149	2156592	coding
INH	*kasA*	Rv2245	2518115	2519365	coding
INH	*nat*	Rv3566c	4007331	4008182	coding
INH	*ahpC*	Rv2428	2726193	2726780	coding
INH, ETH	*mshA*	Rv0486	575348	576790	coding
INH, ETH	*fabG1* upstream		1673300	1673439	intergenic
INH, ETH	*fabG1*	Rv1483	1673440	1674183	coding
INH, ETH	*inhA* upstream		1674184	1674201	intergenic
INH, ETH	*inhA*	Rv1484	1674202	1675011	coding
INH, ETH	*ndh*	Rv1854c	2103042	2101651	coding
RMP	*rpoB*	Rv0667	759807	763325	coding
RMP	*rpoC*	Rv0668	763370	767320	coding
EMB	*embA*	Rv3794	4243233	4249810	coding
EMB	*embB*	Rv3795	4246514	4249810	coding
EMB	*embC*	Rv3793	4239863	4243147	coding
EMB	*aftA*	Rv3792	4239863	4243147	coding
EMB	*embA* upstream		4243148	4243232	intergenic
PZA	*pncA*	Rv2043c	2289241	2288681	coding
PZA	*rpsA*	Rv1630	1833542	1834987	coding
PZA	*rpsA* upstream		1833380	1833541	Intergenic
CPR	*tlyA*	Rv1694	1917940	1918746	coding
LIN, CPR	*rrl* upstream		1473383	1473657	intergenic
LIN, CPR	*rrl*	Rvnr02	1473658	1476795	rRNA
LIN	*rplC*	Rv0701	800809	801462	coding
CLS	*ald*	Rv2780	3086820	3087935	coding
FQs	*gyrA*	Rv0006	7302	9818	coding
FQs	*gyrB*	Rv0005	5240	7267	coding
PAS	*thyA*	Rv2764c	3073680	3074471	coding
PAS	*folC*	Rv2447c	2746135	2747598	coding

Table 1.1 continued

Drug	Gene	Rv No.	Gene start	Gene stop	Type
DLM	*fbiA*	Rv3261	3640543	3642881	coding
DLM	*fbiB*	Rv3262	3640543	3642881	coding
DLM	*fbiC* upstream		1302682	1302930	intergenic
DLM	*fbiC*	Rv1173	1302931	1305501	coding
BDQ	*atpE* upstream		1460997	1461044	intergenic
BDQ	*atpE*	Rv1305	1461045	1461290	coding

No. = number, SM = streptomycin, AMK = amikacin, KAN = kanamycin, CPR = capreomycin, INH = isoniazid, ETH = ethionamide, RMP = rifampicin, EMB = ethambutol, PZA = pyrazinamide, LIN = linezolid, CLS = cycloserine, FQs = fluoroquinolones, PAS = para-aminosalicylic acid, DLM = delamanid, and BDQ = bedaquiline

1.8 Molecular typing of MTBC strains

1.8.1 Spoligotyping and MIRU-VNTR typing

Spacer oligonucleotide typing (spoligotyping) is a PCR-based technique used for genotyping and studying the geographical distribution of MTBC strains. It is based on polymorphisms at the direct repeat (DR) locus[8,44-46]. This locus consists of a series of well-conserved 36 bp DRs that are separated by unique and non-repetitive spacer sequences of 34–41 bp[44,45]. For typing MTBC strains, a PCR amplification of the whole DR locus is done. Amplicons are hybridized to a membrane with 43 covalently bound synthetic oligonucleotides representing the polymorphic spacers identified in *M. tuberculosis* H37Rv and *M. bovis* BCG. The hybridization signals are detected by chemiluminescence and the spacers are visualized. The number of present/absent spacers is used to identify individual clinical strains[8,45-47]. Spoligotyping has many advantages, including simplicity, high reproducibility, and economical feasibility. Furthermore, it requires a small amount of DNA (20 to 50 ng) that might be extracted directly from acid fast bacilli (AFB) positive smears as well as the fact that results are reported in a simple manner (binary or octal numeral). However, the main limitation of this technique is that it has less discriminatory power than mycobacterial interspersed repetitive units-variable number of tandem repeats (MIRU-VNTR) typing and whole genome sequencing (WGS)[45-47].

MIRU-VNTRs are mini-satellite–like structures within the genome of MTBC bacteria that are composed of 40–100 bp repetitive sequences[8,47]. MIRU-VNTRs are scattered into 41 locations throughout the genome[8,46,47]. Based on the sequence analysis of each of those MIRU-VNTR loci, 24 were found to display variations in tandem repeat copy numbers and were thus selected for genotyping of MTBC clinical strains[44,45]. The MIRU-VNTR technique involves PCR amplification of each MIRU locus[47]. The final result is a multi-digit numerical code corresponding to the repeat number at each analyzed locus. The development of this method has permitted mmediate mapping of MTBC transmission, comparison of inter-laboratory data, and global database

construction[44-46]. MIRU-VNTR typing is a reliable and efficient technique[44-46]. Its discriminatory power is typically proportional to the number of loci evaluated[44-46]. For the past 15 years, MIRU-VNTR typing has been the gold standard for MTBC genotyping[44-46].

Spoligotyping and MIRU-VNTR typing results can be analyzed and visualized using different mathematical models, on-line tools, and websites (e.g. MIRU-VNTR*plus* database), besides to computer-based calculations and software programs (e.g. BioNumerics software). These analysis and visualization tools group MTBC strains according to similarity and hierarchical relationship[48].

1.8.2 Whole genome sequencing

WGS is one of the state-of-the-art techniques in molecular biology[46,47,49]. It is currently used in epidemiological studies of MTBC, as it has greater resolution power and can accurately determine transmission chains and outbreaks in comparison with classical genotyping methods[47,49-52]. WGS-based phylogenetic classification of MTBC strains is less prone to distortion as well[49,50]. Furthermore, WGS data offer information on mutations that confer drug resistance, virulence factors, and pathogenicity, as well as enabling discovery of new drug resistance mutations[47,49,51-53]. It also enables the differentiation between relapse and re-infection[47,49,51,54]. By using WGS, the evolutionary history of MTBC can easily be exposed in addition to the global population structure and distribution[44,46,49,55,56]. However, one major limitation to using WGS-based genotyping is the inherent difficulty of data standardization and integration into a readily accessible and expandable classification scheme[47,49]. WGS is expected to become a new standard for routine typing of MTBC strains in the near future, including in low-income, high-burden TB settings[46,47].

1.9 Population structure and genetic diversity of MTBC

On the basis of the presence or absence of the MTBC-specific deletion (TbD1) genomic region, clinical strains of MTBC are classified into ancient or modern phylogenetic clades[3,4,6,57]. Clinical strains of the ancient clade conserve the TbD1 genomic region; however, the TbD1 genomic region is deleted in modern strains[3,4,6,57]. Each of the two clades contain three lineages (L). The modern clade contains L2 (East-Asian), L3 (East-African-Indian) and L4 (Euro-American) while the ancient clade contains L1 (Indo-Oceanic), L5 (*M. africanum* I), and L6 (*M. africanum* II) (Figure 1.5a)[8,57]. Additionally, a novel phylogenetic lineage of MTBC, that appears to be intermediate between the ancient and the modern clades, has been described in Ethiopia and referred to as L7 (Ethiopian lineage)[4,57].

Strain collections from different parts of the world indicate that MTBC involves separate and geographically diverse lineages[4,21,58]. Strains of L2 and L4 have a worldwide distribution, with L2 strains being more prevalent in Asia and Europe while L4 strains are mostly found in Europe and

Africa[55,56,58,59]. Strains of L3 are primarily found in East Africa and in West, Central, and South Asia [3,4,55,56,58-60]. Strains of L1 circulate around the Indian Ocean and the Philippines, whereas strains of L5 and 6 are limited in their geographical distribution to West Africa and strains of L7 are restricted to Ethiopia (Figure 1.5b)[3,4,55,56,58-60].

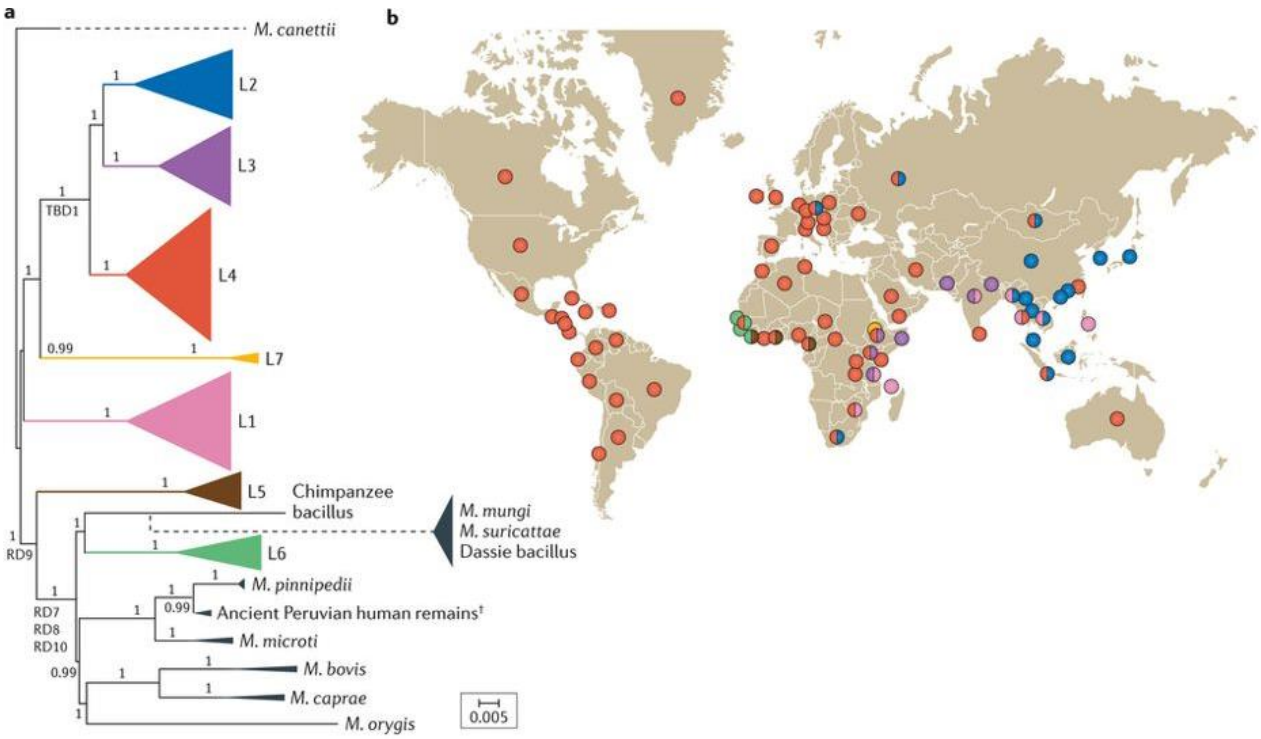


Figure 1.5: The global population structure and geographical distribution of the MTBC lineages. (a) Genome-based phylogeny showing the human- and animal-adapted lineages in colors and grey, respectively. The MTBC specific deletion 1 (TbD1) is indicated for the modern clade (L2-4), likewise, the deletion of the region of difference 7 (RD7), RD8, RD9 and RD10 are shown for the animal-adapted MTBC bacteria. **(b)** The global distribution of the seven human-adapted MTBC lineages. Source: Gagneux[4].

Recently, the MTBC bacteria have been classified into generalist and specialist lineages/sub-lineages based on their geographical distribution or ability to survive in certain environments (i.e. width of the ecological niche)[4,56]. Some lineages/sub-lineages are restricted in their geographical distribution to certain areas, such as L5-7 in West and East Africa, while others, like LAM and Haarlem sub-lineages of L4, are distributed worldwide[4,56].

1.10 Epidemiology of TB in animals and humans in the Sudan

TB in ruminants is a zoonotic disease and has a substantial economic importance as it can cause financial losses of up to $3 billion per year worldwide[1,61,62]. These losses are consequential losses of decreased milk, meat, and hide production and condemnations in slaughterhouses and

abattoirs[1,63,64]. Control and eradication of TB in ruminants are achieved by application of the test-and-slaughter policy and/or abattoir-based surveillance[1,27,61]. Despite all intensive control and management efforts directed against animal TB, the disease has not been eradicated in many countries and continues to be a real public health problem with global perspective[1,27,61]. Wildlife reservoirs, limited sensitivity and specificity of diagnostic tests, large herd sizes, free animal movements, and trade are the major constraints against eradication of TB in ruminants[1,27,61].

Over 80% of cattle and human populations in Africa live where bovine TB is poorly controlled[63,65]. Hence, zoonotic transmission of TB is thought to be common in some parts of Africa[63,65]. In the Sudan, TB was reported in ruminants for the first time during the early 1900s[63,66,67]. Many years later, in the 1950s, it was detected again, but in wild animals and game birds in a zoo in central Sudan[63,66,67]. Recently, a few observational studies have indicated that TB prevalence in cattle is low[63,66,67].

In humans, TB is a major cause of morbidity and mortality in the Sudan[68]. In 2017, the estimated incidence of TB was 77 per 100,000 population with a mortality of 12 per 100,000 population[14]. The total number of notified cases was 21,054, of which 74% were pulmonary TB cases and 49% were bacteriologically confirmed[14]. Overall, TB treatment coverage was 66% and treatment success rate was up to 78%[14]. WHO reported that the prevalence of MDR/RR TB was 2.9% and 13% in new and retreatment cases, respectively, in the Sudan[14]. However, former epidemiological studies found that 6%–30% of the investigated cases are due to MDR MTBC strains[69-75].

For treatment of TB in the Sudan, the national TB control program (NTCP) policy was administration of INH, RMP, PZA, and SM daily for two to three months until the patient becomes smear-negative, followed by eight months of INH and EMB[70,76]. However, some patients were put on a 12-month regimen excluding RMP[70,76]. Thioacetazone (TCN) was previously used instead of EMB[70,76]. Smear-negative pulmonary TB patients and non-severe extra-pulmonary cases were given INH and EMB daily for 12 months, supplemented by daily SM injections during the initial phase. CFX, OFX, CLS, ETH, and AMK are used for treatment of MDR TB[70,76]. Regarding molecular epidemiology of MTBC, only very few studies have investigated the genetic diversity of MTBC strains using spoligotyping[70,77-79]. However, genotyping based on spoligotyping is less discriminative in comparison to MIRU-VNTR and WGS and frequently shows events of homoplasy, i.e. identical typing schemes in phylogenetically unrelated strains leading to conflicting phylogenetic relationships[80].

1.11 Rationale

In Eastern Sudan, knowledge about the prevalence of TB in ruminants, including cattle, sheep, goats and camels is fragmentary and its epidemiology is not very well understood. It is also unknown whether these ruminants are an important source of TB in humans. This gap of knowledge is due to the lack of systemic surveys and monitoring programs, inadequate animal health infrastructures, and a traditional extensive grazing system for raising animals [63,66].

Pulmonary TB in humans is endemic in Eastern Sudan, with high incidence rates that reached up to 275 per 100,000 population in the past few years[68,81]. Nevertheless, this area remains a blind spot with regard to drug resistance rates, circulating MTBC genotypes, and current transmission dynamics, including ruminants as potential reservoirs for human TB infections.

In Eastern Sudan, TB diagnosis depends on clinical examination (i.e. symptoms), chest X-ray, and smear microscopy[82]. No laboratory tests such Xpert MTB/RIF (Cepheid) or polymerase chain reaction (PCR) are available for routine diagnosis in TB centers[82]. Moreover, culture and DST are only possible at the NTRL in the capital Khartoum. WHO[83] showed that smear microscopy has variable sensitivity and specificity. Therefore, it is highly likely that NTCPs that do not use culture or molecular tests often underestimate TB rates. Furthermore, since DSTs are inaccessible in Eastern Sudan, drug resistance TB, particularly MDR TB and XDR TB, are not detected and thus their burden and dynamics are unknown. Besides, mutations that mediate drug resistance in MTBC strains are not investigated and no detailed data, e.g. using WGS, are available. The NTCP policy for TB treatment might lead to further selection of resistances, high failure, and uncontrolled transmission of MDR TB and XDR TB. On top of that, molecular epidemiological studies are very rare in the Sudan. Only very few studies have investigated the genetic diversity of MTBC strains using spoligotyping[70,77-79].

Strains of MTBC L3 are widely prevalent in the Sudan[70]. Additionally, the number and the variety of the detected spoligotypes suggest that L3 strains have probably been circulating in the country for a long time[70]. In order to understand TB molecular epidemiology in Sudan, one needs to understand the population structure and history of L3 strains within a wider geographical setting. Additionally, nothing is actually known regarding the pathogenicity, virulence, and DST among L3 strains or whether L3 strains are associated with chains of MDR TB and XDR TB transmission and outbreaks/epidemics. From a genetic point of view, strains of L3 have the characteristic CAS-specific RD750 deletion and belong to one of the following sub-lineages or spoligotypes: CAS1-Delhi (ST26), CAS1-Kilimanjaro or CAS1-Kili (ST21), CAS1-variant (ST25), and CAS2 (ST288)[57,59]. Nevertheless, a robust classification of L3 sub-lineages based on 24-loci MIRU-

VNTR is not available while genomic markers to capture the whole genetic diversity based on WGS are missing.

1.12 Objectives

Although TB is endemic in Eastern Sudan as Abdallah and Ali[68] and WHO[81] reported, the genetic diversity and transmission dynamics of MTBC bacteria, including MDR strains and mutations that confer drug resistances, are only poorly investigated. Another intriguing question is whether TB in ruminants and zoonotic transmission contributes significantly to the TB epidemiology in humans.

By using state-of-the-art molecular genotyping techniques, such as 24-loci MIRU-VNTR and WGS, my PhD thesis had the following main goals:

1. To estimate the prevalence of tubercle granulomatous lesions and identify the infecting MTBC strains in ruminants slaughtered at the two slaughterhouses of Kassala
2. To investigate the positive predictive value (PPV) of ZN smear microscopy and to identify the genotypes and DST patterns of MTBC strains causing pulmonary TB in humans in Eastern Sudan
3. To investigate the global population structure of MTBC L3 strains, their geographic distribution, and possible association with MDR TB

2 Materials and methods

2.1 Study area and setting

This study was conducted in Eastern Sudan; a region that shares international borders with Ethiopia, Eritrea, and Egypt and national borders with four states in the Sudan (Figure 2.1). Kassala was selected to investigate TB in animals. While for investigating TB in humans, the public hospitals of Kassala, Port Sudan, and El-Gadarif were surveyed, since TB is endemic in humans in Eastern Sudan[68]. Port Sudan is located at the Red Sea coast; latitude $19^0$37′N and longitude $37^0$13′E at an elevation of 10 meters above sea level. Nearly, 579,942 people live Port Sudan. It is around 675 km from the capital Khartoum. Port Sudan has a petroleum refinery and docking facilities that handle the bulk of the country's external trade. Kassala is 575 km to the southwest Port Sudan at latitude $15^0$27′N and longitude $36^0$24′E at 507 meters elevation above sea level. A population of approximately 500,000 people reside in Kassala. The seasonal Gash River divides Kassala into eastern and western parts. Kassala is famous with many agricultural activities and with its big market of fruits. Besides, it is an important trade center. The distance from Kassala to the Eritrean border is about 30 Km. El-Gadarif is about 200 km southwest Kassala; latitude $14^0$02′N and longitude $35^0$23′E at an elevation of 608 meters above sea level. It is a commercial center for the cotton, cereals, sesame seeds, and fodder produced in the surrounding area. The Gash Irrigation Project is located to the northeast of El-Gadarif. Light industries in El-Gadarif include cotton ginning and spinning mills and soap factories. El-Gadarif is nearly 160 Km away from the Ethiopian border. People of Eastern Sudan are mainly Beja, Arabs, Nubians, West Africans, and small minorities of Asians and Europeans[84].

In Eastern Sudan, up to 1,500 human TB suspected cases (i.e. individuals presenting with symptoms suggestive of TB) are recorded annually in each of the three surveyed public hospitals[68,85,86]. Roughly, half of these cases are smears positive[68,85,86]. However, HIV is not a big health problem in Eastern Sudan as it is detected in only ≤0.5% of all tested individuals despite the fact that amongst TB patients HIV prevalence is 18.3%[85,87]. Overall, 17% of all registered TB cases in the Sudan have known HIV status[12]. Eastern Sudan suffered from civil war for more than 10 years and it ended in 2006[82,84].

Eastern Sudan falls within the Sahelian climate zone of Africa[88]. Temperature ranges from a mean minimum of 17^0C in January to a mean maximum of 47°C in April and May[88]. Annual rainfall is concentrated in a single relatively short autumn season from June to September and amounts to around 680 mm per annum. Soils are dark, heavy, deep cracking vertisol[88].

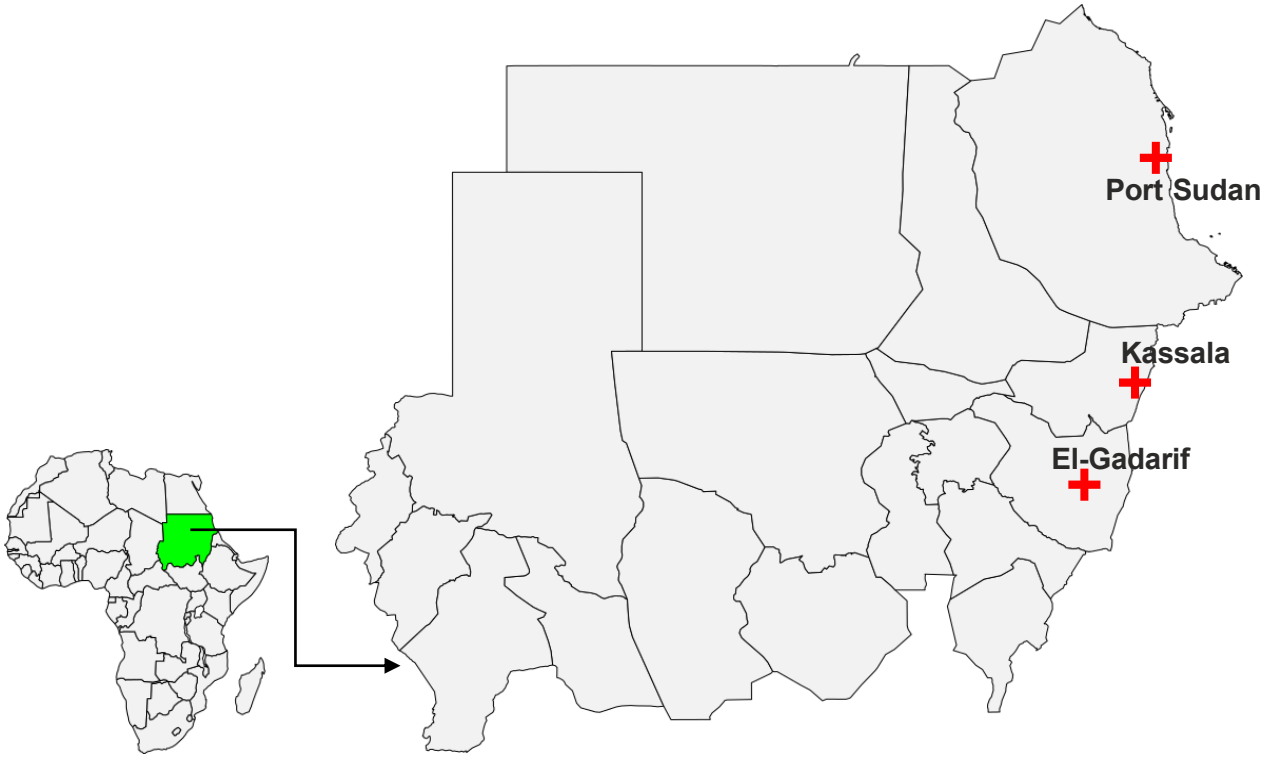


Figure 2.1: Map of the Sudan showing the three sampling sites. The three sampling sites are indicated by Red Cross.

2.2 TB in carcasses of ruminants

2.2.1 Study design, sample size calculation, and sampling

A cross-sectional study design was used in this survey for a period of six months (between June and November 2014)[89]. Kassala was purposively selected since it is an important animal production area in the Sudan. Around, 5 million heads of animals are scattered in the pasture of Kassala state, of which 835,416 heads are cattle, 2,000,334 heads are sheep, 1,661,568 heads are goats, and 659,614 heads are camels[90]. A considerable number of these animals are sold every year in local markets for domestic consumption and for export to international markets[91,92]. Kassala has two slaughterhouses which are East and West Gaash slaughterhouses. The two slaughterhouses were conveniently included in the survey as about 7,500 head of animals, including 2,500 cattle, 4,000 sheep, 950 goats, and 50 camels, are slaughtered and processed for meat per month in these slaughterhouses[89,93].

The actual sample size (n) was calculated based on the following parameters: 95% level of confidence (CI), ±5% desired level of precision, and the expected prevalence of granulomatous lesions suggestive of TB in carcasses of slaughtered animals[89]. The prevalence of TB suspected lesions in carcasses of slaughtered animals in Kassala has not been investigated previously and

it is currently unknown. Therefore, n was determined by assuming an expected prevalence of 50% and by using the formula:

$n = (1.96)^2 \times P_{exp} \times (1- P_{exp})/d^2$

Where:

n	= required sample size
$(1.96)^2$	= constant
P_{exp}	= expected prevalence rate, 50%
d	= desired absolute precision, ±5%

The calculated n was 384 animals from East Gaash slaughterhouse and the same number from West Gaash slaughterhouse. The calculated n was multiplied by 3 to minimize bias and to account for the effect of randomness and representativeness[89]. The final calculated n was 1,152 cattle, sheep, goats, and camels from each of the two slaughterhouses[93].

For selection of study animals, a by-animal species convenient stratification was employed in a first step[89]. In a second step, individual animals were selected by systematic random sampling. Since each slaughterhouse was visited 72 times (three visits per week), 16 animals (four of each cattle, sheep, goats, and camels) were selected on each study day in each slaughterhouse[89]. To determine the sampling interval, the total number of each animal species to be slaughtered on that specific day was divided by four and every N^{th} animal was selected after random selection of the first animal. Consequently, 710 cattle, 729 sheep, 700 goats, and 165 camels were included in the study. For camels, normally less than four were always slaughtered per day in each slaughterhouse. To compensate for this, the required n per day was achieved by selecting more cattle or sheep or goats. The selected animals were identified using a permanent marker, kept separately until slaughtering time and released for slaughter one after another[93].

2.2.2 Post-mortem examination

Post-mortem inspection was done visually, by palpation, and by *in situ* slicing of certain lymph nodes (LN), organs, and muscles[94,95]. In the head, masseter muscle, tongue, and left and right mandibular LNs, parotid LNs, retropharyngeal LNs, and atlantal LNs were inspected[94,95]. In the thorax, anterior and posterior mediastinal LNs and left, right, cranial, and medial bronchial LNs, and the lungs were inspected[94,95]. In the abdomen, the mesenteric and hepatic LNs were examined as well as the liver, small intestine, and the kidneys[94,95]. In the carcass, the left and the right prescapular LNs, prefemoral LNs, superficial inguinal (supra mammary or scrotal) LNs, udder or scrotal contents, and joints were examined in detail under a bright-light source[94,95]. The lobes of the left and the right lungs were inspected and palpated. Then, each lobe was sectioned into

slices of about 2-cm-thick using sterile surgical blades to facilitate the detection of lesions[94,95]. Similarly, LNs and other organs were palpated and sliced into thin sections and inspected for presence of visible lesions[94,95]. Whenever gross lesions suggestive of TB were detected, the infected tissue/organ was classified as having TB lesions[93].

2.3 TB in humans in Eastern Sudan

2.3.1 Ethical considerations and consent

The study was approved by the National Research Ethics Committee, Federal Ministry of Health, Khartoum, the Sudan, and by the Institutional Review Board of the Faculty of Medicine, University of Khartoum, Khartoum, the Sudan (No. 85-03-09). For suspected TB patients to participate in the study, a written informed consent for was obtained from them or their respective guardians in case of children and illiterate individuals.

2.3.2 Study population and sample collection

In a cross-sectional study for the periods from June to November 2014 and from January to July 2016, TB suspects who visited the outpatient department of Kassala, Port Sudan, and El-Gadarif teaching hospitals in search of medical help and produced smear-positive sputum were included in this study. Each recruited patient provided two sputum samples at separate time points, including a spot sample and an early morning sample. If at least one of the two samples showed visible AFB under the microscope (i.e. smear-positive), the samples were pooled and kept frozen at -20ºC. The maximum period of samples storage was six months. An amount of ≤2 ml of each pooled samples was transferred to a screw-capped Eppendorf tube shortly before shipping to the National Reference Laboratory (NRL) for Mycobacteria in Borstel, Germany. The samples were sent to the NRL in two separate batches. Samples collected between June and October 2014 arrived in the NRL in three days whereas the samples collected between January and July 2016 arrived in the NRL after three weeks. All samples were shipped at room temperature.

2.4 Materials

All materials used in laboratory analyses in this study are listed in Annexes 1 to 8.

2.5 Methods

2.5.1 Sample decontamination

The collected granulomatous TB-suggestive tissue samples from carcasses of slaughtered animals and collected sputum samples from humans were decontaminated by sodium hydroxide and the mucolytic agent N-Acetyl-L-Cysteine (NALC-NaOH) according to the procedure described by the guidelines of DIN[96] and Siddiqi and Rüsch-Gerdes[97]. The granulomatous tissue samples

collected from each of the two carcasses were first pooled, minced, and homogenized. Then, an equal volume of NALC-NaOH was added to each sputum sample and to the homogenized tissue. The mixture was incubated for 20 minutes on a shaker, then a volume of 10 to 20 ml sterile phosphate buffer (PBS) (0.067 mol/l, pH 6.8) and 1 to 2 drops of a 5.0% Tween 80 were added. The tubes were put in a centrifuge for 20 minutes (3,500×g). Finally, the supernatant was poured off and the pellet was resuspended in 1.0 ml sterile PBS.

2.5.2 Smear microscopy

Smears were stained by Kinyoun and auramine O stains as described by GLI[98]. A volume of 0.25 ml of the resuspended granulomatous tissue and sputum pellets were used for smear microscopy. It was dropped on a clean slide, air dried, and heated for fixation. Samples collected in 2014 were stained using the Kinyoun method in conjunction with an automated staining system (ZN Aerospray® TB Slide Stainer/Cytocentrifuge, Wescor, Logan, UT, USA) and read using light microscopy (oil immersion lens, 100x magnification). Samples collected in 2016 were stained using auramine O staining and were read with a LED microscope (40x). For auramine O staining, heat-fixed slides were covered with the fluorochrome stain for 15 min, then rinsed in water, covered with 0.5% acid-alcohol for 30-60 seconds for decolorization, and rinsed in water again. For counterstaining, the slides were covered with potassium permanganate for 2 minutes, washed with water, and air dried. Results were recorded as smear-positive or smear-negative according to DIN[96] and WHO[99].

2.5.3 Culturing

The collected sputum samples were cultured into liquid and onto solid media[96-98,100]. For culturing into liquid media, a volume of 0.5 ml of the resuspended samples' pellets were inoculated into mycobacterial growth indicator tube containing 7H9 liquid media (MGIT; Becton-Dickinson) primed with growth supplement (bovine albumin, dextrose, catalase, oleic acid, and polyoxyethyline state) (OADC™; Becton-Dickinson) and antibiotics (polymyxin B, amphotericin B, nalidixic acid, trimethoprim, and azlocillin) (PANTA™; Becton-Dickinson). Prior inoculation, the MGIT PANTA was reconstituted with 15 ml of the MGIT OADC and 0.8 ml of the liquid antibiotic supplement was added to each tube. The tube was tightly recapped, mixed well, and incubated in the BACTEC MGIT 960 instrument[97]. For culturing onto solid media, 0.25 ml of the resuspended pellets were inoculated onto each of Löwenstein-Jensen (LJ) and Stonebrink slopes (Artelt-Enclit, Oelzschau, Leipzig, Germany). The solid media were supplemented with antibiotics (Artelt-Enclit, Oelzschau, Leipzig, Germany). The MGIT tubes were incubated at 37ºC for a maximum of 42 days while the LJ and Stonebrink slopes were incubated for 56 days. Positive cultures were examined

microscopically for AFB with Kinyoun stain and by culturing on blood agar for contamination. Positive MGIT tubes that demonstrated the presence of contaminants only were not processed further. Culture results were reported as positive for mycobacteria, negative if no growth occurred in at least one of the liquid and solid media, and contaminated if all three cultures were contaminated.

2.5.4 DNA extraction

For qPCR, DNA was extracted using QIAamp® DNA Mini kit (250) (Qiagen GmbH, Hilden, Germany). The remaining of the decontaminated suspension of each sample was transferred to a 2 ml Eppendorf tube and centrifuged at 15,000×g for 5 minutes. The supernatant was discarded and 180 µl tissue lysis buffer (ATL) together with a volume 20 µl proteinase K were added to the tube and vortexed for 10 seconds. Then, 200 µl lysis buffer (AL) were added to the tube which was incubated in a boiling water bath for 20 minutes and thereafter sonicated for 15 minutes in order to enhance adequate lysis of bacterial cells. This was followed by adding 200 µl ethanol to the sample to enable DNA precipitation. The mixture was transferred to a filter column and centrifuged at 6,000×g for 1 minute, followed with washing steps by washing buffer AW1 and AW2. A volume of 500 µl of AW1 was added on the filter and centrifuged at 6,000×g for 1 minute. Then, 500 µl of AW2 were added on the filter and centrifuged at 16,000×g for 2 minutes. Finally, the DNA was eluted from the column by using 80 µl AE buffer which was added on the filter, incubated for 5 minutes at room temperature and centrifuged at 6,000×g for 1 minute. The finished sample was transferred to a new Eppendorf tube and stored at -20 until used.

For conducting LPA (GenoType *Mycobacterium* CM) and LPA (GenoType *Mycobacterium* MTBC), 1 ml of each positive MGIT culture was centrifuged at 10,000×g for 15 minutes at room temperature, the supernatant was discarded, and the pellet was suspended in 500 µl distilled water. In case of solid media, a loopful with colonies were transferred into a tube containing 500 µl distilled water. The suspended pellet was boiled for 20 minutes and thereafter sonicated for 15 minutes. Extracted DNA were stored at -20°C before further use.

For WGS, DNA was extracted from Stonebrink cultures using the *N*-acetyl-*N,N,N*-trimethyl ammonium bromide (CTAB; final concentration, 40 mM) and NaCl (final concentration, 0.1 M) (CTAB-NaCl) method according to Somerville et al.[101]. Colonies of grown MTBC strains were suspended in 400 µl of Tris-EDTA (TE) buffer, pH 8.0, which was aliquoted into 1.5 ml screw cap tubes. The viable MTBC bacteria were killed by incubation in a water bath at 80ºC for 20 minutes. After that, 50 µl lysozyme (10 mg/ml) was added to the tube, gently mixed with a pipette and incubated at 37ºC overnight. Volumes of 70 µl of 10% SDS and 5 µl proteinase K (10 mg/ml) were

added to the mixture, gently mixed, and incubated at 65ºC for 10 minutes. The CTAB/NaCl was incubated together with the samples for warming. After that, 100 µl of 5M NaCl, 100 µl CTAB/NaCl were added, gently mixed, and incubated at 65ºC for 10 minutes, followed by adding 750 µl of chloroform/isoamylacohol, mixing by inversion, and centrifugation at 10,000×g for 5 minutes. The aqueous supernatant was added into fresh Eppendorf tubes containing 450 µl ice-cold isopropanol and mixed by inversion. DNA was precipitated at -20°C for at least 30 minutes followed by centrifugation at 10,000×g for 15 minutes at room temperature. The supernatant was removed and the pellet was washed with 1 ml ice-cold 70% ethanol. Much of the ethanol was removed and discarded and the tubes were tilted to allow the pellet to air dry. The pellet was rehydrated with a volume of 80 µl TE.

2.5.5 qPCR

An in-house qPCR detecting MTBC and nontuberculous mycobacteria (NTM) strains was conducted using the Rotor-Gene 2000 (Corbett Research, Mortlake, Australia) and ABI TaqMan Universal PCR Master Mix as previously described by Hillemann et al.[102]. Briefly, the experiments of the qPCR were done under reaction conditions of 95°C for 10 minutes and 40 to 50 two-step cycles consisting of 92°C for 15 seconds and 60°C for 1 minute. The primer concentrations in the reaction mixtures of 15 µl were 500 nM for each forward and reverse primer and 200 nM for each probe. A total of 1.5 µl of a 1:100 dilution of chromosomal DNA was used as a PCR template.

2.5.6 Line probe assay

LPA GenoType *Mycobacterium* CM (common mycobacteria) and LPA GenoType *Mycobacterium* MTBC were used to classify the isolated *Mycobacterium* strains into NTM or MTBC and to identify the species of the MTBC strains, respectively. The two LPAs were conducted according to the instruction of the manufacturer (Hain Lifescience GmbH, Nehren, Germany). With the exception of the use of different primers, the procedure of the two LPAs is similar. Briefly, the test was performed by using 35 µl of a primer-nucleotide mixture (provided with the kit), amplification buffer containing 2.5 mM MgCl2, 1.25 U of hot start *Taq* polymerase (Qiagen, Hilden, Germany), and 5 µl of the heat-inactivated suspension in a final volume of 50 µl for PCR. The amplification protocol consisted of 15 minutes of DNA denaturing at 95°C, followed by 10 cycles comprising 30 seconds at 95°C and 120 seconds at 58°C, an additional 20 cycles comprising 25 seconds at 95°C, 40 seconds at 53°C, and 40 seconds at 70°C, and a final extension at 70°C for 8 minutes. Hybridization and detection were performed in an automated washing and shaking device (Profiblot, Tekan, Maennedorf, Switzerland). The program was started after mixing 20 µl of the amplification products with 20 µl of denaturing reagent (provided with the kit) for 5 minutes in

separate wells of a plastic trough and 1 ml of pre-warmed hybridization buffer was added, followed by a stop to put the membrane strips into each well. The hybridization procedure was performed at 45°C for 30 minutes, followed by two washing steps. For colorimetric detection of the hybridized amplicons, streptavidin conjugated with alkaline phosphatase and substrate buffer was added. After the final washing, strips were air dried and fixed on paper.

2.5.7 Sanger sequencing

For identification of the species of the isolated NTMs, sequencing of the 16S rRNA and/or internal transcribed spacer (ITS) DNA fragments was performed as per the standard procedure described by Richter et al.[103] and Richter et al.[104]. The complete PCR products were sequenced on an automated DNA sequencer (ABI 377 and 3500xl; Applied Biosystems) by cycle sequencing using the Big Dye RR Terminator cycle sequencing kit (Applied Biosystems). The resulting sequences were aligned and compared to the sequences of the International Nucleotide Sequence Database Collaboration.

2.5.8 Drug susceptibility testing

The susceptibility of all of the MTBC strains to SM 1 µg/ml, INH 0.1 µg/ml, RMP 1 µg/ml, EMB 5 µg/ml, and PZA 100 µg/ml was tested using the BACTEC MGIT 960 system as described by Siddiqi and Rüsch-Gerdes[97]. To conduct the test, five MGIT cultures and along with a growth control (GC) tube were prepared by adding 0.8 ml of BACTEC 960 SIRE supplement aseptically. The recommended concentration of each drug was aseptically added to each corresponding MGIT tube as well. Then, 0.5 ml of the well-mixed culture suspension (inoculum) was added into each of the drug containing MGIT tubes. For the GC tube, the suspension (inoculum) was diluted to 1:100 using a sterile saline then a volume of 0.5 ml was added. The the inoculated broth was mixed very well by inverting the MGIT tubes 5–6 times. The ready MGIT tubes were finally placed in the BACTEC MGIT 960 instrument. The results were automatically interpreted by the BACTEC MGIT 960 instrument and reported as either susceptible or resistant. For PZA testing, the BACTEC MGIT 960 PZA kit was used and the concentration of the drug was 100 µg/ml. Prior to inoculating the PZA set tubes, 0.8 ml of BACTEC MGIT 960 PZA supplement was added to both GC and PZA tubes and 100 µl of PZA drug solution was added to the PZA tube. Inoculum suspension was prepared and added to the GC and PZA tubes. Using predefined algorithms, readings were automatically interpreted by the BACTEC MGIT 960 instrument and reported as either susceptible or resistant. All strains which were resistant to at least one first line drug were subjected to OFX 2 µg/ml, AMK 1 µg/ml, and CPM 2.5 µg/ml susceptibility testing according to Rüsch-Gerdes et al.[105].

The ‘unloaded drug set report’ lists the growth units, time to result and susceptible, resistant or invalid result. To check the purity of the innocula used for DST, the suspensions (grown in MGIT 7H9 medium) were plated on sheep blood agar plates and checked after 48 hours for signs of growth. Any growth on the plate indicated that the suspension was contaminated[97].

Strains with a mutation in the *embB* codons 306, 406 or 497, but with a susceptible EMB DST results were further evaluated by MIC determination using the following concentrations in the BACTEC MGIT 960 system: 1.25, 2.5, 3.75, and 5.0 µg/ml. Growth after 10 Days at 5.0 µg/ml were considered as EMB resistance[97,106].

2.5.9 Spoligotyping

The spoligotyping was performed according to the spoligotype kit supplier’s instructions (Ocimum Biosolutions Company, Ijsselstein, the Netherlands). The DR region was amplified by PCR using oligonucleotide primers (DRa: 5’ GGT TTT GGG TCT GAC GAC 3’ and DRb: 5’ CCGAGA GGG GAC GGA AAC 3’) derived from the DR sequence. The DRa is biotinylated at the 5’-end. A total volume of 25 µl of the following reaction mixture was used for the PCR: 12.5 µl of HotStarTaq Master Mix (Qiagen: this solution provides a final concentration of 1.5 mM MgCl2 and 200 µM of each deoxynucleotides triphosphates), 2 µl of each primer (20 pmol each), 5 µl suspension of heat-killed cells (approximately 10 to 50 ng), and 3.5 µl distilled water. The mixture was heated for 15 min at 96°C and subjected to 30 cycles of 1 minute at 96°C, 1 minute at 55°C, and 30 seconds at 72°C. The amplified product was hybridized to a set of 43 immobilized oligonucleotides, each corresponding to one of the unique spacer DNA sequences within the DR locus. After hybridization, the membrane was washed twice for 10 min in 2× SSPE (1× SSPE is 0.18 M NaCl, 10 mM NaH2PO4, and 1 mM EDTA [pH 7.7])/0.5% sodium dodecyl sulphate at 60°C and then incubated in 1:4000 diluted streptavidin peroxidase (Boehringer) for 45 to 60 minutes at 42°C. The hybridized DNA was detected by an enhanced chemiluminescence method (Amersham) and by exposure to an X-ray film (Hyperfilm ECL, Amersham).

2.5.10 24-loci MIRU-VNTR typing

The 24-loci MIRU-VNTR typing was conducted according to the standardized protocol[107]. Briefly, MIRU-VNTR alleles were amplified using Quadruplex PCR Kit (Genoscreen, Lille, France). Fragment analysis using the GeneScan™ 1200 LIZ dye as a size standard (Life Technologies, Darmstadt, Germany) was carried out on a capillary sequencer 3130xL and 3500xL for the genetic analyzer. The GeneMapper software version 3.7 (Life Technologies, Darmstadt, Germany) was used to determine the copy number of MIRU-VNTR alleles.

2.5.11 Whole genome sequencing

The WGS was done according to the instructions of the manufacturer using Nextera (XT) kit (Illumina, San Diego, CA, USA) for library preparation. First, the quality and quantity of the isolated genomic DNA was measured and determined by Qubit™ 2.0 (Life Technologies, Darmstadt, Germany). It was next standardized and diluted using MilliQ water to a final concentration of 0.2 ng/µl from which a volume of 5 µl was used to create the library. Tagmentation buffer and sequence adaptors were added for fragmentation of the genomic DNA. Amplification of the labeled DNA and addition of further adaptor sequences and sample barcodes to the tagmented DNA fragments was done and the amplicon was purified using AMPure XP beads and ethanol. The purified Nextera (XT) library was transferred into a new tube and pooled. The length of the pooled DNA fragments was determined by the 2100 bioanalyser and the High Sensitivity kit from Agilent technologies (Santa Clara, USA). The pooled DNA libraries were loaded into the cartridges of the MiSeq Reagent kit and a run was started. The strain was sequenced with a minimum coverage of 50-fold. After completion of the run, the sequence data were output to FastaQ.

Analysis of the WGS data was conducted by using MTBseq pipeline as described by Kohl et al.[52]. Briefly, the pipeline workflow is comprised of 10 steps for sample specific and comparative analyses (Figure 2.2). During the first seven steps, sample specific analysis, including mapping, base call recalibration and realignment of reads around insertions or deletions, and variant calling are executed using algorithms and tools such as BWA-MEM algorithm and SAMtools. Moreover, all of the detected SNPs and the subsequent amino acid substitutions and association to antibiotic resistances besides to insertions and deletions are annotated with respective metadata in the sample specific report file. Before generating of a descriptive statistics report, the overall dataset performance is examined in the TBvariants step. The succeeding last step of the sample specific analysis permits the phylogenetic classification of the input sample(s). Finally, the comparative analysis of multiple samples is executed during the three last steps (i.e.TBjoin, TBamend, and TBgroups). The output is a list of all positions for which a variant was detected in any of the input samples and repetitive regions, resistance associated genes, and the presence of other variants within a window of 12 bp within the same dataset are all filtered. The final step of the comparative analysis groups the input samples according to the number of distinct SNP positions. In addition, Fasta formatted sequences are generated as direct input for targeted applications (e.g. tree reconstruction algorithms). The pipeline uses *M. tuberculosis* H37Rv genome (NC_000962.3) and corresponding metadata as reference by default[52].

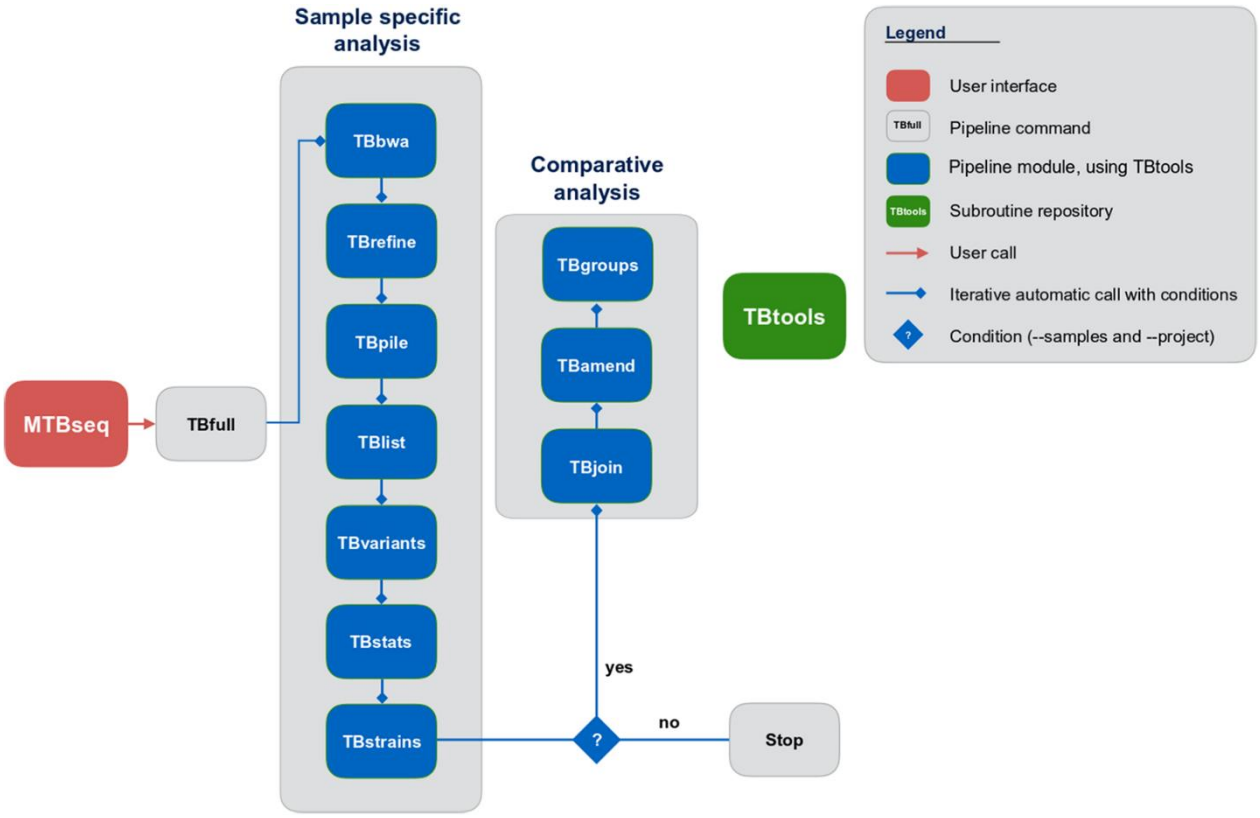


Figure 2.2: A simplified illustration of the steps and the workflow of the MTBseq pipeline. Source: Kohl et al.[52]

2.5.12 Data analyses and visualization

All collected data were entered, cleared, coded, and stored electronically in a Microsoft Excel for Windows 2010 database. The Statistical Package for Social Sciences (SPSS) for Windows version 20.0 (SPSS Inc., Chicago, Illinois, USA) was used for all appropriate statistical analyses including frequencies and proportions. A logistic regression model was used to assess differences of proportions of clustered strains, MDR strains, and clustered MDR strains between the different DCMGs.

Molecular typing data were analyzed using BioNumerics version 7.5 software (Applied Maths, St. Martens, Belgium) according to the manufacturer's instructions.

For prediction of the complete sequence of 16S gene of the *Mycobacterium* strain isolated from sheep carcasses, de novo assembling of the fully sequenced genome was carried out. Reads were trimmed with Trimmomatic version 0.35 (PMCID: PMC4103590) for Q30 in a sliding window of 20 bp and nextera adapter contamination, discarding all reads shorter than 100 bp[108]. Trimmed paired and unpaired reads were assembled with SPAdes (PMCID: PMC3342519) version 3.6.2 using the built in read error correction based on k-mer frequencies and the option –careful[109]. The

16S gene sequences were predicted by the RNAmmer Prediction Server from the assembled contigs which were filtered for containing ribosomal sequences[110].

3 Results

3.1 Slaughterhouse investigation for TB

Generally, information about the prevalence and genetic diversity of MTBC strains in animals is scanty in Africa[65]. Furthermore, the contribution of animal TB to human TB incidence is unknown[65]. In the Sudan, a few studies have shown that MTBC are not very common in cattle carcasses in the central and western parts of the country[63,66,67]. Nevertheless, in the eastern part of the country, nothing is known about the prevalence of MTBC strains in ruminants' carcasses. Therefore, this study was conducted to investigate the prevalence of TB-suggestive lesions and to identify the infecting MTBC strains in ruminants slaughtered in Kassala.

3.1.1 Study population

A total of 40,089 ruminants, including cattle, sheep, goats, and camels, were slaughtered during the study period (June - November 2014) at the two surveyed slaughterhouses in Kassala, Eastern Sudan[111]. The majority of these animals (81.0%) were slaughtered at East Gaash slaughterhouse, while the remaining animals (19.0%) were slaughtered at West Gaash slaughterhouse. For cattle, goats, and camels, the male to female ratio was 1:1, while for sheep the ratio was 1:2[111]. Slaughtered cattle, sheep, and goats were at least one year old, but slaughtered camels were at least five years old. Nevertheless, younger animals can be slaughtered, but only to relieve pain and suffering (e.g. due to accidents)[111]. Most of the slaughtered animals were indigenous breeds in addition to a few exogenous breeds of Eritrean and Ethiopian origin[111]. Animals that are raised for the purpose of meat production in Kassala are kept in the pasture (e.g. Albutanah plains) where grass and water are abundant during rainfall seasons. However, in dry seasons, animals are trekked for long distances in search of grass and water and only a few of them are kept at home or in close proximity to the villages and are provided with feed and water[112-114]. Herd/flock size (i.e. number of animals in the Herd/flock) is sustained by natural breeding. Neither intensive or feedlot farming nor artificial insemination (AI) are practised for meat production in Kassala[112-114]. Herds/flocks are mainly composed of females together with a few males, as most of the produced male animals are taken to export markets at a specific age after weaning, while females are kept for breeding and are sold for domestic consumption at the end of their production lives[112,113]. Of all animals slaughtered during the study period (n=40,089), 5.7% (n=2,304) were selected during antemortem (i.e. before slaughtering) and their carcasses were thoroughly inspected for presence of TB-suggestive lesions[89,93-95]. Characteristics of the selected animals are presented in Table 3.1.

Table 3.1: Frequencies and distributions of inspected carcasses at the slaughterhouses of Kassala

Variable	No. of inspected carcasses	% of inspected carcasses
Slaughterhouse		
East Gaash	1468	63.7
West Gaash	836	36.3
Species		
Cattle	710	30.8
Sheep	729	31.6
Goat	700	30.4
Camel	165	7.20
Breed		
Gaash (cattle)	270	11.7
Butana (cattle)	232	10.1
Cross (cattle)	81	3.50
Karour (cattle)	70	3.00
Barka (cattle)	57	2.50
Garrage (sheep)	161	7.00
Dubassy (sheep)	193	8.40
Gaash (sheep)	277	12.0
Arrit (sheep)	56	2.40
Horro (sheep)	42	1.80
Baladi (goat)	150	6.50
Garrage (goat)	306	13.3
Nubi (goat)	244	10.6
Bushari (camel)	108	4.70
Anaafy (camel)	57	2.50
Age group (years)		
1 - 3	503	21.8
> 3 - 6	784	34.0
> 6	1017	44.2
Sex		
Male	813	35.3
Female	1491	64.7
Total	**2304**	**100**

3.1.2 Prevalence of TB lesions

The overall prevalence of TB-suggestive lesions was 0.1% (2/2304). By-species prevalence was 0.0% in the carcasses of slaughtered cattle, goats, and camels and 0.3% (2/729) among sheep carcasses. The detected lesions had the typical appearance of caseous necrosis. They were whitish, scattered, and enclosed in lightly congested grey or yellowish fibrous tissue congregating all-together and were observed in the liver and lungs of one carcass and in the peritoneal cavity, covering the whole abdomen and viscera in the second carcass (Figure 3.1).

Figure 3.1: Detected TB-suggestive lesions in the carcasses of slaughtered animals. The lesions were found in the liver and lungs (left) and the peritoneal cavity (right) of two carcasses of sheep in Kassala. Black arrows denote to the TB-suggestive lesions.

3.1.3 Isolation of mycobacteria

A strain of *Mycobacterium* species was isolated from the samples (n=3) collected from one of the two carcasses. However, no *Mycobacterium* species strains could be isolated from the samples (n=4) collected from the second carcass, and cultures were rather extensively contaminated with bacteria and yeasts. Furthermore, MTBC members such as *M. bovis*, *M. caprae*, and *M. tuberculosis* could not be isolated from any of the cultured samples.

3.1.4 Molecular identification

As Figure 3.2 shows, the banding pattern (band 1, 2, 3, and 10) that appeared on the LPA strip of the GenoType *Mycobacterium* CM test was unique. It was different from the banding patterns of all 23 *Mycobacterium* species that can be identified by this LPA[115]. Likewise, the strain could not be identified by blasting its 16S rRNA and ITS sequences at the National Center for Biotechnology Information (NCBI) database. The two DNA sequences (i.e. 16S rRNA and ITS) did not align with any known validly described *Mycobacterium* species.

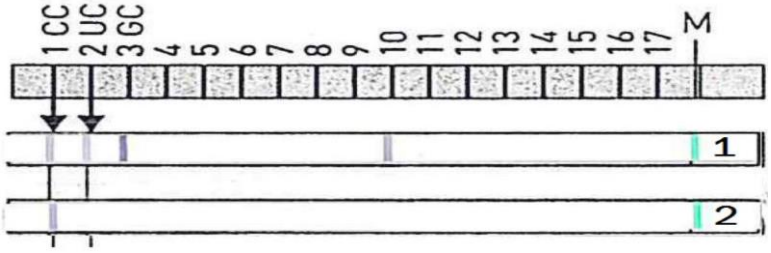


Figure 3.2: Bands of GenoType *Mycobacterium* CM. LPA showing the unique banding pattern of the unknown *Mycobacterium* species isolated from TB-suggestive lesions collected from a carcass of slaughtered sheep in Kassala. **1-** strain and **2-** negative control

WGS showed that the isolated *Mycobacterium* species had 4.8 mbp. Moreover, metagenomic sequence classification by Kraken revealed that 71.6% of the reads were unclassified and the closest relative was the *Mycobacterium sinense* JDM601. Additionally, Kraken showed that only 0.2% and 0.1% of the reads were identical to the genome of *M. tuberculosis* H37Rv and *M. intracellulare* MOTT-64, respectively. The complete predicted sequence of the 16S allocates this strain to slow-growing mycobacteria, with the closest similarity to members of the *M. terrae* group. Figure 3.3 depicts the alignment of the predicted sequence of the 16S in the NCBI Blast database.

Figure 3.3: Alignment of the 16S rRNA gene sequence of the strain (Query) and *Mycobacterium sinense* strain JDM601 (Sbjct).

As mentioned above, the prevalence of TB-suggestive lesions in carcasses of ruminants, including cattle, sheep, goats, and camels, was very low in Kassala. An unknown *Mycobacterium* species was sampled/isolated from the lesions of one carcass but no MTBC strain could be detected.

3.2 TB in humans in Eastern Sudan

The genotyping results of samples collected from the carcasses of ruminants showed no evidence for zoonotic MTBC transmission in Kassala, Eastern Sudan. In the following, the study focused on human-derived specimens to investigate the main causative agent of TB since the disease represents a significant health problem in humans with high incidence rates[68]. In addition, TB centers in the region lack proper diagnostics[82]. Therefore, it is very likely that TB rates, particularly MDR TB and XDR TB rates, are underestimated and the true burden of the disease is likely unknown. Further, mutations that mediate drug resistance in MTBC are not investigated and no detailed data, e.g. using LPAs or WGS, are available. Thus, standard laboratory methods and modern sequencing technologies were jointly used to analyse the MTBC population structure in Eastern Sudan and to determine MDR TB and XDR TB rates.

3.2.1 Study population

Sputum samples (n=383) investigated in this study were collected from three sites in Eastern Sudan, namely Kassala (34.7%, n=133), El-Gadarif (42%, n=161), and Port Sudan (23.3%, n=89) (Table 3.2). TB patients who were recruited were between 4 and 105 years old with a median age of 35 years old (interquartile range of 25; 45). Additionally, most of the patients (64%, n=245) were men. Moreover, retreatment and newly diagnosed TB cases were 5.5% (n=21) and 81.5% (n=312), respectively, while treatment history for 13% (n=50) of the patients was not known (Table 3.2).

Table 3.2: Demographic characteristics of TB patients investigated in Eastern Sudan

Characteristic	No. of recruited patients	% of recruited patients
Origin		
Kassala	133	34.7
El-Gadarif	161	42.0
Port Sudan	89	23.3
Sex		
Male	245	64.0
Female	123	32.1
Not available	15	3.9
Age (years)		
< 25	76	19.8
25 to 40	170	44.4
> 40	108	28.2
Not available	29	7.6
Treatment history		
Retreatment	21	5.5
New	312	81.5
Not available	50	13.0
Total	**383**	**100**

3.2.2 Smear microscopy, culture, and PCR in Germany

The results of the smear microscopy conducted at local laboratories in Eastern Sudan and at the NRL in Germany were discordant. Nearly one-third of the samples (n=123, 32.1%) that were smear-positive with visible AFB at local laboratories in Eastern Sudan were categorised as smear-negative with no visible AFB on the repeat smear microscopy at the NRL in Germany (Figure 3.4). Of these smears that had no visible AFB at the NRL, 89 (72.4%) were scored 1+, 22 (17.9%) were scored 2+, and seven (5.7%) were scored 3+ in the Sudan. Furthermore, the smear scores of five (4.0%) of the samples were missing.

Half of the cultured sputum samples (51.2%, 196/383) showed growth of mycobacteria, of which 87.2% (171/196) and 7.2% (14/196) were identified as *M. tuberculosis* and *M. intracellulare*, respectively. The other 5.6% (11/196) of the *Mycobacterium*-positive cultures had more than one species of bacteria, including seven cultures with *M. tuberculosis* and *M. intracellulare*, one culture with *M. tuberculosis* and *M. fortuitum*, one culture with *M. tuberculosis* and *M. asiaticum*, one culture with *M. tuberculosis* and a *Corynebacterium* species, and one culture with *M. intracellulare* and an unknown NTM (Figure 3.4). The remaining cultures of samples were either contaminated (13.3%, 51/383) or showed no mycobacterial growth i.e. negative (35.5%, 136/383), of which 169 were subjected to an in-house qPCR for the detection of mycobacterial DNA (Figure 3.4). Three-quarters (75.1%, 127/169) tested positive for MTBC DNA and 4.7% (8/169) for NTM DNA.

3.2.3 Positive predictive value of smear microscopy

A quarter (25.2%) of the samples that did not have visible AFB by smear microscopy at the NRL in Germany revealed growth of mycobacteria. This proportion increased to 71% when PCR and culture results were taken together into account (Annex 9). As Figure 3.4 depicts, 90.7% (331/365) of the samples were positive for mycobacteria, either by culture or by PCR and 9.3% (34/365) were negative (no evidence of mycobacteria). Additionally, 15.6% (57/365) of the samples had no evidence of MTBC either by culture or by PCR, resulting in a positive predictive value of 84.4% (308/365, 95% CI from 80.7 to 88.1) of smear microscopy for pulmonary TB in Eastern Sudan. A total of 32 samples showed evidence of NTM by culture (n=24) or PCR (n=8) (Annex 9).

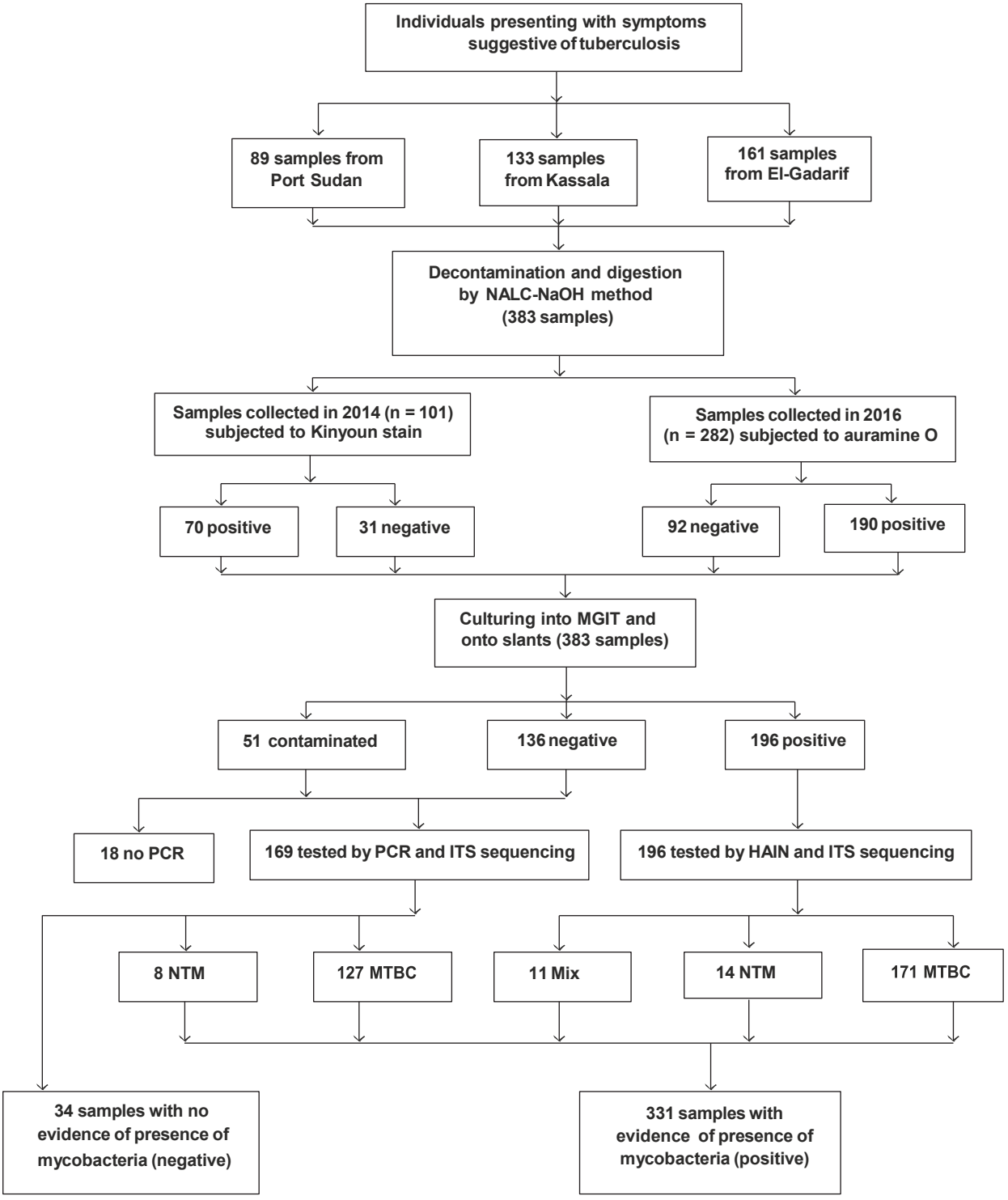


Figure 3.4: Flow of the work from sampling to culture and PCR results. NALC-NaOH = sodium hydroxide/N-acetyl-cysteine, MGIT = mycobacteria growth indicator tube, PCR = polymerase chain reaction, HAIN = line probe assay for GenoType CM and GenoType MTBC, MTBC = *Mycobacterium tuberculosis* complex, NTM = nontuberculous mycobacteria, and Mix = two different species of bacteria grew on the same culture.

3.2.4 Genetic diversity among MTBC strains

Spoligotyping and 24-loci MIRU-VNTR typing results were available for 97.8% (177/181) of all of the MTBC strains. The combined results of the two typing methods classified the strains into four phylogenetic lineages (L1–4) (Annex 10 and 11). L3 (Delhi/CAS) strains were the most prevalent MTBC strains and comprised 73.4% (130/177) of all strains, while L4 strains were the second-largest group of strains with 22.5% (40/177) and were further classified into the following sub-lineage/genotypes: LAM (3.4%, 6/177), S-type (1.1%, 2/177), Sudan H37Rv-like (2.8%, 5/177), UgandaII (2.3%, 4/177), Haarlem (2.8%, 5/177), Cameroon (1.1%, 2/177), X-type (2.3%, 4/177), and TUR (2.3%, 4/177), as well as 4.5% (8/177) Euro-American, i.e. strains with no defined sub-lineage. For L1 and L2, only five (2.8%) and two (1.1%) strains were found, respectively (Annex 10 and 11).

As Table 3.3 shows, MTBC strains with identical MIRU-VNTR and spoligotyping profiles were assigned to molecular clusters as surrogate marker for putative transmission networks. Each cluster was assigned to a unique MLVA 15-9 code using the MIRU-VNTR*plus* on-line database[116]. Overall, 117 strains were assigned to 28 clusters that comprised 2–42 strains/patients resulting in a clustering rate of 66.1%. The MTBC L3 strains with 1557-32 and 4534-32 MLVA 15-9 codes were most prevalent and accounted for 42 (23.7%) and 16 (9.0%) of all MTBC strains, respectively (Table 3.3).

Table 3.3: The most prevalent and clustered MLVA 15-9 strain type in Eastern Sudan

Lineage/sub-lineage	No. of strains (%)	24-loci MIRU-VNTR typing profile	15-9 MLVA code
Delhi/CAS	42 (23.7%)	232236442244225153353743	1557-32
	16 (9.0%)	232236442244225153352743	4534-32
	4 (2.3%)	232236442264225133353743	1649-32
	3 (1.7%)	242236442244225153353743	10425-32
	3 (1.7%)	242247432244225113342543	1064-32
	3 (1.7%)	232226442244225153353743	12560-32
	3 (1.7%)	232232442244225153353643	17280-32
Sudan H37Rv-like	3 (1.7%)	224213322334226153335522	2248-62

No. = number, CAS = Central Asian Strain, and MLVA = Multiple-Locus Variable number tandem repeat Analysis

Following the use of spoligotyping and 24-loci MIRU-VNTR techniques, a WGS approach was further applied to investigate the genetic diversity of the identified clusters and to explore molecular

determinants of drug resistances, as well as to determine the rate of direct patient-to-patient transmission events. A total of six samples were excluded, as only less than 50% of their reads could be mapped to the genome of *M. tuberculosis* H37Rv[52]. In total, 96.7% (175/181) of the strains could be used to explore genetic drug resistance markers, of which 173 were used to determine the rate of direct patient-to-patient transmission events based on a concatenated sequence alignment comprising 11,987 SNPs. The use of WGS confirmed the major lineages identified by the MIRU-VNTR and spoligotyping. Moreover, it enabled the assigning of subgroups with an SNP bar code nomenclature that was recently introduced by Coll et al.[57] (Table 3.4 and Figure 3.5).

Table 3.4: SNP-based classification of MTBC strains from Eastern Sudan

Genotype	Coll lineage (associated MIRU type)	No. of strains	Non-MDR	MDR
Lineage 1	1.1.3 (EAI)	1	0	0
	1.2.2 (EAI)	3	0	0
Lineage 2	2.2.1 (Beijing)	1	1	0
Lineage 3	3 (Delhi/CAS)	121	16	18
	3.1.1 (Delhi/CAS)	7	0	0
Lineage 4	4.3.1 (LAM)	4	0	0
	4.3.3 (LAM)	2	0	0
	4.1 (X-type)	2	1	0
	4.1.1.1 (X-type)	2	1	0
	4.1.2.1 (Haarlem)	5	0	0
	4.4.1.1 (S-type)	2	0	0
	4.6.1.1 (Uganda II)	4	1	0
	4.6.2 (not defined)	1	0	0
	4.6.2.2 (Cameroon)	2	0	0
	4.6 (not defined)	6	0	0
	4.2.2 (TUR)	4	1	0
	4.8 (not defined)	1	0	0
	4.8 (Sudan H37Rv-like)	1	0	0
	4.9 (Sudan H37Rv-like)	4	0	0

No. = number, EAI = East-African-Indian, CAS = Central Asian Strain, and LAM = Latin American Mediterranean

A pairwise SNP distance of ≤12 SNPs between any two MTBC strains was used to identify putative epidemiologically linked patients for recent transmission as suggested by Walker et al.[51]. Based on this threshold, a total of 46.8% (n=81, 95% CI from 39.5 to 54.5) of the strains were grouped into 30 genome clusters comprising two to 8 strains/patients (Figure 3.5). Moreover, 22 genome clusters were identified with ≤5 SNP distance between any two strains.

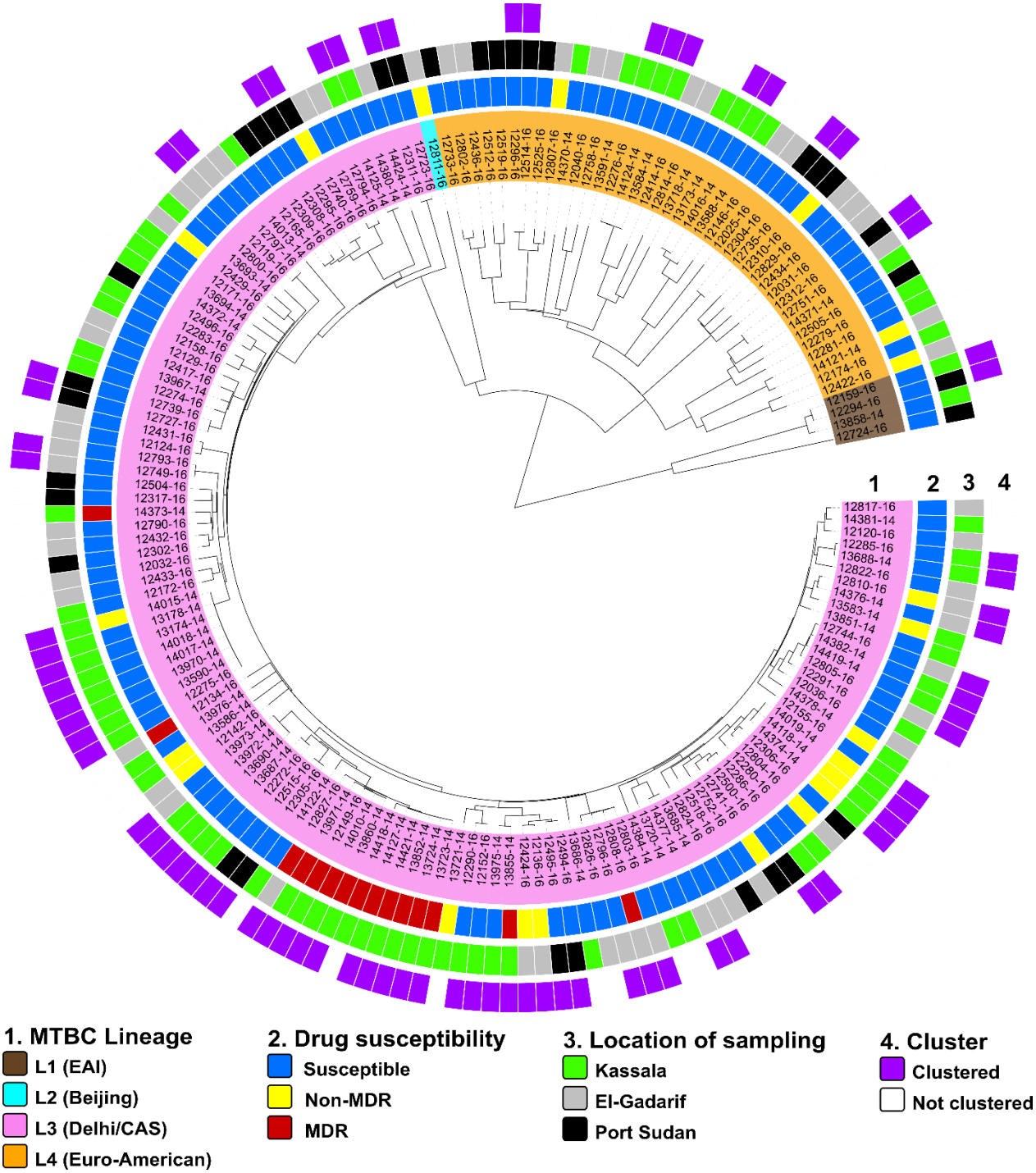


Figure 3.5: The phylogenetic diversity of MTBC strains in Eastern Sudan. A maximum likelihood (ML) tree calculated based on 11,987 SNPs. **(1)** The tree shows MTBC strains belong to 4 different lineages (L1-4), **(2)** first line drugs genotypic DST results, **(3)** geographical location of sample collection, and **(4)** clustered and none clustered strains (distance of ≤12 SNPs).

As expected, clustering rates using the MIRU-VNTR marker were higher in comparison with WGS-inferred clusters among strains of L3 and L4. However, with either of the two methods (i.e. MIRU-VNTR or WGS), L3 strains were observed to have the highest clustering rate (P=0.0061 for 12 bp WGS-based clusters and P≤0.0001 for MIRU-VNTR clustering rates) (Table 3.5). Specifically, L3 strains were observed to have a clustering rate of 53.1% (68/128, 95% CI from 44.1% to 61.9%), considering a 12 bp SNP threshold, while L4 strains were noticed to have a rate of 27.5% (11/40, 95% CI from 15.1% to 44.1%) and L1 strains a rate of 50.0% (2/4, 95% CI from 9.2% to 90.8%)[51]. None of the L2 strains was assigned to any WGS or MIRU-VNTR cluster (Table 3.5).

3.2.5 Drug susceptibility and resistance among MTBC strains

Both, the phenotypic and genotypic drug susceptibility testing (pDST and gDST) results for first line antibiotics were available for 96.7% (175/181) of the MTBC strains. Correlation between the two DSTs is provided in Table 3.5. Overall, 39 (22.3%, 95% CI from 16.5% to 29.3%) of the samples showed resistance to at least one of the five tested antibiotics using pDST, while the rest of the strains (77.7%, n=136) were fully susceptible. Out of the resistant strains, 18 were MDR and represented 10.3% of all isolated MTBC strains. Additionally, 21 (12.0% of all detected MTBC strains) were non-MDR, with 18 harbouring resistance to a single antibiotic and three to multiple antibiotics. Resistances against SM were detected in 20.6% (36/175) of the MTBC strains and were mediated by mutations in the *rspL* (Lys43Arg, Lys88Arg, and Lys88Met), the *gidB* (e.g. Ala138Val), and the *rrs* 514 a/c genes. All INH resistant strains (11.4%, 20/175) were observed either with a mutation in the *katG* gene (Ser315Thr and Ser315Asn) coding for a catalase (drug activator) or in the promotor region of the drug target *InhA* (-15 c/t). Strains with the mutation *inhA* -15 c/t are considered to be low-level INH resistant with MICs around 0.4 mg/l but would be also considered resistant towards the second line drugs ETH and prothionamide[117]. pDSTs for RMP were found in 11.4% (20/175) of all MTBC strains (18 MDR and two non-MDR). All of these strains harboured drug resistance-conferring mutations in the *rpoB* gene (Ser450Leu, His445Tyr, His445Asn, and His445Asp). One EMB resistant strain (0.6%, 1/175) was found with the mutation Gln497Arg in the *embB* gene. However, 13 additional mutations associated with EMB resistance were detected in the *embB* gene, i.e. Met306Ile, Met306Val, and Gly406Asp, but with MICs ranging from 1.25 to 5 µg/ml, strictly classifying these strains as phenotypically susceptible[106]. With regard to PZA, one strain (0.6%, 1/175) was identified with the mutation Gln10Arg in the *pncA* gene, coinciding with phenotypic PZA resistance, while all other strains were genotypically considered to be wild-types in the *pncA* gene and the promotor region.

Table 3.5: Clustering rates using MIRU-VNTR genotyping and WGS among MTBC strains from Eastern Sudan

Lineage/sub-lineage	Strains typed by MIRU-VNTR					Strains typed by WGS (%)				
	No. of strains	No. of clusters	Clustered (%)	95% CI for clustered %	Unique (%)	No. of strains	No. of clusters	Clustered (%)	95% CI for clustered %	Unique (%)
EAI	5	1	2 (40.0)	7.30 - 83.0	3 (60.0)	4	1	2 (50.0)	9.20 - 90.8	2 (50.0)
Beijing	2	0	0 (0.0)	NA	2 (100.0)	1	0	0 (0.0)	NA	1 (100.0)
Delhi/CAS	130	20	100 (76.9)	68.6 - 83.7	30 (23.1)	128	24	68 (53.1)	44.1 - 61.9	60 (46.9)
LAM	6	2	4 (66.7)	24.1 - 94.0	2 (33.3)	6	1	2 (33.3)	6.00 - 75.9	4 (66.7)
S-type	2	1	2 (100)	NA	0 (0.0)	2	1	2 (100)	NA	0 (0.0)
Sudan H37Rv-like*	5	1	3 (60.0)	17.0 - 92.7	2 (40.0)	4	1	3 (75.0)	21.9 - 98.7	1 (25.0)
Uganda	4	1	2 (50.0)	9.20 - 90.8	2 (50.0)	4	1	2 (50.0)	9.20 - 90.8	2 (50.0)
Haarlem	5	1	2 (40.0)	7.30 - 83.0	3 (60.0)	5	1	2 (40.0)	7.30 - 83.0	3 (60.0)
Cameroon	2	0	0 (0.0)	NA	2 (100.0)	2	0	0 (0.0)	NA	2 (100.0)
X-type*	4	0	0 (0.0)	NA	4 (100.0)	2	0	0 (0.0)	NA	2 (100.0)
TUR*	4	0	0 (0.0)	NA	4 (100.0)	4	0	0 (0.0)	NA	4 (100.0)
Not defined*	8	1	2 (25.0)	4.50 - 64.4	6 (75.0)	11	0	0 (0.0)	NA	11 (100.0)
All L4 combined	40	7	15 (37.5)	23.2 - 54.2	25 (62.5)	40	5	11 (27.5)	15.1 - 44.1	29 (72.5)
Total	**177**	**28**	**117 (66.1)**	**58.6 - 72.9**	**60 (33.9)**	**173**	**30**	**81 (46.8)**	**39.5 - 54.5**	**92 (53.2)**

NA = not applicable, CAS = Central Asian Strain, LAM = Latin American Mediterranean, EAI = East-African-Indian, and * = the sub-lineage of one Sudan H37Rv-like, two X-type, and four TUR strains was defined by MIRU-VNTR but these strains were defined as Euro-American by WGS (counted with as sub-lineage not defined by WGS).

Using genetic resistance markers, high sensitivity and specificity for the WGS-based resistance prediction could be achieved (Table 3.6). Nevertheless, varying levels of EMB MICs in genotypic resistant strains with *embB* mutations in codon 306, 406, or 497 resulted in a very low positive predictive value (7.14%) (Table 3.6). Furthermore, *gidB* mutations, including Phe12Ser, Ile114Ser, Arg39Pro, Trp45STP, Iso114Ser, Ser136STP, as well as deletions at positions 4408101, 4408017, and 4408116 other than the *rpoB* mutation Met445Tyr, were considered as mutations with an unclear effect due to the lack of available confidence. However, these strains were phenotypically resistant to the respective drug. Overall, 91 mutations conferring drug resistance were detected by using PhyResSE and by screening for mutations in resistance mediating genes (Annex 12)[43,52]. Delhi/CAS strains harboured the vast majority of these mutations (94.5%, 86/91), whereas the rest of the mutations were found in the Beijing, X-type, Uganda, and TUR strains. The EAI, LAM, S-type, Haarlem, Sudan H37Rv-like, and Cameroon strains had no drug resistant conferring mutations, i.e. fully susceptible to all screened drugs (genotypic wild-type).

Table 3.6: Correlation between phenotypic and genotypic resistances to the first line antibiotics in MTBC strains from Eastern Sudan

Drug	DST R			DST S			Se	Sp	PPV	NPV	Unknown mutations
	gR (TP)	gS (FN)	gU (FN)	gR (FP)	gS (TN)	gU (TN)					
SM	27	0	9	1	138	0	75.0%	99.3%	96.4%	93.9%	9/175 (5.1%)
INH	20	0	0	0	155	0	100%	100%	100%	100%	0/175 (0.0%)
RMP	19	0	1	0	155	0	95.0%	100%	100%	99.4%	1/175 (0.6%)
EMB	1	0	0	13*	161	0	100%	92.5%	7.14%	100%	0/175 (0.0%)
PZA	1	0	0	0	174	0	100%	100%	100%	100%	0/175 (0.0%)

SM = streptomycin, INH = isoniazid, RMP = rifampicin, EMB = ethambutol, PZA = pyrazinamide, DST = drug susceptibility testing, R = resistant, S = susceptible, gR = genetically resistant, gS = genetically susceptible, gU = genetical resistance unknown, *See methods, Se = sensitivity which is true positive (TP) divided by TP plus false-negative (FN), Sp = specificity which is true negative (TN) divided by TN plus false positive (FP), PPV = positive predictive value which is TP divided by (TP+FP), and NPV = negative predictive value which is TN divided by (TN+FN)

With regard to recent transmission, 77.8% (14/18) of the MDR strains and 42.9% (9/21) of the non-MDR strains were associated with transmission chains with a 12 bp SNP threshold[51]. The linked MDR strains congregated in two SNP clusters (i.e. Groups A and B) as Table 3.7 illustrates. Two MDR strains (i.e. Group C) were clustered but with genotypic wild-type strains. Additionally, three MDR strains were unique, i.e. not clustered. The detected clustered non-MDRs, belonged to six clusters, of which only two clusters were formed by non-MDR strains. In addition, 12 non-MDR strains were unique.

Table 3.7: Detected mutations that mediate resistances against the tested first line anti-TB drugs in 18 MDR MTBC strains from Eastern Sudan

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	≤12 SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
1	13852/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
2	13723/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (ggc/gAc), Gly406Asp (ggc/gAc)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
3	14421/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
4	14127/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
5	13857/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
6	14375/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
7	13724/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group A
8	12827/16	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	R	embB Met306Val (atg/Gtg)	R	pncA Gln10Arg (cag/cGg)	Delhi/CAS	Delhi/CAS	CL	Group B
9	12149/16	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group B
10	14010/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Group B
11	13971/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Group B
12	13860/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Group B

Table 3.7: continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
13	12808/16	S	WT	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Asp (cac/Gac)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Group C
14	13855/14	R	rpsL Lys43Arg (aag/aGg)	R	fabG1 Thr4Iso (aca/aTa)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Group C
15	12134/16	R	gidB Ala138Val (gcg/gTg)	R	katG Ser315Asn (agc/aCc)	R	rpoB His445Asn (cac/Cac)	R	embB Gln497Arg (cag/cGg)	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Uq
16	14418/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Uq
17	14373/14	S	WT	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Uq
18	13719/14	R	WT	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser431Thr (agc/aCc)	S	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Un

SM = streptomycin, INH = isoniazid, RMP = rifampicin, EMB = ethambutol, PZA = pyrazinamide, pDST = phenotypic drug susceptibility testing, gDST = genotypic drug susceptibility testing, R = resistant, S = susceptible, WT = wild-type, Uq = unique, Un = unknown, CL = clustered, and Gr = grouped

Taken together, the results of this part of my thesis showed that 90.7% (331/365) of the sputum samples were positive for mycobacteria, either by culture or by PCR, and 9.3% (34/365) were negative (i.e. no evidence of mycobacteria). Additionally, the positive predictive value for MTBC detection by smear microscopy was 84.4% (308/365) in Eastern Sudan. Moreover, approximately half (51.2%, 196/383) of the specimens revealed growth of mycobacteria, with the majority of these strains (n=171, 87.2%) being *M. tuberculosis*. Any drug resistance was detected in 22.3% (39/175) of all *M. tuberculosis* strains, including 10.3% (18/175) MDR strains. MTBC L3 strains dominated the MTBC population structure (73.4% of the strains) and were observed as having the highest clustering rate (53.1% based on WGS pairwise distances). Furthermore, MTBC L3 strains harboured the majority of the observed drug resistances (non-MDR and MDR).

3.3 The global population structure of MTBC lineage 3

The dominance of MTBC L3 strains in Eastern Sudan and their association with drug resistance underline previous reports on the epidemiological relevance of this lineage in North/East Africa and West and South Asia[4,60,118]. Thus, reconstruction of the phylogeny of L3 strains from these geographical settings is crucial for a proper understanding of TB epidemiology in these regions. Results can inform on putative differences in pathogenicity, reveal ongoing transmission networks, and provide a robust classification of L3 sub-lineages.

To reveal the global population structure of L3, a total of 1,685 strains from 22 different countries were analysed. A number of collaborators provided 24-loci MIRU-VNTR data and different studies were included (Annex 13 and 14). Of the 1,685 included strains, 710 were typed at the Research Center Borstel, Germany. One-third (37.2%, n=627) of the strains were from African countries, 31.2% (n=526) from European countries, 26.1% (n=440) from Asian countries, and the geographical origins of 5.5% (n=92) of the strains were not available.

3.3.1 24-loci MIRU-VNTR-based phylogeny

The MIRU-VNTR typing results of all strains were uploaded to the www.miru-vntrplus.org on-line database to confirm their genotypes and to assign multiple locus VNTR analysis 15-9 codes (MLVA 15-9 codes) to each strain according to the 24-loci MIRU-VNTR nomenclature[116]. A total of 675 different MLVA 15-9 types were identified, of which 519 could be assigned to a single strain and 156 to two strains or more (Annex 15). Strains with identical MIRU-VNTR profiles were assigned to molecular clusters as surrogate marker for recent transmission. The seven largest MLVA 15-9 clusters are shown in Table 3.8. Nonetheless, the full MLVA 15-9 codes of 381 strains could not be retrieved due to missing loci.

Table 3.8: The largest MLVA 15-9 clusters of lineage 3

No.	MLVA 15-9 type	Number	%
1	1064-32	81	4.8
2	1557-32	68	4.0
3	1061-32	33	2.0
4	1220-15	27	1.6
5	4534-32	22	1.3
6	1212-32	21	1.3
7	1449-32	20	1.2

To visualise the global population structure of L3, a minimum spanning (MS) tree was calculated using 24-loci MIRU-VNTR typing data (Figure 3.6). In the MS tree, five Delhi/CAS Miru Groups (DCMGs) or clonal complexes could be defined. Each DCMG represented a separate sub-lineage (SL) and included closely related strains that originated from one parental central node. In addition to the distinct DCMGs 1–5, a putative ancestral and basal sub-lineage (i.e. DCMG6) that contained strains with diverse 24-loci MIRU-VNTR patterns was also defined. DCMGs 1–5 were associated with the following main MLVA 15-9 types: 1557-32, 1061-32, and 4534-32 (DCMG1), 1064-32, 1295-32, and 1422-32 (DCMG2), 1449-32 and 154-62 (DCMG3), 1220-15 (DCMG4), and 1212-32 (DCMG5), respectively. DCMG6 strains emerged from the central node in a strong star-like manner and likely represented a larger ancestral diversity of L3 strains. Moreover, strains of the parental central node share the 1231-32 MLVA 15-9 type and originate from Pakistan (n=1), India (n=1), Ireland (n=1), and Germany (n=2).

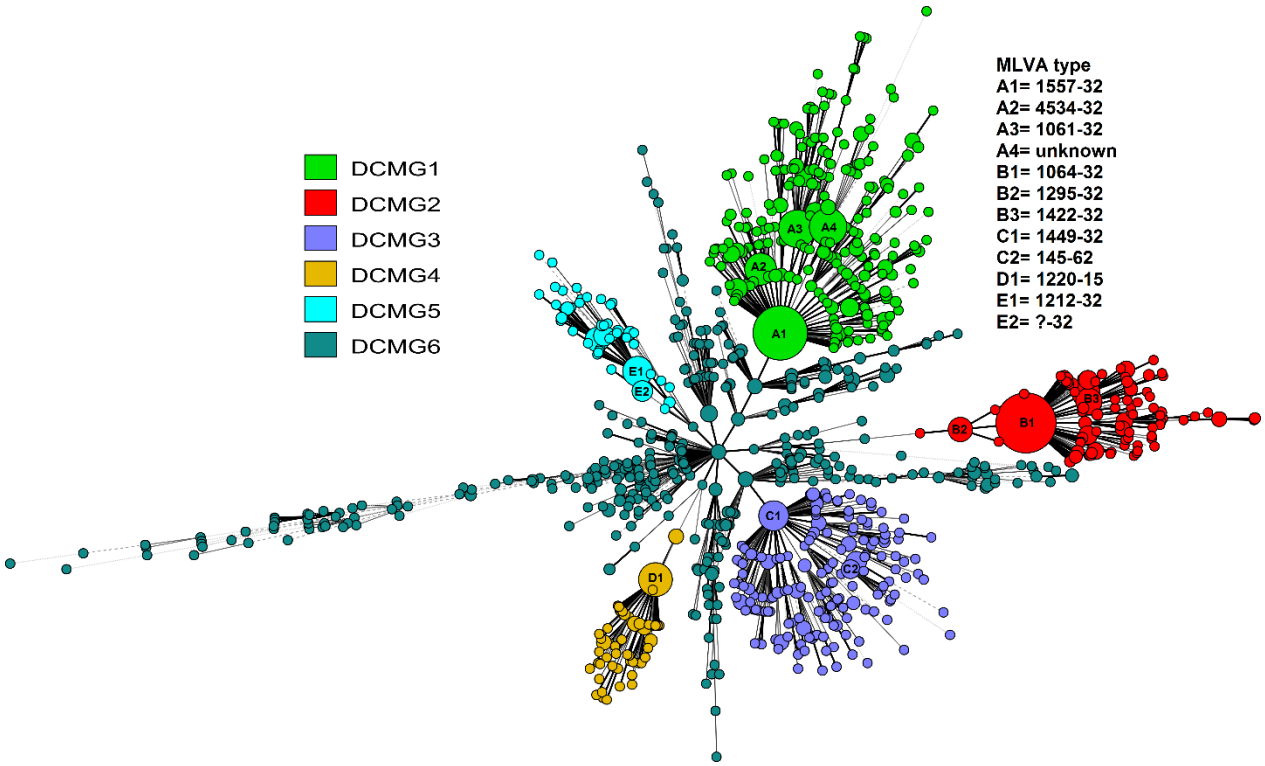


Figure 3.6: Delhi/CAS MIRU-VNTR Groups (DCMGs). Minimum spanning (MS) tree calculated based on 24-loci MIRU-VNTR data of 1,685 clinical strains of MTBC L3 originating from 22 countries. The MS tree shows the different DCMGs and the global population structure of MTBC L3. Strains with similar 24-loci MIRU-VNTR pattern are grouped together in one node (i.e. cluster). The five distinct DCMGs were defined in the MS tree. Each group is indicated by a different color. The length of the branches denotes to the number of allele changes between two patterns. Solid lines represent 1, 2 or 3 changes while gray dashed lines represent 4 changes and gray dotted lines represent 5 or more changes.

By grouping strains with identical 24-loci MIRU-VNTR patterns into a molecular cluster, 908 L3 strains were clustered in a total of 195 clusters with an overall clustering rate of 53.9% (95% CI from 51.5 to 56.2). Cluster sizes varied from clusters containing two strains to clusters containing 81 strains. The highest clustering rate was observed among the strains of DCMG2 (77.9%, 218/280), while the lowest clustering rate was among the strains of DCMG6 (30.3%, 140/462). Strains of DCMG1 and 3–5 showed intermediate clustering rates that ranged from 38.2% to 67.3% (Figure 3.7).

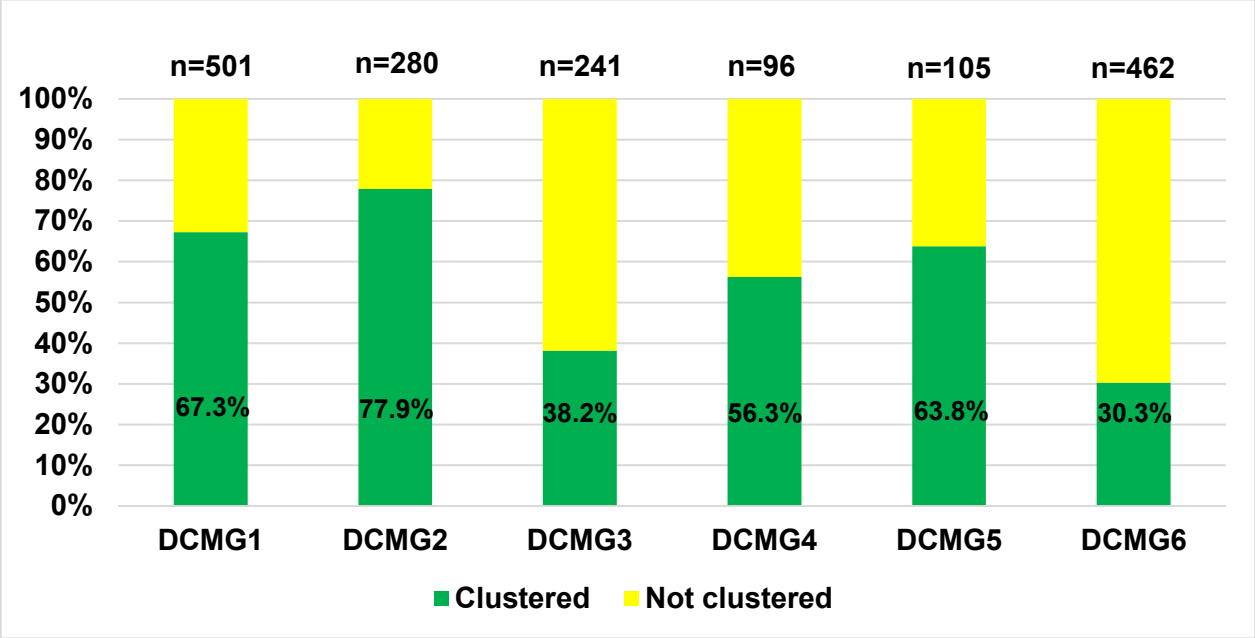


Figure 3.7: Clustering rates among DCMGs. Proportions of clustered and not clustered strains are shown in green and yellow. Proportions (%) of clustered strains are indicated. n = total number of strains in each DCMG.

A comparison of the proportions of clustered strains in each DCMG using a logistic regression analysis revealed significant differences and showed that strains of DCMG 1–5 have an increased probability of being clustered in comparison to strains of DCMG6 (i.e. the reference group [OR=1]), with odds ratios (OR) ranging from 1.4 to 8.1 and a P-value of ≤0.034 (Table 3.9).

Table 3.9: Logistic regression analysis comparing the proportions of clustered strains among DCMGs

DCMGs	No. of strains	No. of clustered (%)	OR	P-value	95% CI for OR
					Lower - Upper
6	462	140 (30.3)	ref		
1	501	337 (67.3)	4.7	<0.001	3.599 - 6.062
2	280	218 (77.9)	8.1	<0.001	5.728 - 11.41
3	241	92 (38.2)	1.4	0.036	1.024 - 1.969
4	96	54 (56.3)	3.0	<0.001	1.887 - 4.635
5	105	67 (63.8)	4.1	<0.001	2.559 - 6.327

DCMGs = Delhi/CAS Miru Groups, No. = number, OR = odds ratios (i.e. increased or decreased probability [OR ≠1]), CI = confidence interval, and ref = odds ratio equals 1

Out of the 195 clusters, 76 contained L3 strains that were found in two to 10 different countries (e.g. strains of cluster B1 [see Figure 3.6] and no. 103 [see Table 3.10] having the 1064-32 MLVA code were found in India, Tanzania, Kenya, Ethiopia, Sudan, Zambia, Eswatini, Germany, Sweden, and Ireland). The full MLVA 15-9 code could be retrieved for L3 strains in all clusters,

except for the strains in clusters 119, 182, 184, and 185 due to one or two missing MIRU-VNTR loci (Table 3.10).

Table 3.10: 24-loci MIRU-VNTR clusters of L3 strains reported in at least two different countries

No. of strains	No. of cluster	DCMG	MLVA code	Country	No. of non-MDR (%)	No. of MDR (%)
5	2	4	2218-15	India, Germany	1 (20.0)	1 (20.0)
4	3	4	3667-15	India, Pakistan, Germany	1 (25.0)	0 (0.0)
2	5	4	12410-15	India, Sweden	0 (0.0)	0 (0.0)
4	6	4	3354-15	India, Pakistan, Germany	0 (0.0)	0 (0.0)
2	7	4	3354-15	Pakistan, Germany	0 (0.0)	1 (50.0)
5	8	4	7122-15	Tanzania, Kenya, Cameroon, Germany	0 (0.0)	0 (0.0)
27	9	4	1220-15	Pakistan, Germany, Sweden	8 (29.6)	5 (17.9)
3	13	1	4539-32	Ethiopia, Germany	0 (0.0)	0 (0.0)
68	14	1	1557-32	Ethiopia, Sudan, Germany, Sweden, Ireland	8 (11.8)	2 (2.94)
2	15	1	1062-32	Sudan, Germany	0 (0.0)	1 (50.0)
3	16	1	4538-32	Sudan, Germany	0 (0.0)	0 (0.0)
3	22	1	7186-32	Ethiopia, Sudan	1 (33.3)	0 (0.0)
33	26	1	1061-32	Ethiopia, Germany, Sweden	2 (6.06)	10 (30.3)
6	29	1	7785-32	Ethiopia, Germany	0 (0.0)	0 (0.0)
2	33	1	6974-32	Sudan, Sweden	1 (50.0)	0 (0.0)
2	36	1	4535-32	Ethiopia, Sudan	1 (50.0)	0 (0.0)
3	38	1	15106-32	Ethiopia, Sudan	0 (0.0)	0 (0.0)
8	40	1	1649-32	Ethiopia, Sudan, Iraq, Germany	0 (0.0)	1 (12.5)
2	41	1	11471-32	Ethiopia, Germany, Sweden	0 (0.0)	1 (50.0)
4	46	1	12560-32	Sudan, Germany	0 (0.0)	1 (25.0)
2	50	1	12087-32	Sudan, Sweden	1 (50.0)	0 (0.0)
2	51	1	1557-25	Ethiopia, Sudan	0 (0.0)	0 (0.0)
22	53	1	4534-32	Sudan, Germany, Sweden, Ireland	1 (4.54)	16 (72.7)
3	54	1	9259-32	Sudan, Germany	1 (33.3)	0 (0.0)
9	55	1	1557-419	Ethiopia, Germany, Sweden	1 (11.1)	0 (0.0)
2	56	1	1219-32	Ethiopia, Germany	0 (0.0)	0 (0.0)
5	59	1	9048-332	Ethiopia, Germany	0 (0.0)	0 (0.0)
4	61	1	12525-32	Sudan, Sweden	0 (0.0)	0 (0.0)
4	68	1	6755-32	Ethiopia, Germany	0 (0.0)	0 (0.0)
5	74	6	4536-32	Germany, Sweden	0 (0.0)	0 (0.0)
3	79	6	12557-32	India, Sweden, Germany	0 (0.0)	0 (0.0)
5	83	6	1231-32	India, Pakistan, Germany, Ireland	0 (0.0)	0 (0.0)
2	85	6	17366-32	India, Germany	1 (50.0)	0 (0.0)
2	87	5	9069-32	Pakistan, Ethiopia	0 (0.0)	0 (0.0)
21	88	5	1212-32	Tanzania, Kenya, Ethiopia, Germany, Sweden, Ireland	0 (0.0)	0 (0.0)
14	90	5	1212-25	Ethiopia, Senegal, Germany, Sweden	3 (21.4)	0 (0.0)
3	92	5	5070-25	Kenya, Sweden	0 (0.0)	0 (0.0)
9	94	5	12760-32	Sudan, Germany	0 (0.0)	0 (0.0)
2	95	6	12569-32	Germany, Sweden	0 (0.0)	0 (0.0)
4	97	6	8966-32	Iraq, Germany	0 (0.0)	3 (75.0)

Table 3.10 continued

No. of strains	No. of cluster	DCMG	MLVA code	Country	No. of non-MDR (%)	No. of MDR (%)
81	103	2	1064-32	India, Tanzania, Kenya, Ethiopia, Sudan, Zambia, Eswatini, Germany, Sweden, Ireland	2 (2.47)	22 (27.2)
3	105	2	1064-145	Ethiopia, Germany	0 (0.0)	2 (66.7)
8	107	2	4439-32	Tanzania, Kenya, Eswatini, Sweden	0 (0.0)	0 (0.0)
16	110	2	1422-32	Tanzania, Kenya, Mozambique, Congo, Germany	0 (0.0)	1 (6.25)
3	114	2	2136-32	Kenya, Germany	0 (0.0)	0 (0.0)
6	116	2	1064-25	Tanzania, Kenya	0 (0.0)	1 (16.7)
4	119	2	?-545	Tanzania, Kenya, Germany	0 (0.0)	0 (0.0)
7	121	2	12794-32	Kenya, Germany	0 (0.0)	1 (14.3)
2	122	2	7210-229	Mozambique, Eswatini	0 (0.0)	0 (0.0)
13	124	2	1295-32	Kenya, Ethiopia, Germany, Ireland	0 (0.0)	1 (7.69)
3	127	2	17651-32	Kenya, Germany	0 (0.0)	0 (0.0)
4	128	6	3612-32	Germany, Ireland	0 (0.0)	0 (0.0)
2	132	6	3627-32	Pakistan, Germany	0 (0.0)	1 (50.0)
3	135	6	9139-32	Uzbekistan, Turkmenistan	1 (33.3)	2 (66.7)
20	140	3	1449-32	India, Pakistan, Germany, Ireland	2 (10.0)	5 (25.0)
5	142	3	317-32	Pakistan, Germany, Ireland	0 (0.0)	1 (20.0)
7	145	3	145-62	Turkmenistan, Germany	3 (43.9)	1 (14.3)
2	146	3	5847-32	Germany, Sweden	0 (0.0)	0 (0.0)
4	147	3	1320-32	Pakistan, Germany	1 (25.0)	0 (0.0)
4	148	3	9072-32	Ethiopia, Germany, Sweden	0 (0.0)	0 (0.0)
3	154	3	1449-145	Iraq, Germany	0 (0.0)	2 (66.7)
9	155	3	3549-145	Iraq, Germany	0 (0.0)	8 (88.9)
5	156	3	1534-145	India, Pakistan, Germany, Ireland	0 (0.0)	0 (0.0)
2	158	6	2215-32	Germany, Ireland	0 (0.0)	0 (0.0)
3	159	6	3633-32	India, Pakistan, Germany	0 (0.0)	0 (0.0)
2	160	6	3635-32	India, Pakistan	0 (0.0)	0 (0.0)
3	163	3	17165-32	Germany, Ireland	0 (0.0)	0 (0.0)
4	164	3	5085-32	India, Germany	1 (25.0)	0 (0.0)
3	168	6	3644-32	Pakistan, Germany	0 (0.0)	1 (33.3)
6	170	6	4727-32	India, Iran, Germany	0 (0.0)	0 (0.0)
4	174	6	9928-32	Ghana, Germany	0 (0.0)	1 (25.0)
2	179	6	3649-32	Pakistan, Ireland	1 (50.0)	0 (0.0)
3	182	6	?-32	Ethiopia, Sudan	1 (33.3)	0 (0.0)
2	183	6	345-32	Ethiopia, Germany	0 (0.0)	0 (0.0)
2	184	2	?-32	Tanzania, Ghana	0 (0.0)	0 (0.0)
2	185	2	?-25	Tanzania, Kenya	0 (0.0)	0 (0.0)

No. = number, DCMG = Delhi/CAS Miru Group, and MLVA = multiple loci VNTR analysis

3.3.2 Geographical distribution of DCMGs

To analyse the global distribution of DCMGs, the country of origin's geographic coordinates were assigned to each of the analysed clinical strains with the software Ridom MLVA Compare v1.03 (Ridom, Münster, Germany). The geographical distribution of the strains of the DCMGs differed between Africa, Asia, and Europe (Figure 3.8 and Table 3.11). Strains of DCMG1 were dominant in North/East Africa (Sudan, Ethiopia, and Djibouti) and West Asia (Saudi Arabia and Iraq) (North/East Africa West Asian DCMG), whereas strains of DCMG2 were almost totally confined to East Africa (East-African DCMG). The largest number of strains of DCMG2 was detected in Tanzania and Kenya, with small numbers reported in Ethiopia, Djibouti, and the Sudan. However, DCMG3 strains were restricted to Asia, being mainly prevalent in South (India and Pakistan), Central Asia (Turkmenistan), and West (Saudi Arabia and Iraq) Asia (Asian DCMG). Strains of DCMG4 were mainly prevalent in South Asia in India and Pakistan (South Asian DCMG), but strains of DCMG5 were found in Ethiopia, Djibouti, Kenya, and Tanzania in East Africa in addition to Saudi Arabia in West Asia (East Africa West Asian DCMG). Strains of DCMG6 were prevalent in most of the 22 investigated countries in this study, though larger numbers were found in South and West Asia. Strains of all DCMGs were seen in Europe in Germany, Ireland, and Sweden and were mostly found among immigrants (Figure 3.8 and Table 3.11)[119-121].

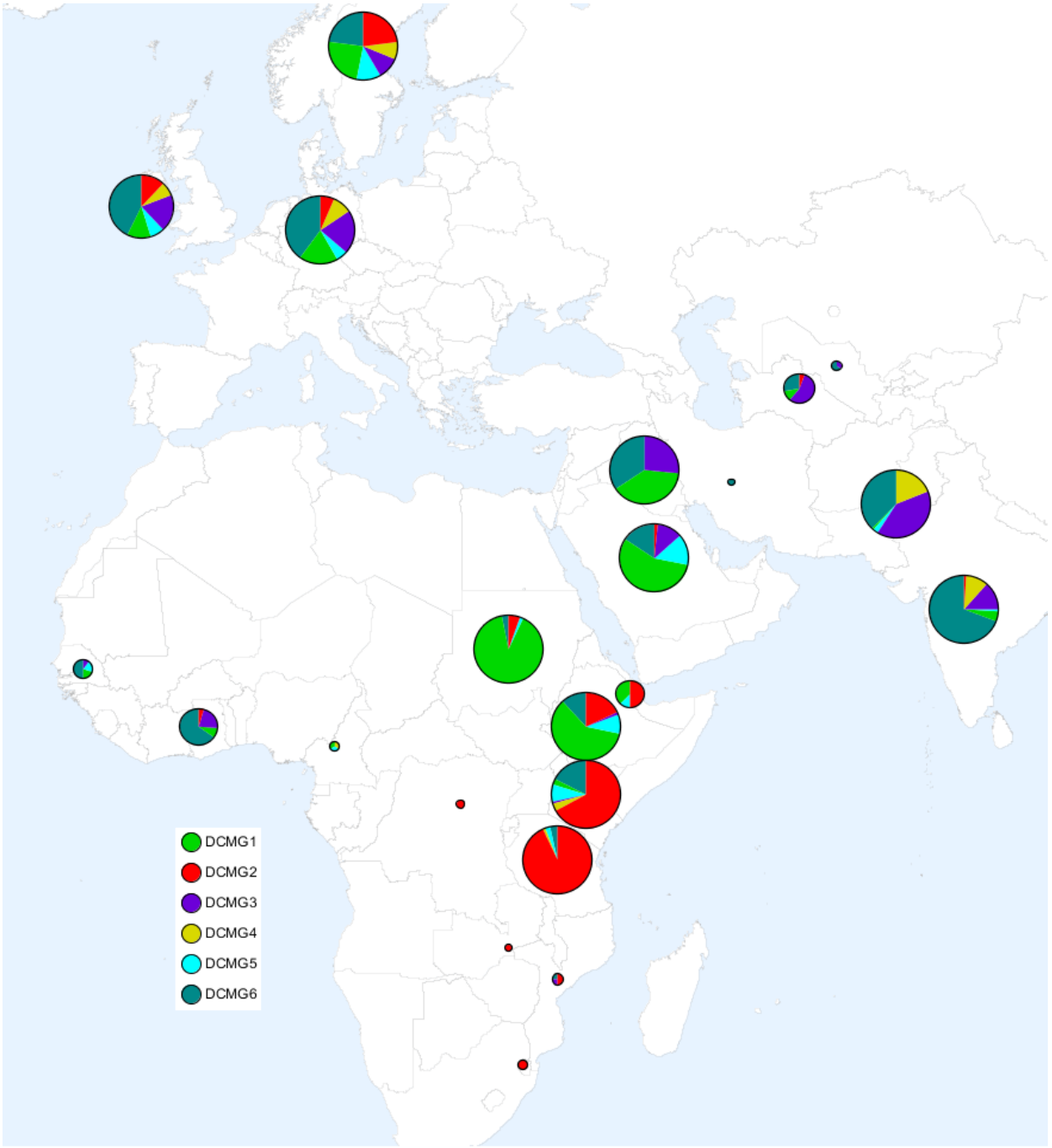


Figure 3.8: The global geographical distribution of DCMGs. Each pie chart represents a country and the size of each pie chart is proportional to the number of strains. Each DCMG is indicated by a different color.

Table 3.11: Proportional geographical distribution of strains of DCMGs in Asia, Africa, and Europe

DCMGs	Southern Asia	Central Asia	Western Asia	Northern Africa	Eastern Africa	Central Africa	Western Africa	Southern Africa	Europe	Unknown origin	Total
1	6 (1.2%)	2 (0.4%)	97 (19.4%)	117 (23.4%)	153 (30.5%)	1 (0.2%)	4 (0.8%)	0 (0.0%)	99 (19.8%)	22 (4.4%)	**501**
2	1 (0.4%)	1 (0.4%)	2 (0.7%)	7 (2.5%)	207 (73.9%)	2 (0.7%)	1 (0.4%)	3 (1.1%)	52 (18.6%)	4 (1.4%)	**208**
3	63 (26.1%)	11 (4.6%)	32 (13.3%)	0 (0.0%)	5 (2.1%)	0 (0.0%)	6 (2.5%)	0 (0.0%)	99 (41.1%)	25 (10.4%)	**241**
4	34 (35.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	5 (5.2%)	1 (1.0%)	0 (0.0%)	0 (0.0%)	46 (47.9%)	10 (10.4%)	**96**
5	4 (3.8%)	0 (0.0%)	19 (18.1%)	2 (1.9%)	35 (33.3%)	1 (1.0%)	2 (1.9%)	0 (0.0%)	35 (33.3%)	7 (6.7%)	**105**
6	119 (25.8%)	7 (1.5%)	42 (9.1%)	4 (0.9%)	51 (11.0%)	0 (0.0%)	20 (4.3%)	0 (0.0%)	195 (42.2%)	24 (5.2%)	**462**
Total	**227**	**21**	**192**	**130**	**456**	**5**	**33**	**3**	**526**	**92**	**1,685**

3.3.3 Drug susceptibility and resistance among strains of DCMGs

Phenotypic DST data were available for 63.5% (1,070/1,685) of the analysed L3 strains. Four MDR TB cohort studies and a national drug surveillance were included in the analysis. A proportion of 58.5% (626/1,070), with a 95% CI from 55.6 to 61.5, of the strains that had pDST data available were fully susceptible to all first line drugs. Nevertheless, 21.5% (231/1,070) were L3 MDR strains and 19.9% (213/1,070) were L3 resistant strains but non-MDR (Figure 3.9 and Table 3.12).

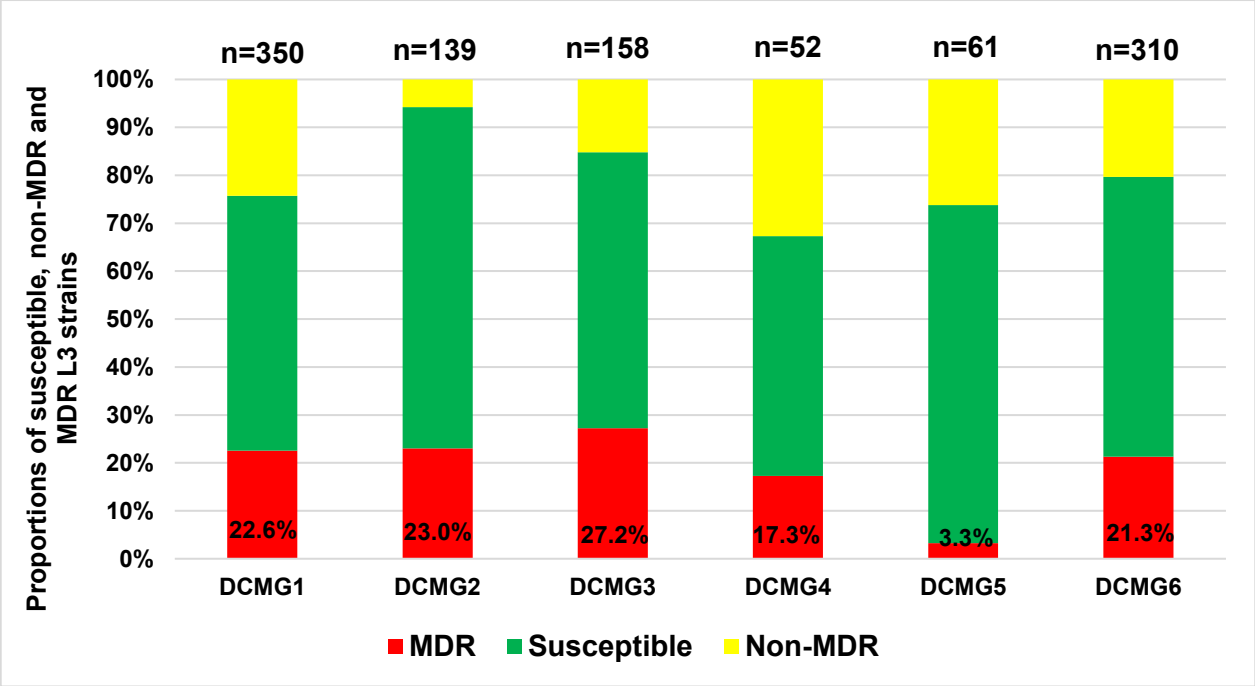


Figure 3.9: Drug susceptibility pattern among the strains of DCMGs. Proportions of fully susceptible, resistant but non-MDR, and MDR L3 strains are shown in different colors. Proportions (%) of L3 MDR strains are indicated. n = total number of strains with pDST results available in each DCMG.

Proportions of L3 MDR strains among the six DCMGs were as follows: 22.6% (79/350), 23.0% (32/139), 27.2% (43/153), 17.3% (9/52), 3.3% (2/61), and 21.3% (66/310), respectively (Figure 3.9 and Figure 3.10). Significant statistical differences were observed between these proportions by using a logistic regression analysis (Table 3.12). L3 MDR strains were more likely to occur in DCMG 1–4 and 6 with an OR of 6.17–11.0 and a P-value of ≤0.024 in comparison to DCMG5 (i.e. the reference group [OR=1]) (Table 3.12).

Table 3.12: Numbers and proportions of susceptible, non-MDR, and MDR L3 strains in each DCMG and logistic regression analysis for MDR L3 strains among of each DCMG

DCMGs	Susceptible		Non-MDR		MDR				Total
	No.	% (95% CI)	No.	% (95% CI)	No.	% (95% CI)	OR	P-value	
5	43	70.5 (58.1-80.5)	16	26.2 (16.8-38.5)	2	3.30 (0.30 -11.6)	ref	-	61
1	186	53.1 (47.9-58.3)	85	24.3 (20.1-29.0)	79	22.6 (18.5-27.3)	8.60	0.003	350
2	99	71.2 (63.2-78.1)	8	5.80 (2.60-11.3)	32	23.0 (16.8-30.7)	8.80	0.004	139
3	91	57.6 (49.8-65.0)	24	15.2 (10.4-21.8)	43	27.2 (20.9-34.7)	11.0	0.001	158
4	26	50.0 (36.9-63.1)	17	32.7 (21.5-46.3)	9	17.3 (9.20-29.9)	6.17	0.024	52
6	181	58.4 (52.8-63.7)	63	20.3 (16.2-25.2)	66	21.3 (17.1-26.2)	7.89	0.005	310
Total	626	58.3 (55.6-61.5)	213	19.9 (17.5-22.3)	231	21.6 (19.0-24.0)	-	-	1,070

DCMGs = Delhi/CAS Miru Groups, No. = number, OR = odds ratios (i.e. increased or decreased probability [OR ≠1]), CI = confidence interval, and ref = odds ratio equals 1

The overall clustering rate of L3 MDR strains was 62.8% (145/231) (Table 3.13). With the exception of DCMG5, which had only two unique MDR strains, the proportions of clustered L3 MDR strains varied per DCMG. The highest clustering rate was noticed among the MDR strains of DCMG2 (93.8%, 30/32) and the lowest was among the strains of DCMG6 (45.5%, 30/66). L3 MDR strains of DCMG1 and 2 were three to 18 times more likely to be clustered than L3 MDR strains of the reference group (i.e. DCMG6), with P-values of 0.002 and <0.001 (Table 3.13). The smallest cluster contained two MDR strains, whereas the largest cluster contained 22 strains (Figure 3.10).

Table 3.13: Numbers and proportions of clustered and unique MDR strains in each of the DCMGs and logistic regression analysis for clustered MDR L3 strains

DCMGs	Unique			Clustered					Total
	No.	%	95% CI	No.	%	95% CI	OR	P-value	
6	36	54.5	42.6-65.9	30	45.5	34.0-57.4	ref	-	66
1	23	29.1	55.6-61.5	56	70.9	60.0-79.8	2.92	0.002	79
2	2	6.20	0.72-21.2	30	93.8	78.8-99.3	18.0	<0.001	32
3	21	48.8	34.6-63.3	22	51.2	36.8-65.4	1.30	0.560	43
4	2	22.2	5.30-55.7	7	77.8	44.3-94.7	4.20	0.087	9
5	2	100	-	0	0.00	-	-	-	2
Total	86	37.2	31.3-43.6	145	62.8	56.4-68.8	-	-	231

DCMGs = Delhi/CAS Miru Groups, No. = number, OR = odds ratios (i.e. increased or decreased probability [OR ≠1]), CI = confidence interval, and ref = odds ratio equals 1

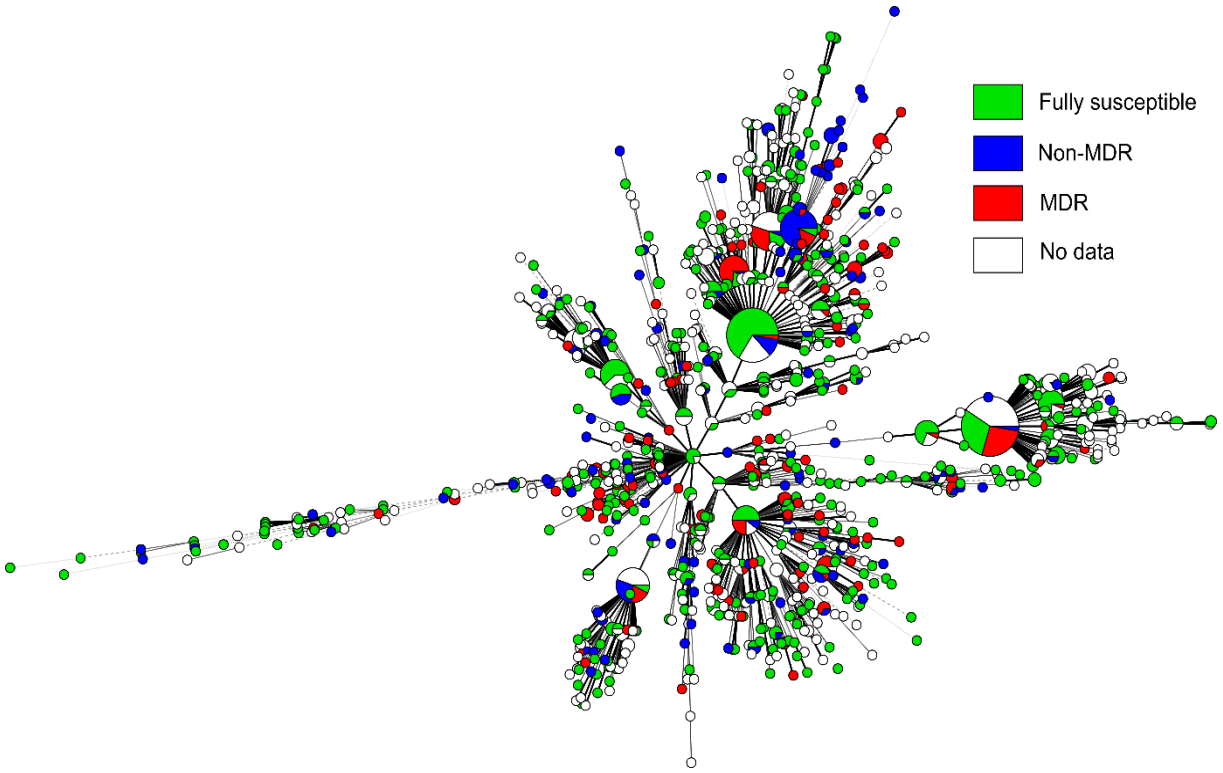


Figure 3.10: Clusters of drug resistant strains among DCMGs. Minimum spanning (MS) tree calculated based on 24-loci MIRU-VNTR data of 1,685 clinical strains of lineage 3. The MS tree is showing the clusters of drug resistant strains among the DCMGs. Fully susceptible, resistant but not MDR, and MDR are indicated by different colors. The length of the branches denotes to the number of allele changes between two patterns. Solid lines represent 1, 2 or 3 changes while gray dashed lines represent 4 changes and gray dotted lines represent 5 or more changes.

3.3.4 Genome-based phylogeny of MTBC lineage 3

To analyse the phylogeny at a higher resolution and to define genome-based differences among strains from different geographical regions, 159 representative strains were selected from all DCMGs for WGS (Figure 3.11). The proportions of selected strains from each DCMG were as follows: 7.0% (35/501) from DCMG1, 9.4% (24/241) from DCMG2, 10.0% (27/280) from DCMG3, 8.3% (8/96) from DCMG4, 8.6% (9/105) from DCMG5, and 12.1% (56/462) from DCMG6 (Figure 3.11). These strains originated from 13 different African, Asian, and European countries. For each strain, at least 97% of the genome was mapped to the *M. tuberculosis* H37Rv reference genome (GenBank ID: NC_000962.2), with at least 75% variant frequency[52]. The combined analysis of all SNPs of the 159 strains resulted in 12,262 variable positions.

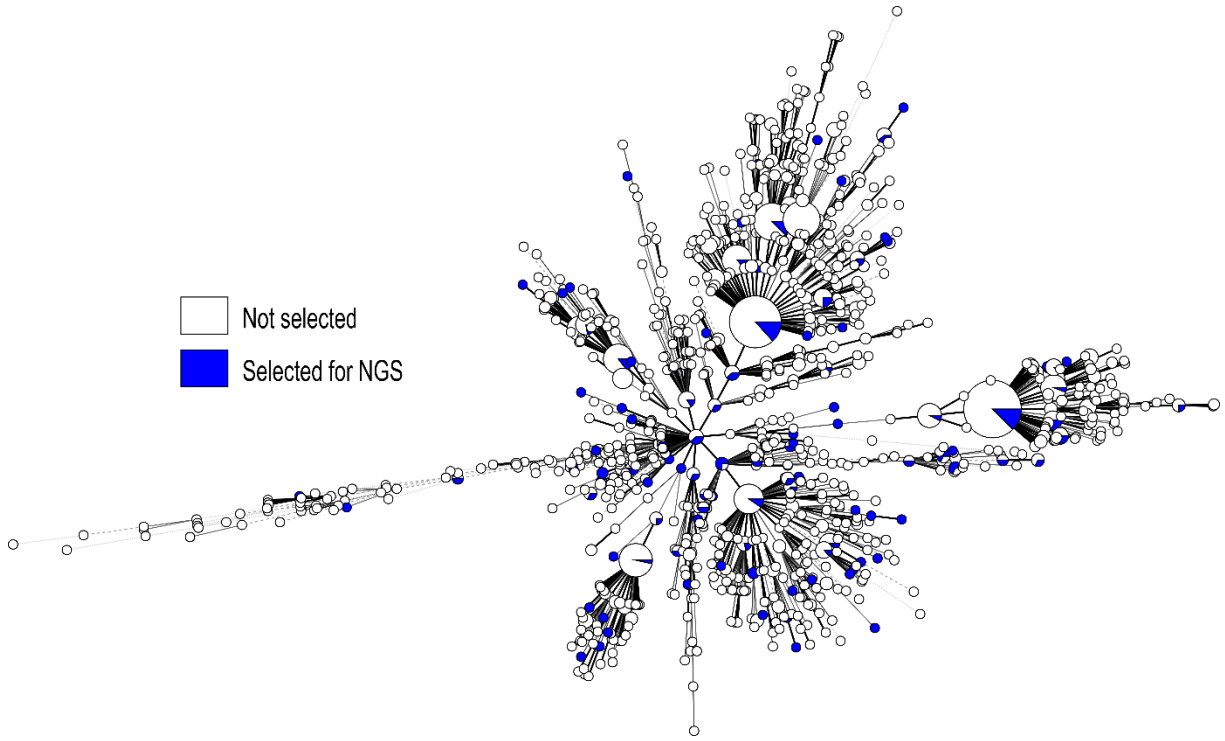


Figure 3.11: Selected lineage 3 strains for whole genome sequencing. Minimum spanning (MS) tree calculated based on 24-loci MIRU-VNTR data of 1,685 clinical strains of L3. The MS tree is showing strains of L3 that were selected for whole genome sequencing. The length of the branches denotes to the number of allele changes between two patterns. Solid lines represent 1, 2 or 3 changes while gray dashed lines represent 4 changes and gray dotted lines represent 5 or more changes.

A maximum likelihood (ML) approach based on the SNP data was used to reconstruct the phylogeny of L3. The SNP-based groups (i.e. sub-linages) were defined/inferred based on the ML tree topology, MIRU-VNTR clonal complexes, and by considering the geographical information of strains' origin using an in-house customised pipeline (Figure 3.12). The SNP-based ML groups were mostly congruent with the 24-loci MIRU-VNTR phylogenetic classification (Chi square = 642.74, P-value <2.2e-16). However, for some strains, the SNP-based grouping and the 24-loci MIRU-VNTR phylogenetic classification were different. Additionally, the ML showed a high support of more than 90% for the main branches and, thus, a robust topology. The genomic analysis of the population structure allowed for the classification of L3 strains into eight subgroups (Figure 3.12). The length of the branches in the ML tree was proportional to the substitution rate and, consequently, provided insight into the diversity within and among the subgroups (Figure 3.12).

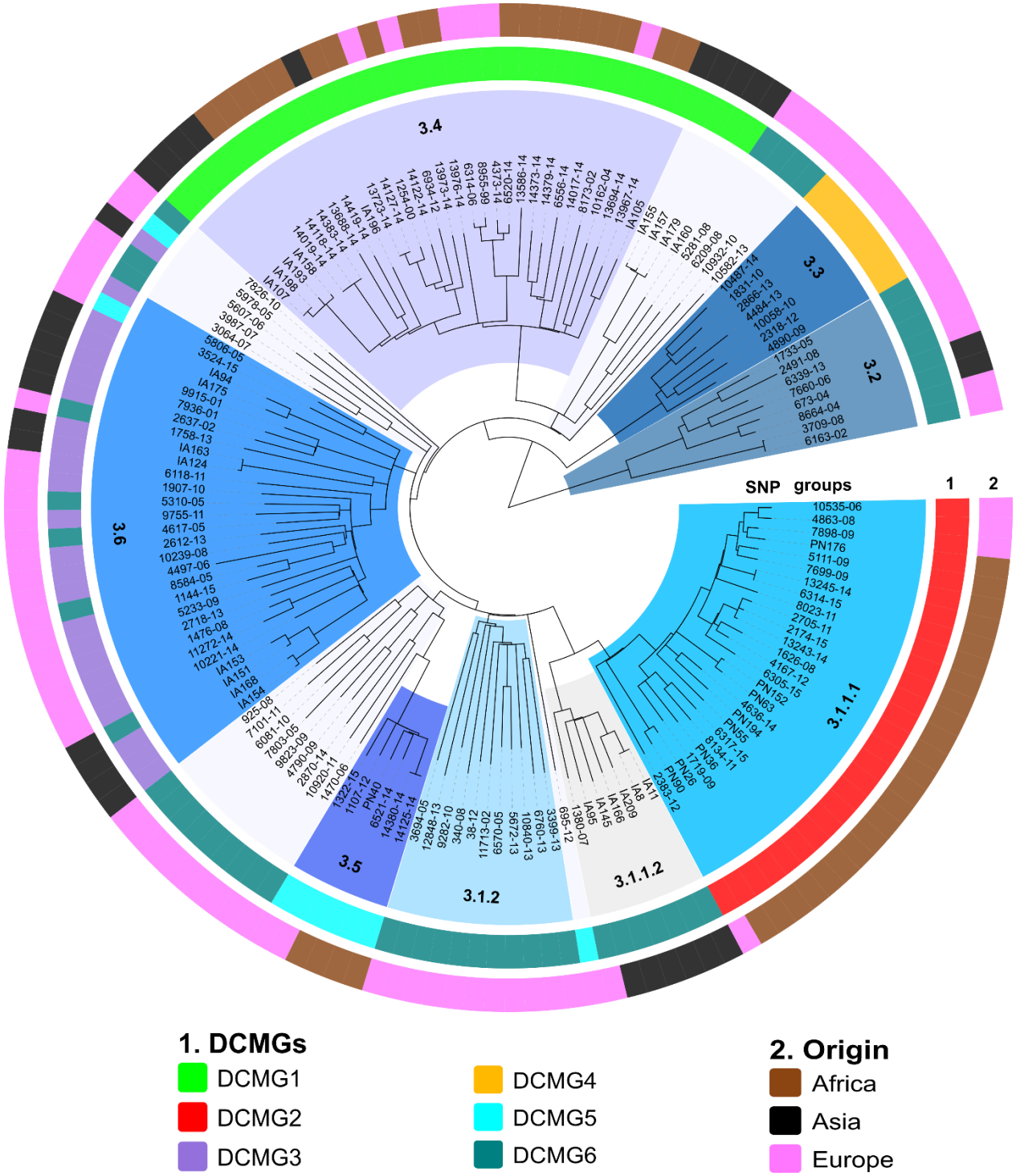


Figure 3.12: Global phylogeny of lineage 3. Maximum likelihood (ML) tree calculated based on 12,262 concatenated SNPs. The tree shows **(1)** the eight groups of L3 based on SNPs, **(2)** the six DCMGs of L3 based on 24-loci MIRU-VNTR, and **(3)** the geographical origin of the strains.

SNP-based groups were defined by eight to 102 specific SNPs occurring in many genes that were putatively associated with genes encoding for critical functions for the survival of MTBC strains, such as virulence (e.g. Rv1417 and Rv1661), transmissibility (e.g. Rv3084, Rv0197, and Rv2180c), dormancy (e.g. Rv3660c, Rv2284, and Rv1736c), and induction of immune response (e.g. Rv1283c and Rv1283c), or affect the level of resistance against SM and RMP (e.g. Rv1630 Rv3795) and pathogenesis (e.g. Rv1829) (Annexes 16–24)[122-128]. Group-specific SNPs were also found in genes with unknown functions (Annexes 16–24).

4 Discussion

The Sudan is a high TB burden country with an estimated incidence of TB cases of 77 per 100,000 population in 2017 (compared to 6.2 per 100,000 population in Germany)[14]. Particularly, in Eastern Sudan, TB represents a major health problem with estimated incidences of up to 275 per 100,000 population in the past few years[68,81]. Regardless of this disease situation, little is known about the population structure of the disease-causing MTBC strains, their genotypic drug resistance profiles, or their possible historic origin in this region. To date, only fragmentary spoligotyping data are available to infer the local population structure of infecting strains, while data on current transmission dynamics are completely missing[70,77-79]. In the same context, knowledge about the prevalence of TB in ruminants is incomplete and its epidemiology is not very well understood in Eastern Sudan.

This thesis aims to improve our knowledge of the TB epidemic in Eastern Sudan by, firstly, studying the prevalence of TB among ruminants in the two slaughterhouses of Kassala, since the role of ruminants as potential reservoirs for human TB infections is unclear in this area. To determine whether animals could be a relevant source of zoonotic TB, standard methods were employed to investigate the prevalence of lesions suggestive of TB in slaughtered animal carcasses. A very low prevalence of lesions suggestive of TB in carcasses of ruminants was observed in Kassala (0.1%, 2/2304). Additionally, no MTBC members were detected in the investigated samples (n=7). These findings suggest that the incidence of TB in ruminants is rather low. Therefore, ruminants are an unlikely source of human TB in the region.

Apart from animal TB, this thesis aimed to identify the genotypes of the most prevalent MTBC strains that cause pulmonary TB in humans in Eastern Sudan, as well as to shed light on their respective drug resistance patterns using state-of-the-art methods (e.g. WGS). The molecular characterisation of MTBC strains from TB patients revealed that L3 strains were the predominant MTBC strains causing up to 73.4% (130/177) of the investigated TB cases. Moreover, the occurrence of likely patient-to-patient transmission of L3 MDR strains was revealed by identifying clusters of genetically closely related strains by WGS (77.8%, 14/18). These findings raised the concern about the high prevalence and successful ongoing transmission of L3 strains in North/East Africa and other parts of the world and whether this lineage disposes over dedicated virulence or transmission factors that render it more successful.

To this end, the global population structure and distribution of MTBC L3 strains and their association with MDR TB were investigated in order to reveal whether there were L3 sub-types that were potentially more virulent and more transmissible (i.e. by means of clustering rates) or L3

sub-types that were associated with distinct resistance profiles (i.e. an association with MDR TB transmission chains or outbreaks). Based on a collection of 1,685 strains from 22 countries, five DCMGs/clonal complexes, in addition to a putative ancestral and basal sub-lineage, could be defined by using MIRU-VNTR typing. Furthermore, the global phylogeny was reconstructed, identifying major SNP groups and capturing the whole genomic diversity among L3 strains. Eight different SNP-based groups were defined (i.e. WGS-inferred L3 sub-lineages). Moreover, the worldwide distribution of strains of the defined DCMGs differed between Africa, Asia, and Europe. The likely reason for this phenomenon is ancient and recent human migration events as already described for strains of other MTBC lineages, such as MTBC L2 and L4[6,55,56,129,130]. Drug susceptibility data were available for 1,070 (63.5%) of the clinical strains of L3, of which 21.5% were MDR. Significant differences were noticed between the proportions of L3 MDR strains among the DCMGs. Additionally, the overall clustering rate of L3 MDR strains was 62.8% (145/231), discovered by means of MIRU-VNTR typing. This finding demonstrates that more than half of MDR TB cases caused by L3 strains are actually due to patient-to-patient transmission, especially in East Africa where 93.8% (30/32) of L3 strains were connected in MIRU-VNTR typing networks.

Together, the aim of this thesis was to enhance our understanding of the TB epidemiology in Eastern Sudan, both in humans and in animals, as well as to give insights into the MTBC L3 population structure. Detailed discussions on the respective parts of this thesis follow below.

4.1 Slaughterhouse investigation

4.1.1 TB in carcasses of ruminants slaughtered in Kassala

In this study, the prevalence of granulomatous lesions suggestive of TB in slaughtered ruminants was very low. Only 0.1% (2/2304) of the inspected carcasses were found with granulomatous lesions in Kassala during a period of six months (June to November 2014). No indication for presence of MTBC bacteria was found in samples collected from the lesions of the two carcasses suggesting that inter-species transmission of TB between domestic animals and humans is not contributing to TB incidence in humans (77 per 100,000 population)[14].

Infections with *Mycobacterium* species are frequently detected in farm animals in different geographical areas of the world where control and management measures are not in place[131]. In these areas, mycobacterial infections result in major economic consequences and significant financial losses due to decreased milk and meat production, as well as the condemnation of infected organs and carcasses[1,131]. Other concomitant losses occur due to proscription of livestock trade at national and international levels[1]. Moreover, animal reproduction and breeding are directly and adversely affected by mycobacterial infections[1,63,64].

In the Sudan, previous studies found that 0.01%–8.0% of the inspected carcasses of slaughtered ruminants had granulomatous lesions suggestive of TB. The detected lesions were either limited to specific organs such as, the lungs, liver, and lymph nodes, or were generalised miliary lesions found throughout the whole carcass[63,66,67,132]. However, studies investigating TB in dairy and fattening farms are very rare in the Sudan. Ayman et al.[133] reported that 1.9% (11/569) of the investigated dairy cattle were tuberculin skin test (TST)-positive reactors in Eastern Sudan. Generally, the estimated prevalence of lesions suggestive of TB in carcasses of slaughtered ruminants ranged from very low estimates to 16% in Africa[62,132,134,135]. *M. bovis*, *M. farcinogenes*, *M. tuberculosis*, *M. avium*, and *Nocardia* species, as well as rapidly and slowly growing NTMs, were the species most commonly isolated from the lesions[66,67,132,135]. In the present study, the samples collected from granulomatous lesions of one carcass revealed growth of an unknown *Mycobacterium* species closely related to *M. terrae*, while samples of the second carcass were extensively contaminated with bacteria and yeasts (i.e. no mycobacterial growth could be obtained from any of the cultures). This supported rather the findings of Ehsan and Nganwa[63] and Aljameel et al.[67] who reported a low prevalence of MTBC among slaughtered domestic animals in the Sudan.

Environmental risk factors and emaciation due to, for example, malnutrition, diseases, and endoparasites, can boost the susceptibility of animals to MTBC bacteria[1,63]. Herein, environmental risk factors were not investigated, hence, any solid conclusions regarding their effect on the findings of this study could not be made. Whereas emaciated/wasted animals are normally excluded during antemortem examinations due to public health regulations. However, some of these excluded animals might be slaughtered and prepared for human consumption illegally and away from the supervision of the authorities. This practice is common in the Sudan and is locally known as “Keery slaughtering”. TB-suggestive lesions could perhaps be more prevalent in carcasses of ruminants that are slaughtered illegally.

Post-mortem (i.e. after slaughtering) inspection is a very crucial technique for detection and monitoring of TB in animals[1,94,95]. It is also cost-effective and affordable for developing countries with limited resources[61,64]. However, a variable sensitivity, ranging from 28% to 64%, has been reported for post-mortem inspection[61,64,134]. These divergent sensitivity estimates could explain the low prevalence of TB-suggestive lesions observed in this study. Moreover, the study was carried out in a short period (six months); therefore, seasonal variations in the prevalence of animal TB could not be ruled out[136]. Using TST before slaughtering would enhance the accuracy of post-mortem inspections and increase the overall efficiency of TB-lesion detection since granulomatous lesions are likely to be detected in the internal organs of animals that are TST-positive, while

extended slaughterhouse-based surveys will explore whether seasonal variations in the prevalence of granulomatous lesions suggestive of TB exist or not.

MTBC bacteria were not detected in the investigated slaughterhouse samples in the present study. This means that the risk of zoonotic transmission of these bacteria via the consumption of infected animals' meat is apparently negligible or highly unlikely. Nevertheless, there is, potentially, a very slight risk of the zoonotic transmission of unknown species of mycobacteria. Moreover, when the meat is prepared through unsupervised Keery slaughtering, the risk of zoonotic transmission of MTBC and NTM bacteria is possibly higher than when the meat is prepared through supervised slaughtering. This finding also implies that producers will not bear financial losses due to decreased meat production and/or the condemnation of infected organs and carcasses. Moreover, local and international trade in livestock from Eastern Sudan will not be restricted due to animals' TB. In spite of that, the occurrence of TB in animals that are raised in farms for milk or meat and in animals that are grazing in the pasture should be monitored.

In Kassala, TB control programs, such as the test-and-slaughter policy, are not applied[63]. Accordingly, it is logical to assume that a percentage of slaughtered ruminants might have visible TB-suggestive lesions in their organs. To the contrary, none of the inspected carcasses was infected with any of the MTBC bacteria. Radostits et al.[1] indicated that the prevalence of MTBC infections is low in indigenous animals due to natural resistance and in animals that are at pasture all the year round. The absence of MTBC among animal carcasses in Kassala is possibly related to the fact that most of the slaughtered ruminants are indigenous breeds (e.g. zebu type cattle). In addition, these ruminants are at pasture year-round in search of fodder and water.

4.1.2 Strengths and limitations

A strength of this study is found in the fact that a large number of different animal species at the two slaughterhouses in Kassala were investigated. Since the animal husbandry is largely similar in Eastern Sudan, findings of this study could be generalised to the same animal species slaughtered and prepared for human consumption in slaughterhouses in other areas of Eastern Sudan (e.g. El-Gadarif). In contrast, one of the limitations is that no dairy farms were investigated, therefore, TB prevalence could not be compared between animals raised for meat and animals raised solely for milk production. Additionally, the study was conducted in a short period (only six months). This could have resulted in the underestimation of TB prevalence due to seasonal variations. Only post-mortem inspection was used as the means for investigating TB. However, tubercle granulomatous lesions are small in size during the early stages of the disease and might not be visible with the naked eye.

4.1.3 Conclusions and outlook

In conclusion, there was no evidence for the presence of MTBC bacteria in the carcasses of slaughtered ruminants in Kassala. This indicates that the prevalence and incidence of MTBC infections in animals are seemingly very low. Therefore, the zoonotic transmission of MTBC bacteria is unlikely. One of the carcasses that were found with typical TB-suggestive lesions was infected with an unknown *Mycobacterium* species. This strain could possibly be pathogenic to humans. To develop a better understanding of the epidemiology of MTBC infections in animals in the Sudan, extended slaughterhouse-based surveys combined with TST testing of animals pre-slaughtering are warranted. Surveys should also include animals that are raised for milk production.

4.2 TB diagnostics and epidemiology in humans in Eastern Sudan

Eastern Sudan suffers from a high-burden of human TB[14,68,81]. Diagnostic assays, such as culture and molecular techniques, testing for the presence of MTBC bacteria in clinical sputum specimens are not routinely applied[82]. Rather, sputum samples are consistently tested by smear microscopy, which has variable sensitivity and specificity aside from its difficulties in result interpretation[82,137].

By using state-of-the-art laboratory techniques for phenotypic and genotypic DST and for MTBC genotyping, this study could show that the positive predictive value of smear microscopy in Eastern Sudan was as high as 90% and that TB patients are predominantly infected with *M. tuberculosis* strains that belong to L3 (Delhi/CAS). In addition, drug resistance to at least one first line drug and MDR TB cases occurred almost exclusively among L3 strains, underlining the epidemiological importance of strains of this MTBC lineage for the TB epidemic and for regional TB control programs in Eastern Sudan. The high clustering rate and the fact that any drug resistance and MDR genotypes are mostly found among L3 strains further suggest that recent transmission might be fuelled by drug resistant strains through limited diagnosis and infection control capacities.

4.2.1 Positive predictive value of smear microscopy

To investigate the positive predictive value of smear microscopy in Eastern Sudan, all collected sputum samples (n=383) were cultured into liquid and onto solid cultures at the NRL, Borstel, Germany. Available culture-negative and contaminated samples were further tested using qPCR. Approximately, 10% of the samples that were smear-positive in the investigated TB centers in the Sudan did not actually have any evidence of MTBC or NTM strains when culture or qPCR were used as the gold standards. This showed that the positive predictive value of smear microscopy

in Eastern Sudan was 90% for the detection of mycobacteria, regardless of them being MTBC or NTM. Furthermore, the positive predictive value for the detection of MTBC strains alone was 84.4%. There was a disagreement between the results of smear scores determined at local centers in the Sudan and at the NRL in Germany. One-third of the samples that were smear-positive in the Sudan did not contain visible AFB at the NRL in Germany but were culture- or qPCR-positive.

In countries with limited resources, TB diagnosis largely relies on clinical findings, chest X-ray, and smear microscopy[12,14,138]. Culture and molecular tests are often not available in routine practice, especially in peripheral or remote TB centers[12,14]. In Eastern Sudan, clinical findings, chest X-ray, and smear microscopy are the only means of TB diagnosis[82]. Rieder et al.[137] found that the prevalence of TB in low-income countries was between 10% and 15% among self-presenting presumptive TB patients by using smear microscopy. This means, when the specificity of smear microscopy is as high as 98%, the positive predictive value will be in the range of 75%–91% in the context of passive case finding[137]. In Africa and Asia, reported positive predictive values were noticeably different. In Ethiopia, Kenya, and Nigeria, as well as in India positive predictive values ranged from 60% to 98%[139-143]. TB prevalence and the accuracy of smear microscopy are the main reasons behind these noticeable differences between positive predictive values in Africa and Asia. Unfortunately, neither prevalence nor specificity were investigated herein. Therefore, any firm conclusions regarding the underlying cause of the observed positive predictive value of smear microscopy for MTBC detection in Eastern Sudan could not be ascertained.

In general, the diagnostic accuracy of smear microscopy for the detection of TB cases is affected by many factors, where faulty microscopy and the presence of NTMs are the most important[99,144]. In this study, the sub-optimal diagnostic accuracy of smear microscopy could possibly and equally be ascribed to smear reading and staining errors and to the presence of NTMs. All, except one, of the culture- and qPCR-negative sputum samples (n=33) were smear-negative, as there were no visible AFB on repeat examination at the NRL. Smear reading and staining errors are best addressed through adequate training, regular supervision, and national external quality assurance schemes[99,144]. Xpert MTB/RIF, other DNA-based techniques, or culture can discriminate between MTBC and NTM strains[145].

The sputum sample of one in every seven patients diagnosed with smear-positive pulmonary TB at the investigated local TB centers in the Sudan did not actually have any evidence of the presence of MTBC bacteria when using culture or qPCR at the NRL in Germany. Regardless, these patients were initiated on TB therapy with potentially toxic drugs that may result in many undesirable adverse effects (e.g. fibrosis or lung cancer)[146,147]. These side effects may, in turn,

lead to the deterioration of general health or even to death[146,147]. However, in low-resource settings with limited diagnostics and treatment options there is little one can offer patients with potentially fatal lung disease, with the exception of acute pneumonia and TB. This needs to be compared to the fact that individuals receiving treatment for TB and their families will experience significant economic costs associated with the treatment and loss of income[147,148]. Costs are also accrued by the health care system[147,148].

Routine diagnostic tests usually test for the presence/absence of MTBC bacteria in patients' sputum specimens only. However, knowing the genotype of the infecting MTBC strain can also be of great importance for the treatment and control's success. For instance, certain genotypes are associated with MDR TB in different world settings or may expose a higher transmission risk due to increased virulence[55,149].

4.2.2 MTBC population structure in Eastern Sudan

MTBC bacteria are classified into different species and lineages/sub-lineages based on genetic characteristics[9]. This genomic diversity can influence pathogenesis and the acquisition of anti-TB drug resistance[57]. Moreover, some lineages/sub-lineages tend to harbour a high number of drug resistant strains[55,149]. Therefore, molecular surveillance approaches are key to investigating the effectiveness of NTCPs.

By using different molecular genotyping methods, this study showed that L3 strains play a major role in TB epidemiology in Eastern Sudan, infecting around 75% of all TB patients who provided samples. Previous studies did report strains that belong to different spoligotypes of L3 in the Sudan but they, however, had lower proportions (i.e. from 13% to 56%)[60,70,77-79]. Furthermore, MTBC L3 strains were reported in East Africa (e.g. Ethiopia), as well as in West and South Asia (e.g. Saudi Arabia, Pakistan, and India)[60,150-155]. Nevertheless, the proportions of L3 strains were lower than 75% (e.g. 38.8% in Ethiopia and 29.4% in Saudi Arabia, and 61% in Pakistan) except in India, where proportions reached up to 84%[151,154,156,157]. The regional expansion of L3 strains in Eastern Sudan could possibly be attributable to both recent transmission (i.e. a high clustering rate) and its' evolutionary history with multiple historical introductions from other East-African countries and West or South Asia (reflected by the diversity of detected L3 genotypes)[6,60,118,153,155,158].

The following events might serve as an explanation for the MTBC population structure in Eastern Sudan: I) the migration of people from Saudi Arabia, Ethiopia, and Eritrea; II) the arrival of Beja people; and III) ancient relations between the people of the Sudan and India. There has been an extensive legal and illegal transboundary movement of people from Saudi Arabia, Ethiopia, and

Eritrea to the Sudan[159]. Due to ethnic warfare, many Rashaida people migrated from Saudi Arabia to Eastern Sudan around the year 1850[160,161]. Now, Rashaida move constantly and illegally between Sudan, Eritrea, and North-Western Ethiopia[160,161]. Moreover, the largest number of foreigners in the Sudan has been from Ethiopia and Eritrea[59,162]. These people are either residents (i.e. they live in camps or within the Sudanese local community) and/or are transiting through the country (e.g. to Europe or USA)[159]. This transboundary movement probably play(ed) a major role in TB epidemiology and dynamics in the Sudan. An example of the transmission of MTBC L3 strains beyond international borders is the detection of strains of this lineage in Ireland, Italy, the Netherlands, and Denmark among migrants from East-African countries (e.g. Somalia, Ethiopia, and Eritrea)[119-121]. Additionally, Beja people, who live in Eastern Sudan, claim that they originated from India, which might be one source for the introduction of particular L3 strains into Eastern Sudan[163,164]. Furthermore, the historical relations between the Sudan and India started as early as 5,000 years ago[163,164]. This ancient interaction might be another possible source for the introduction of L3 strains into the Sudan. Further, it is intriguing to assume that human genetics might be an important factor influencing the susceptibility of people from Eastern Sudan to L3 strains and also leading to high transmission rates. Many polymorphisms in the human homolog of the natural-resistance-associated macrophage protein 1 (NRAMP1) gene have been found to be associated with an increased risk for TB in people from East and West Africa[165].

The findings of this study clearly show that the role of MTBC L1, L2, and L4 strains in TB epidemiology in Eastern Sudan is limited, as only 2.8%, 1.1%, and 22.5% of the MTBC strains belong to these lineages, respectively. This is in agreement with the findings of Chihota et al.[60], Sharaf Eldin et al.[70], Sharaf Eldin et al.[77], Eldirdery et al.[78], and Khalid et al.[79], who showed the presence of strains of these lineages in the Sudan by spoligotyping. In this study, among the detected L4 strains, some were classified as belonging to LAM, S-type, Sudan H37Rv-like, UgandaII, Haarlem, Cameroon, X-type, and TUR sub-lineages. This is attributed to the combined usage of different molecular methods (i.e. MIRU-VNTR typing and WGS) which have more discriminative power than spoligotyping[50,80]. Strains belonging to different sub-lineages of L4 have been found in the Sudan's neighbouring countries, in East-African countries, and worldwide[56,166-169]. Additionally, the few numbers of L1 and L2 strains reported in this study are in line with the findings of Chihota et al.[60], Sharaf Eldin et al.[70], Sharaf Eldin et al.[77]. These two lineages occur more frequently in East Asia and Eastern Europe, and some clones of L2 are particularly associated with MDR and XDR TB outbreaks in Eastern Europe[4,8,55,149,158].

The limited role of L1, L2, and L4 in TB epidemiology in Eastern Sudan could possibly be due to a lower virulence, as they may not adapt as well to local human hosts. To clarify this point, host-

pathogen interaction studies should be conducted. Another possible explanation could perhaps be that L1, L2, and L4 did not have enough time to spread in the Sudan. Brynildsrud et al.[129] and Rutaihwa et al.[130] indicated that European colonial expansion introduced strains of these MTBC lineages into Africa, while L3 strains were likely introduced into Sudan from ancient times.

Knowing the genotype of MTBC strains is highly significant for identifying the origin of the disease, as well as for better understanding of its epidemiology and dynamics[46]. Equally important is acquiring information regarding the DST patterns of the MTBC strains, since DST enables the prescription/administration of the most relevant and effective anti-TB therapy[16,37-39]. Doubtless, DST-guided anti-TB therapy is the key to the success of the treatment and control of TB. Furthermore, an investigation of resistance rates gives insight into MDR TB trends and helps to identify regions where targeted actions should be concentrated.

4.2.3 MDR TB in Eastern Sudan

NTCPs' efforts to treat and control TB and prevent its further spreading have been challenged by the emergence of drug resistant MTBC strains in different world settings[11,170]. MDR and XDR TB outbreaks have had massively devastating consequences on the progress towards TB eradication[4,11,170].

In this study, 22.3% (39/175) of the MTBC strains investigated have resistance to at least one of the five tested first line anti-TB drugs, with 10.3% (18/175) MDR and 1.1% (2/175) RR strains. As the Sudan is a resource-limited country, TB diagnosis procedures are not optimal[82]. In peripheral TB centers, neither infrastructure nor sufficient financial support exists to enable pDST or gDST[82]. Consequently, TB due to drug resistant bacteria (e.g. MDR strains) often goes undetected and is not correctly treated[82]. This failure/difficulty to detect and treat MDR TB and other forms of drug resistant TB cases could perhaps explain the observed high proportions of any resistances and MDR in this study. Moreover, it may increase the number of MDR TB cases, as well as lead to the emergence of XDR TB, over time in Eastern Sudan and in the whole country. Furthermore, defaulting treatment and/or the irregular taking of anti-TB drugs could be another possible explanation for anti-TB drug resistances noticed in this study. Studies of Abu Shanab[171] and Ali and Prins[172] reported that 14%–57% of patients default TB treatment. Reasons for defaulting treatment include the long period of drug administration (i.e. at least six months) and the side effects of these drugs, as well as social stigma, lack of motivation, poor awareness, and health staff's poor approach[171]. In Eastern Sudan, socioeconomic deprivation, poverty, and low standards of living are most likely among the main causes of acquiring MDR and non-MDR TB parallel to nutritional deficiencies and ignorance and unawareness. The observed high proportion

of any drug resistance in this study is less than the resistance rates reported in previous studies in the Sudan (39% to 66.5%), while the noted proportion of MDR strains is in range of previous reports (6% to 21.7%)[69-75]. Nevertheless, the WHO showed that MDR/RR rates were 2.9% in new and 13% in previously treated TB cases in 2017 in the whole country[14]. All of the available data might suggest a higher rate of MDR TB compared to the recent WHO estimates.

The heterogeneity of MDR strain reports in the Sudan could be attributed to the sample size investigated in each study and/or to the laboratory technique used for pDST. Abdul-Aziz et al.[72], Khalid et al.[73], and Sharaf Eldin et al.[77] investigated between 53 and 90 samples from Khartoum and Kassala, while the other studies investigated 100–235 samples from Khartoum, Gazira, Kassala, Port Sudan, and camps for displaced people, whereas, for pDST, only one study used the BACTEC MGIT 960[69-71,74,75]. In East Africa, the proportions of MDR strains ranged from 0.2%–31.4%, while for the whole continent of Africa, including MDR TB hotspots such as South Africa, MDR TB was between 1.2%–11.7%[173-179]. These variations could be ascribed to differences in the efficiency of TB control and MDR transmission in each investigated area or country. Some geographical areas/countries characterised by high numbers of drug resistant TB and are even listed in the high MDR TB burden countries[12,14]. Furthermore, the prevalence of HIV/AIDS is another important factor to be considered, as it is high across sub-Saharan Africa[180]. The overall quality of health systems and the performance of NTCPs could be added to the factors that influence the prevalence of resistant TB. A well-performing NTCP is able to control the spreading of MDR TB and, therefore, low number of such cases exist in the community. In the low- and middle-income settings of sub-Saharan Africa, male sex, HIV, diabetes mellitus, a history of known contact with TB, alcohol use, malnutrition, and low socioeconomic status are the main predictors of resistant TB[4,178].

In this study, neither pDST nor gDST showed any resistances or mutations to the tested second line anti-TB drugs. The question of whether XDR TB does not truly occur in Eastern Sudan or the failure to detect it is due to inadequate sampling remains unanswered. The WHO's recent reports indicated that two XDR cases were found and confirmed using laboratory tests during the past two years in the Sudan[12,14]. Nevertheless, neither one of the two XDR TB cases started on treatment[12,14]. To better understand the epidemiology and dynamics of MDR and XDR TB in Eastern Sudan, adopting continuous surveillance is necessary. This will only be meaningful if pDST or gDST and systematic data collection and analyses are conducted routinely.

In this study, 10.3% (18/175) of the confirmed TB cases were caused by MDR strains and another 12.0% (21/175) of the cases were due to infection with non-MDR strains. Due to the lack of diagnostic capacity, these cases were falsely considered to be due to infection with susceptible

MTBC strains and, hence, were treated with first line drugs, expecting patients to become smear-negative in two to three months[70,76]. If they remained smear-positive, the first-line drugs were replaced by a second line regimen (i.e. MDR TB treatment), posing a high risk of further resistance development and MDR TB transmission[70,76]. Both decisions, to give the first line drugs and to switch to second line drugs, were based on assumptions and were not guided or supported by laboratory testing. This may lead to a high risk of treatment failure, the emergence of resistance, and the further spreading of MDR strains, notably in the hospital setting and within the community in case patients are treated in ambulatory care. In addition, as the treatment was unnecessarily lengthy in the absence of proper drug choice through laboratory testing, it generated additional financial costs for both patients and the healthcare system[147,148].

A number of measures have been proposed to counteract the emergence of drug resistant MTBC strains such as reducing time-to-diagnosis or shifting resources to more comprehensive second-line pDSTs[181]. WGS is a promising tool for genome-based resistance prediction yet, currently, it is only available for specialised TB centers in high-income countries[181]. Further, the transition to WGS as a routine diagnostic practice will be largely dependent on the availability of convenient interpretation tools in the future[182,183]. However, the data obtained here showed that WGS has a high sensitivity and specificity for predicting resistance to all first line drugs in the analysed strains, except for EMB. Herein, the molecular analysis rather allowed for the correction of false-negative EMB results for strains harbouring *embB* mutations. Variations of the EMB MIC around the defined breakpoint of 5.0 µg/ml are known to result in a low reproducibility of phenotypic results for strains with *embB* mutations[106].

Furthermore, WGS revealed that the MDR and non-MDR strains developed mutations associated with resistance to all first line drugs. However, no mutations for second line drugs were detected, so far leaving a therapeutic option to control the MDR TB epidemic in Eastern Sudan with improved TB control measures (e.g. tailored therapy). WGS further illustrated a clear domination of the following mutations that mediate resistances against first line drugs: the Lys43Arg in the *rspL* gene (SM), the Ser315Thr in the *katG* gene (INH), the His445Tyr and Ser450Leu in the *rpoB* gene (RMP), and the Met306Ile in the *embB* gene (EMB). Other observed mutations cause resistance against SM in the *rrs* and the *gidB* genes, as well as against INH in the intergenic region of the *fabG*1 gene. Resistances associated with these mutations have been demonstrated in the Sudan and East Africa (e.g. Ethiopia) as well[71-73,184-187].

Proper TB treatment is very important to achieve a cure and to avoid the emergence of drug resistances and has been the cornerstone for successful TB control[16,17,37-39]. However, undiagnosed and/or ineffectively treated TB cases (susceptible or drug resistant TB) foster the

spread of MTBC pathogens in the community, causing additional morbidities and mortalities[188]. Therefore, identifying and stopping ongoing chains of TB transmission are greatly significant to optimise control measures.

4.2.4 MTBC bacteria transmission chains and outbreak

The clustering of MTBC strains is an indication of recent transmission chains and putative ongoing outbreaks[51,189]. These outbreaks can be identified by different molecular genotyping methods, e.g. MIRU-VNTR typing and WGS[11,51,190].

This is the first exploratory study elucidating the transmission dynamics of susceptible and drug resistant MTBC strains in the Sudan. Overall, 66.1% (n=117) and 46.8% (n=81) of the MTBC strains are associated with transmission events by using MIRU-VNTR typing and WGS, respectively. This finding clearly shows that a significant portion of the incident TB cases in Eastern Sudan is mainly due to person-to-person transmission rather than to the reactivation or progression of latent TB to active TB. Comparable clustering rates of MTBC strains have been reported in other regions in Africa (e.g. Ethiopia [63.8%] and South Africa [67%])[155,191]. However, in Asia, higher clustering rates were reported (e.g. Saudi Arabia [86.4%])[153,192]. On the one hand, several factors, such as varying living conditions/environments (e.g. rural or urban setting), demographic differences (e.g. age and sex), health status (e.g. no disease or HIV and diabetes), and personal behaviours (e.g. alcohol and smoking), could possibly cause differences in the rates of the recent transmission of MTBC strains. On the other hand, the reactivation and/or progression of latent TB to active TB also contributes significantly to the TB burden in Eastern Sudan and, hence, should be targeted for TB control, since 33.9% by MIRU-VNTR typing and 53.2% by WGS of the MTBC strains are unique strains.

In this study, it appears that L3 is the main driving force behind the MTBC outbreaks and recent transmission chains among the study's individuals in Eastern Sudan. Precisely, 77% (100/130) and 53.1% (68/128) of the L3 strains are related to each other by MIRU-VNTR typing (20 clusters) and WGS (24 clusters), respectively. The high clustering rate along with the clear dominance of L3 (73.4%) in Eastern Sudan insinuates that it is a greatly successful MTBC lineage and arguably the most significant player in the epidemiology of TB in this region. Moreover, strains of L3 are perhaps the main cause of TB in the whole country, since the findings of this study are concordant with the findings of a previous molecular epidemiological study by Sharaf Eldin et al.[70]. However, to better understand the TB epidemiology in the Sudan and to enable the making of a concrete conclusion, more molecular epidemiological studies exploring the genotypic diversity of MTBC and covering the whole country are needed. The high frequency of L3 strains together with the

high clustering rate noted herein mirror the findings reported in East Africa (Ethiopia, Djibouti, Kenya, and Tanzania)[60,154,155,167,190,193]. This likely indicates that L3 is endemic in Eastern Sudan and East Africa, thus, emphasises the epidemiological importance of L3 at a local level (Eastern Sudan) and on a regional scale (East Africa). It also highlights the necessity for a targeted surveillance and control action.

Aside from its high occurrence and clustering rates, almost all of the anti-TB drug resistances occurred among the strains of L3 (87.2%, 34/39). WGS-based clustering was detected amongst 77.8% (14/18) of the L3 MDR strains. This finding could partly be explained by delayed TB diagnosis and/or inadequate treatment, as an improper drug choice raises the probability of developing drug resistant TB and, subsequently, the risk of its transmission and dissemination (discussed above). Additionally, deprivation, low standards of living, and the socioeconomic status of people are other incriminated reasons to account for the high transmission of MDR TB in Eastern Sudan. Hamed et al.[194] showed that 42% of the population in Eastern Sudan live in households that are below the food poverty line. For this reason, some TB patients, particularly those who live in remote villages, are not able to travel multiple times per week/month in order to receive proper health care and treatment. Moreover, the habit of sharing insufficiently ventilated rooms is common among these poor communities. Therefore, TB can easily spread between the individuals of the household because of the close proximity to the index case and the prolonged duration of contact. This notion is supported by the findings of Oren et al.[189] who were able to prove an obvious association between low socioeconomic status and TB transmission. Additionally, host- and/or pathogen-related genetic factors could also explain the observed high rate of transmission. Strains of L3 and the people of Eastern Sudan have co-evolved for a long time. Consequently, strains of L3 acquired enhanced fitness traits for higher rates of transmissibility and infectivity. A similar observation was made for L2 strains in Eastern Europe[55,149]. In Ethiopia, up to 66.7% of the MDR strains were clustered, indicating ongoing recent transmission, whereas in South Africa, a country that is an MDR TB 'hotspot', the transmission of phenotypically drug resistant strains was noticed in up to 31% of the cases[190,193,195,196]. Overall, Kompala et al.[188] and WHO[197] reported that the global MDR TB outbreak is largely driven by an ongoing, unrecognised primary transmission which accounted for up to 80% of MDR TB cases in resource-limited settings around the year 2010.

Taken together, TB transmission is a major public health concern in Eastern Sudan and, most likely, in the whole country. This warrants an urgent and meaningful intervention to control the community's transmission and outbreaks of susceptible and, especially, MDR MTBC strains. Moreover, the data obtained in this study indicates the relevance of introducing multi-site large-

scale surveillance in the region and, possibly, in the whole country, employing molecular genotyping tools. Integrated molecular epidemiological surveillance would allow for the identification of TB outbreak 'hotspots' and will support the designing and creation of the most impactful interventions, such as active case-finding and contact tracing, to interrupt any ongoing outbreaks.

4.2.5 Strengths and limitations

This study has several strengths. It was a multi-site study conducted in three public hospitals in Eastern Sudan, rendering the findings generalisable to other similar settings. Cultures and molecular diagnostics were performed in an NRL in a high-resource setting. The combination of molecular diagnostics along with one liquid and two solid cultures reduced the likelihood of misclassifying a sample as negative. In addition, both a pDST-using culture and a gDST-using WGS were carried out to investigate the drug susceptibility pattern of MTBC strains. The combined usage of spoligotyping, MIRU-VNTR, and WGS to depict the genetic diversity of the MTBC strains, as well as the direct transmission from person-to-person is one of the main strengths of this study.

The limitations of the study include a lack of data on the sensitivity and specificity of smear microscopy in this setting and unknown TB prevalence among individuals presenting for investigations. Furthermore, prolonged transit time might have resulted in some samples being falsely classified as culture-negative. However, this was addressed by investigating culture-negative samples with molecular methods. Clinical data such as HIV status, chest radiography findings, and treatment outcomes were not available. Mining of such data is difficult and challenging in Eastern Sudan. In addition to the social disgrace of being infected with TB, the culture surrounding it makes TB patients reluctant and unwilling to give information about themselves, especially if guardians need to provide that information. Two samples from each patient were pooled. If one of these two samples was smear-negative and the other one was a scanty smear-positive, the pooled sample might be classified as smear-negative at the NRL.

4.2.6 Conclusions and outlook

In summary, the positive predictive value of smear microscopy reported in Eastern Sudan indicates that a proportion of the individuals recruited in this study received unnecessary treatment. Half of the false-positive results were due to reading and staining errors, while the other half was due to the presence of NTMs. The former needs to be addressed through quality assurance schemes, training, and supervision, and the latter would require more specific confirmatory tests, such as the Xpert MTB/RIF or other molecular diagnostics or cultures, to avoid possible overtreatment[198].

The obtained molecular typing data confirmed that MTBC L3 strains are the major cause of pulmonary TB cases in Eastern Sudan. Additionally, the vast majority of drug-resistant TB cases, including MDR TB cases, was due to infection with L3 strains. The findings further confirmed that more than three-quarters of MDR TB cases were through patient-to-patient transmission. Poverty and weaknesses in TB control and a lack of suitable diagnostics are potential factors that led to the emergence of MDR MTBC transmission in the country. So far, MTBC strains that are resistant to the tested second-line anti-TB drugs were not found; however, monitoring the magnitude and trends of anti-TB drug resistance is crucial. The expansion of MDR strains needs to be carefully traced in the following years and the implementation of rapid molecular DSTs should be considered.

In a wider context, the data obtained herein and in other studies demonstrate that L3 strains are common in the Sudan, the Sudan's neighbour Ethiopia, and East Africa, as well as West and South Asia[4,60,151-155,158,167,190,193]. Likewise, L3 has been frequently isolated and characterised among immigrants in the Netherlands, Denmark, and Ireland[120,121]. This wide distribution of L3 across three continents highlights its epidemiological importance and justifies investigating its genetic diversity to gain insight into its patho-biological traits, resistance mechanisms, transmission capacities, and host adaptations. The global population structure of L3 stains is discussed in the next section.

4.3 Insights into MTBC lineage 3 population structure

The ability of MTBC lineages/sub-lineages to spread among different host populations and across geographical regions varies greatly[4,8,55,56]. Some lineages/sub-lineages are restricted in their spatial distribution (i.e. specialists), while others are geographically widespread and are supposed to be adapted to different host populations (i.e. generalists)[4,56]. Therefore, the MTBC lineage-specific genetic background seems to encode for particular virulence and drug resistance traits[8,55,149]. While all other modern MTBC lineages (L2 and L4) are well investigated, L3 still lacks a robust sub-lineage classification, and data on its geographical distribution and associations to MDR TB are only sparsely available.

By using 24-loci MIRU-VNTR (n=1,685), the data obtained in this study showed that L3 is divided into five DCMGs or clonal complexes in addition to one putative ancestral and basal DCMG. Moreover, based on WGS data (n=159), L3 strains could be classified into eight SNP groups, which is more than the number of SNP groups reported previously[57]. Each of the eight SNP groups has specific mutations (8-102 SNPs). Sub-lineage identification based on the two techniques (i.e. MIRU-VNTR and WGS) was mostly in concordance, thus, providing a robust phylogenetic

classification scheme for future studies. Indeed, the obtained data in this work allowed for the capturing of detailed genetic diversity (i.e. L3 strains are genetically diverse).

Approximately half of the TB cases caused by strains of MTBC L3 are apparently due to patient-to-patient transmission since 53.9% of the analysed strains were clustered (i.e. they shared identical 24-loci MIRU-VNTR patterns). Furthermore, of the strains that had pDST data available, 21.5% were MDR, and the proportions of MDR strains differed significantly between the six DCMGs. The findings of this study confirmed the South-Asian origin of L3, as 52.4% of the strains in this area belonged to DCMG6, but the prevalence of L3 strains in Africa and Europe embodies international human migration over time. Therefore, this study reports important and valuable epidemiological information for regional/international TB surveillance and control programs.

4.3.1 Sub-lineages of MTBC lineage 3

Strains of L3 harbour the CAS-specific molecular signature, namely the RD750 deletion, and belong mainly to one of four spoligotypes (CAS1-Delhi, CAS1-Kili, CAS1-variant, and CAS2)[57,59,118]. However, a high-resolution investigation of the L3 population structure has not yet been performed. To address this question, the largest collection of clinical L3 strains to date from 22 different countries was analysed in this study. The obtained data demonstrated that L3 is diverse, with at least eight WGS-confirmed larger groups and a set of less defined, more ancestral strains. Each of the SNP groups harboured specific mutations in different genes (8-102 SNPs). This genetic diversity may also lead to different patho-biological features (e.g. disease severity), resulting from adaptation to a different host population as seen by the differences in spatial distribution across the world.

Indeed, the data obtained in this work allowed for the performing of a high-resolution analysis of the SNP-based genetic diversity of L3 strains on a global scale. Herein, eight SNP groups were defined, exceeding the previously described SNP groups, including L3.1.1, L3.1.2.1, and L3.1.2.2[57]. The diversity of the analysed strains' origins in this work could possibly explain the noticed genetic diversity. Since the analysed strains originated from 22 different countries, a more detailed SNP-based picture of the global L3 population structure could be captured. Additionally, more than half of the analysed strains (63.3%, n=1,067) originated from Asian and African countries where L3 is frequently being reported (i.e. L3-endemic countries/areas). In contrast, Coll et al.[57] analysed L3 strains from only seven countries, including China, Malawi, the Netherlands, Portugal, Russia, Uganda, and the UK, in addition to strains with unknown origins. It is important to note here that, of these seven countries, only Malawi and Uganda could be considered as L3-endemic or neighbouring L3-endemic regions in Africa[4,60].

Genetic variability in MTBC strains can have an impact on the overall outcome of TB infection[199]. When an SNP (i.e. a genetic variant) occurs in a gene regulating an essential function for the survival of an MTBC strain, it might result in a metabolic cost: either reduced or enhanced fitness (e.g. less virulant or highly transmissible strains)[200]. In this study, the WGS revealed specific SNPs for each of the identified L3 groups/subgroups. The least number of detected mutations per group/subgroup was 8 while the highest number of mutations was 102. Some of these mutations were found in genes associated with various known functions. For instance, SNPs were found in the following genes: *Rv1417* and *Rv0260*c, which are known to influence the virulence, transmissibility, and pathogenesis of MTBC strains either directly or indirectly[122-128].

Moreover, SNPs were detected in the peptide transporter-encoded *Rv1283*c gene that modulates the innate immune response of macrophages infected with an MTBC strain and in the *Rv2241* gene that induces a humoral immune response in TB patients[199,201]. Mutations were also noticed in the *Rv1696*, *Rv3441*c, *Rv3918*c, *Rv0202*c, and *Rv1295* genes that could potentially be new drug targets because of their important roles in *M. tuberculosis* survival as well as in the *Rv3800*c gene which could be a useful candidate for subunit vaccine development[202-204]. Furthermore, SNPs occurred in the *Rv1630* (SM), *Rv0842* (RMP), *Rv3696*c (EMB), *Rv0010*c (PZA), and *Rv0006* (FQs) genes that appear to affect the level of resistance of *M. tuberculosis* to the respective first or second line drugs[205-207]. In the same context, the *Rv0946*c gene, which is involved in the succinylation of different metabolic enzymes in XDR strains, harboured an SNP[208]. Considering all of these mutations, it is tempting to speculate that strains of different L3 groups have differences in their patho-biological features, such as disease severity (e.g. rapidity of weight loss, association with relapse or treatment failure), as well as in the way these strains manifest as disease[54]. Strains of some sub-lineages have been reported to cause classical pulmonary TB in addition to a wide variety of extrapulmonary manifestations (e.g. DCMG1 in Ethiopia and Saudi Arabia)[154,209]. In Pakistan, L3 strains were also associated with extrapulmonary TB at a higher rate than the strains of other MTBC lineages[156].

Furthermore, polymorphisms may result in different phenotypic features for strains of different sub-lineages (e.g. growth rates onto cultures and in mice)[54]. Only two SNP groups (3.2 and 3.5) were found to harbour mutations in epitope genes, suggesting that human T-cell epitopes in the other SNP groups are widely conserved. However, polymorphisms in T-cell epitopes might lead to different pathological outcomes, such as severe lung damage and cavitation, elevating the probability or risk of transmission[54]. Generally, an unambiguous link between the observed polymorphisms and the virulence, transmissibility, pathogenesis, immune response, and antibiotic

resistance, as well as the disease's severity and presentation would require the generation of isogenic mutant strains and the utilisation of appropriate model systems.

4.3.2 Geographical distribution of DCMGs

Strains of MTBC L3 are circulating primarily in Asia and Africa around the Indian Ocean where several countries and areas are known to have a high TB burden[4,12,14]. Individual studies have shown that L3 strains are responsible for causing as high as 84% of both pulmonary and extrapulmonary TB cases in some of these countries/areas (e.g. India)[59,60,118,151,154,156,210-212]. This illustrates the significance of L3 for the TB epidemic worldwide and around the Indian Ocean in particular.

In light of the different dating scenarios for MTBC evolution, it has been speculated that the progenitor of L3 emerged many thousands of years ago (e.g. 6 to 10 thousands years) in South Asia, probably during the time of the Indus Valley Civilisation, located in the area spanning from northwest India to northeast Afghanistan[6]. Since then, the L3 bacterial population began to grow and expand parallel to the expansion of the global human population. This historic long-term expansion is likely reflected by the large diversity and star-like population structure of DCMG6 strains, which are mainly found among TB patients in India and Pakistan (i.e. in 52.4%, 119/227). Especially, DCMG6 and 4, that are associated with the WGS-based sub-lineage 3.2 and 3.3 respectively, are considered evolutionary ancestral L3 strains and are mainly found at the suggested L3 origin in South Asia.

The spreading of L3 strains beyond South Asia is likely mediated by ancient and recent human migration, mirroring the scenarios of the global population expansion of MTBC L2 and 4[6,55,56,129,130,213]. Examples of ancient human migrations that might have introduced strains of L3 into North/East Africa include contacts and trade between people from India and North/East Africa during the time of the Nile Valley Civilisation (3100–332 BC) and Indus Valley Civilisation (2600–1900 BC) (i.e. Nubian and Indian Kingdoms) through to the time of the Axumite Empire of Ethiopia (500 BC to 700 AD) and the Funj Sultanate of the Sudan (1504–1821 AD)[163,164,214-218]. These events might have led to the introduction of DCMG1, which is associated with the WGS-based sub-lineage 3.4, into parts of North/East Africa. Another possible explanation for the occurrence of sub-lineage 3.4 in North/East Africa could be a more recent migration that included the relocation of a sizable Indian diaspora to the Sudan and Kenya during the second half of the nineteenth century[164]. In addition, Indian and East-African soldiers fought side by side during World War II as part of the East-African Campaign of the British army against the Italians at the Battle of Keren[164,217]. These ancient and recent migrations were either via Mesopotamia (i.e. West

Asia) or the Indian Ocean. It is also important to mention that there have been long historic relations between Mesopotamia and the Indus Valley Civilisation[219,220].

Apart from being a historic route for trade caravans from India to Africa and Europe, West Asia has witnessed an extensive movement of regular and irregular migrants from low-income Asian and African countries since the first half of the twentieth century[153,209,221-223]. This extensive movement was driven by discovery of oil and Gulf Wars. This extensive movement was driven by the discovery of oil and the Gulf Wars. The discovery of oil led to an economic revolution and created many jobs, therefore attracting workers from East Africa and South Asia, whereas the Gulf Wars forced many of these workers to return to their home countries. It is possible that migrants transmitted strains of DCMG1, 3, 5, and 6 between African and Asian countries and introduced them into European countries[119-121,224]. For instance, strains of DCMG1 (1649-32 MLVA code) were found in Ethiopia, the Sudan, Iraq, and Germany, and strains of DCMG3 (1449-32, 145-62, and 3549-145 MLVA codes) have been detected in India, Pakistan, Iraq, Turkmenistan, Germany, and Ireland. In contrast, strains of DCMG2 are mainly circulating in East Africa, reflecting the continuous cross border movement of people in this region of the world. Strains of DCMG2 have also been seen outside East Africa (e.g. in Europe) probably isolated from TB patients originating from East Africa[119-121]. The detection of clustered strains with the 1064-32 MLVA code in 10 different countries including Tanzania, Kenya, Ethiopia, the Sudan, Zambia, Eswatini, India, Germany, Sweden, and Ireland confirms the transmission of DCMG2 strains beyond international borders, while strains of DCMG4 are primarily found in India and Pakistan (e.g. detection of strains with the 3667-15 and 3354-15 MLVA code in India, Pakistan, and Germany), showing the transboundary movement of people in South Asia and to Europe. Detection DCMG1-6 strains in European countries reflects international human migration (e.g. strains of the DCMG1 largest cluster with the 1557-32 MLVA code were found in Ethiopia, the Sudan, Germany, Sweden, and Ireland). Studies that reported L3 strains in Europe showed that most of these strains were detected in migrants[119-121]. Overall, this long history of international relationships is highly incriminated for the spreading of L3 strains across Asian and African countries, and in European countries.

4.3.3 Drug resistance among clinical strains of MTBC lineage 3

The emergence of MDR and XDR MTBC strains is a problem that challenges global efforts to control TB[10,11,18]. According to the WHO, many of the countries where L3 strains are most prevalent (e.g. Ethiopia, Kenya, Pakistan, and India) have been defined as high-burden MDR TB countries[12,14]. To reveal whether L3 strains are among the drivers of the high-burden of MDR TB in these countries and to demonstrate whether L3 strains are associated with the transmission of

antibiotic resistant TB, 63.5% (n=1,070) of the strains were analysed for pDST. Anti-TB drug resistances including MDR strains were detected in all of the six DCMGs in this study. Overall, anti-TB drug resistance was found in 41.5% (444/1,070) of the strains, with 21.5% (231/1,070) being MDR strains. However, individual studies reported higher proportions of MDR L3 strains among new and previously treated TB patients[151-153,156,223,225-230]. For instance, in Ethiopia and Kenya (i.e. East Africa), up to 43.5% of L3 strains were found to be MDR, whereas, in India, Pakistan, and Nepal (i.e. South Asia) the proportions of MDR L3 strains reached 48.7%[156,225-230]. In Saudi Arabia and Iraq (i.e. West Asia) the proportions of MDR strains amongst L3 were, approximately, up to one-third[151-153,223].

The proportions of L3 MDR strains among the six DCMGs differed significantly in this study. MDR L3 strains were more likely to occur in DCMG1–4 and 6 than in DCMG5 (P-value of ≤0.024). Moreover, a proportion of 58.9% (136/231) of the MDR L3 strains were passed through transmission networks by means of MIRU-VNTR typing. Hence, strains of L3 are likely capable of driving a global MDR TB epidemic (e.g. 72.7% of the strains of DCMG1 [4534-32 MLVA code] found in the Sudan, Germany, Sweden, and Ireland and 27.2% of the strains of DCMG2 cluster with the MLVA code 1064-32 detected in 10 different countries on three continents were MDR). East Africa can be considered a hotspot region for transmission of MDR L3 strains as 81.3% (26/32) of the MDR strains from this region are clustered by MIRU-VNTR tying. This is supported by the findings of a previous study where it was seen that the transboundary movement of Somali MDR TB patients played a major role in MDR TB transmission in East Africa and beyond[231]. Some studies have shown that 26.3% of L3 MDR strains are clustered in Ethiopia and Kenya[226,229]. Poverty, overcrowding, and ignorance could be the leading factors for drug resistant TB in L3-endemic countries. To effectively control MDR and other forms of drug resistant TB, including cases that are caused by MDR L3 strains, the application of an external quality assurance (EQA) scheme, continuous personnel training, and the strengthening of TB diagnoses, as well as the updating of TB diagnosis and treatment guidelines should be prioritised in affected countries.

4.3.4 Strengths and limitations

The main strength of this study is the number of the analysed L3 strains (n=1,685). To date, this is the largest collection of analysed clinical strains of L3 in one study. In addition, these strains originated from various geographical regions around the world (i.e. from 22 countries in Africa, Asia, and Europe). This made it possible to capture the global picture of the spatial distribution of L3 strains at a high resolution. The second strength of this study is the utilisation of 24-loci MIRU-VNTR genotyping and WGS to investigate the genetic diversity within L3 strains. 24-loci MIRU-VNTR genotyping is the current gold standard for MTBC typing, while WGS is the most advanced

(i.e. ‘state-of-the-art’) technique currently in use in the field of molecular biology. The use of 24-loci MIRU-VNTR genotyping permitted a robust L3 DCMGs identification, whereas the use of WGS allowed for the confirmation of DCMG classification using the MIRU-VNTR technique. However, DCMGs identification (based on MIRU-VNTR typing) and sub-lineage classification (based on WGS) were not fully equal. Additionally, WGS enabled the identification of novel SNP groups/subgroups in addition to revealing sub-lineage-specific mutations.

This work has limitations as well. There is a geographical sampling bias since only samples from Africa, Asia, and Europe were analysed in the present study. This made it impossible to capture the full worldwide picture of the spatial distribution of L3 strains. Studies included in the analysis used different study designs (i.e. 14 cross-sectional, eight population-based, four MDR cohort, and four convenient studies in addition to one national drug survey and four studies with unavailable study designs). The inclusion of MDR cohort studies and a national drug surveillance in the analysis might have inflated the numbers of MDR and non-MDR strains across the identified DCMGs and, hence, led to the over-estimation of drug resistance rates. In addition, some studies used BACTEC MGIT 960 for pDST while other studies used LJ slants. This might have affected the detection sensitivity of drug resistances, including MDR. The unavailability of spoligotyping data hindered the possibility of cross-matching the identified DCMGs and SNP groups using WGS with known spoligotypes at the SITVIT database.

4.3.5 Conclusions and outlook

Indeed, the findings of this work provided a robust phylogenetic classification scheme for future studies. The use of the 24-loci MIRU-VNTR technique and WGS revealed that L3 is genetically diverse, with six DCMGs and eight SNP groups. Each of the DCMGs is mostly a distinct group based on SNPs. In addition, the SNP groups are distinguished by specific mutations in various genes. The generation of isogenic mutant L3 strains alongside the exploitation of appropriate model systems are required to study the association between some of the observed polymorphisms and pathobiological traits, e.g. the association between SNPs in the *Rv3084* gene and virulence and transmissibility. Moreover, the defined DCMGs are associated with geographical locations. DCMG2 is almost exclusively limited to North and East Africa, while DCMG4 is confined to South Asia. However, DCMG1 and 5 are circulating in North and East Africa and West Asia, but DCMG3 is mainly detected in West and South Asia. Strains of DCMG6 are found in almost all investigated regions. In South Asia, most of the L3 strains belong to DCMG6. Moreover, strains of DCMG4, that are associated with the WGS-based sub-lineage 3.3, are also limited to South Asia. Since both DCMG4 and 6 are comprised of the evolutionary ancestral L3 strains, it is most likely that L3 emerged in South Asia, rather than in Central Asia as mentioned

in the literature. The successful introduction of L3 strains into other parts of the world supposedly occurred by human migration over time. Nevertheless, studying the evolutionary scenario of the defined sub-lineages of L3 by estimating the time to the most common ancestor will expose how these distinct sub-lineages diverged from the progenitor of L3. Additionally, it will allow for the precise identification of the historic events that fuelled the dispersal of L3 strains across Africa and Asia. The results indicate that MDR strains occur among the strains of six DCMGs but with different proportions (i.e. MDR strains were more likely to occur in DCMG1–4 and 6 than in DCMG5). Further, the results suggest that more than half of the MDR TB cases caused by MDR L3 strains are probably due to patient-to-patient transmission. Therefore, strains of L3 can possibly drive a global MDR TB epidemic, especially in East Africa. Strengthening NTCPs is critical for preventing the spreading of MDR TB.

To summarise, this work showed that the zoonotic transmission of MTBC strains most likely does not happen in Eastern Sudan. Moreover, MTBC L3 strains are the most predominant MTBC strains infecting humans in Eastern Sudan. In addition, L3 strains are responsible for causing all of the MDR TB cases. Hence, L3 is clearly the most important player in the epidemiology of TB in this region. Moreover, studying the global genetic diversity of MTBC L3 by analysing strains from 22 different countries confirmed that the genetic diversity of this lineage is higher than currently expected and led to a significantly improved classification system that will serve future studies in terms of molecular drug resistance determinants while allowing insight into the evolutionary history of this regionally dominating MTBC lineage.

Zusammenfassung

Tuberkulose in Tieren und Menschen des Ostsudans und genetische Diversität von klinischen *Mycobacterium tuberculosis* Komplex Isolaten der Linie 3

Der Sudan gehört zu den Tuberkulose (TB) Hochinzidenzregionen weltweit (77 TB-Fälle pro 100.000 Einwohnern). Besonders der östliche Teil des Landes erreichte in den letzten Jahren TB-Inzidenzraten von bis zu 275 TB-Fälle pro 100.000 Einwohner. Die Prävalenz von *Mycobacterium tuberculosis* Komplex (MTBK) Stämmen in Wiederkäuern, in der Bevölkerung sowie mögliche Zoonosen sind bisher nur unzureichend untersucht. Weiterhin kann der Ost-Sudan als „Blinder Fleck“ bezeichnet werden in Bezug auf verfügbare Daten über Antibiotika Resistenzprofile oder molekulare Resistenzmarker der multiresistenten (MDR) MTBK Stämme, welche resistent gegen die beiden Antibiotika Isoniazid und Rifampicin sind. Die MTBK Linie 3, Delhi/CAS Genotyp, kommt besonders häufig im Sudan und angrenzenden Ländern vor und gilt als Hauptverursacher der TB in diesen Regionen. Trotzdem ist die regionale als auch die globale genetische Diversität und die Übertragungswahrscheinlichkeit der MTBK L3 Stämme nicht umfassend untersucht. Um diese Frage genauer zu analysieren werden in dieser Arbeit die historische Ausbreitung aber auch rezente Übertragungsketten an Hand der genetischen Diversität der Stämme analysiert. Die gewonnen Daten sollen eine detailliertere phylogenetische Struktur aufzeigen und neue Daten für zukünftige molekulare Resistenztests und Mutationsdatenbanken für Resistenzvorhersagen auf Basis einer Gesamtgenomsequenzierungen (WGS) liefern.

Zunächst wurden 2304 Schlachtkadaver von Wiederkäuern aus zwei Schlachtbetrieben in Kassala auf mögliche TB-Granulome untersucht. Darüber hinaus wurden 383 Sputum-Proben von TB-Patienten in Kassala, Port Sudan und El-Gadarif analysiert (Juni-November 2014 und Januar-Juli 2016). Alle Proben wurden im Nationalen TB-Referenzzentrum in Borstel dekontaminiert und MTBK Stämme in Flüssigmedium (BACTEC MGIT 960) und auf Festmedium (Löwenstein-Jensen (LJ) und Stonebrink) kultiviert. Eine Speziesidentifikation bei positive Kulturen erfolgte mit Line Probe Assays und ITS Gensequenzierung. Kulturen wurden weiterhin für die phänotypische Resistenztestung verwendet, MTBK-DNA diente zur molekularen Klassifizierung (Genotypisierung) der Patientenisolate. Die Genotypisierung erfolgte mittels Spoligotypisierung, 24-loci MIRU-VNTR Typisierung und WGS. Kultur negative und kontaminierte Proben wurden mit einer hauseigenen qPCR für die Detektion von mykobakterieller DNA untersucht. Um die globale Populationsstruktur von MTBK L3 Stämmen zu analysieren wurden 1,685 MTBK Isolate aus 22 Ländern aus Afrika, Asien und Europa mittels 24-loci MIRU-VNTR Typisierung analysiert. Phänotypische Resistenztests von 1,070 Stämmen wurden ausgewertet und 159 repräsentative MTBK L3 Isolate für eine WGS Analyse ausgewählt.

Es konnte gezeigt werden, dass lediglich 0.1% (2/2304) der Schlachkadaver von Wiederkäuern mögliche TB-Granulome auswiesen. Keine der beiden Proben enthielt MTBK-Stämme, in einer der Proben konnte ein Stamm mit Ähnlichkeiten zu *Mycobacterium terrae* nachgewiesen werden. Weiterhin wurden 383 Patientenisolate aus dem Ost-Sudan genotypisiert und in Bezug auf ihr Resistenzprofil untersucht. Molekulare Cluster wurden als Marker für mögliche Übertragungskosten bestimmt. Bakterielles Wachstum konnte in 51.2% (196/383) der Sputum-Proben nachgewiesen werden; 87.2% (171/196) der positiven Proben enthielten *M. tuberculosis,* 7.2% (14/196) wurden als *M. intracellulare* klassifiziert, 5.6% (11/196) waren kontaminiert und erlaubten keine eindeutige Speziesbestimmung. Mittels qPCR wurden weiterhin 75.1% (127/169) als MTBK und 4.7% (8/169) als nicht-tuberkulöse Mykobakterien (NTMs) bestimmt. Die Ergebnisse des Kulturnachweises und der qPCR kombiniert resultierten in einem positiven Vorhersagewert von 90.7% (331/365) für den mikroskopischen Nachweis von Mykobakterien in Sputum-Proben aus dem Ost-Sudan. Die Genotypisierung der verfügbaren MTBK-Isolate zeigte folgende Verteilung der unterschiedlichen Linien: 73.4% (130/177) L3, 22.5% (40/177) L4, 2.8% (5/177) L1, und 1.1% (2/177) L2. Phänotypische DST Daten waren für 96.7% (175/181) der MTBK-Stämme verfügbar, 22.3% (39/175) zeigten mindestens eine Resistenz gegen eines der Erstrangmedikamente, 10.3% (18/175) wurden als MDR eingestuft. Weiterhin waren 77.8% (n=14) der MDR-MTBK Isolate einem molekularen Cluster zugeordnet.

Die Analyse der globalen MTBK L3 Kollektion konnte sechs klonale Komplexe basierend auf der 24-loci MIRU-VNTR Typisierung definieren, Delhi/CAS Miru Gruppen (DCMGs) und acht Subgruppen basierend auf der WGS Analyse. Jeder der acht Genombasierten Subgruppen ist durch spezifische Mutationen charakterisiert. Von den MTBK L3 Stämmen mit verfügbaren phänotypischen Resistenzergebnissen wurden 21.5% als MDR klassifiziert wobei nicht alle DCMGs gleichmäßige MDR-TB Proportionen zeigten.

Die Arbeit lässt vermuten, dass MTBK-Stämme in vermeintlichen TB-Granulomen in Schlachtkadavern von Wiederkäuern in Kassala sehr selten sind und eine Übertragung von Tier auf Mensch sehr unwahrscheinlich ist. Weiterhin zeigt die Arbeit, dass spezifischere Labortests, z.B. Xpert MTB/RIF eingesetzt werden sollten, um TB-Patienten eindeutiger zu bestimmen. Einige Sputum-Proben enthielten beispielsweise keine MTBK Stämme oder NTMs. Weiterhin zeigten die Ergebnisse die Übertragung von MDR-MTBK Stämmen und dass MTBK L3 Stämme die Hauptursache einer Lungentuberkulose im Ost-Sudan sind. Armut und Schwächen im TB-Kontrollsystem sowie der der Diagnostik können als potentielle Faktoren angesehen werden, die zum Auftreten von MDR-MTBK Übertragungen im Sudan führen. Diese Fälle müssen in den

folgenden Jahren weiterhin sorgfältig überwacht werden und die Einführung von schnellen molekularen DSTs wäre wünschenswert.

In dieser Arbeit wurde eine robuste phylogenetische Klassifizierung von MTBK L3 Stämmen erstellt, welche diese Linie in sechs DCMGs und acht genombasierte Subgruppen unterteilt. Stämme der einzelnen DCMGs zeigen weiterhin unterschiedliche geografische Verteilungen. Es konnte der vermeintliche Ursprung dieser Linie in Süd-Asien bestätigt werden (repräsentiert durch die anzestrale Subgruppe L3.2) und weiterhin die Ausbreitung in andere Teile der Welt nachvollzogen werden. Historische Migrationsbewegungen aber auch rezente Übertragung haben die derzeitige Phylogenie der MTBK L3 Stämme maßgeblich beeinflusst; der Effekt der Antibiotikabehandlung ist eher moderat anzusehen, im Vergleich zu MTBK L2 Stämmen zum Beispiel, und betrifft alle DCMGs gleichermaßen mit Ausnahme von DCMG5, welcher niedrigere MDR-TB Proportionen aufweist. Trotzdem sind über die Hälfte der MDR-MTBK L3 Stämme mit einer rezenten Übertragung assoziiert, so dass dieser Mechanismus auch in Ost-Afrika eine große Bedeutung bei steigenden MDR-TB Raten spielt.

Zusammengefasst zeigt die Arbeit, dass Zoonosen keine große Rolle bei der TB-Epidemiologie spielen, MTBK L3 Stämme die Hauptverursacher der Lungentuberkulose im Ost-Sudan sind und 10% der Patienten als MDR-TB Fall eingestuft werden. Die globale Analyse der MTBK L3 Stämme aus 22 Ländern offenbarte zudem eine größere genetischen Diversität als bisher vermutet wurde und führte zu einem signifikant verbesserten Klassifizierungssystem. Dies wird in Zukunft Studien zur molekularen Analyse von Resistenzdeterminanten dienlich sein und erlaubte neue Einsichten in die Evolutionsgeschichte dieser regional dominierenden MTBK Linie.

Summary

The Sudan is a high tuberculosis (TB) burden country (i.e. 77 per 100,000 population). Specifically, the eastern part of the country is endemic with high incidence rates of TB reaching up to 275 per 100,000 population during the past decade. Prevalence of *Mycobacterium tuberculosis* complex (MTBC) strains in ruminants and in humans, as well as potential zoonotic spillovers, are only sparsely investigated. Furthermore, Eastern Sudan is a blind spot with regard to drug resistance rates and virtually no data is available on molecular drug resistance markers of circulating multi-drug resistant (MDR) MTBC strains that are resistant to at least isoniazid and rifampicin. Particularly, strains of MTBC lineage 3 (L3), i.e. Delhi/CAS genotype, are highly prevalent in the Sudan and adjacent countries and are the main causative agent of TB in those regions, however, their regional and global genetic diversity and transmission capacity are likely not fully captured yet. Thus, this thesis aims to provide a better understanding of historic and recent transmission dynamics, the genetic background of MTBC L3 strains and to draw a robust phylogenetic framework which will serve future molecular drug resistance assays and whole genome sequencing (WGS) mutation catalogues.

First, a total of 2,304 carcasses of ruminants slaughtered at the two slaughterhouses of Kassala were inspected for presence of TB-suggestive lesions (June to November 2014). Second, 383 sputum samples were collected from TB patients at Kassala, Port Sudan, and El-Gadarif teaching hospitals (June to November 2014 and January to July 2016). All collected samples were sent to the National Reference Laboratory (NRL) for Mycobacteria, Borstel in Germany, where all samples were decontaminated and cultured into MGIT liquid medium (BACTEC MGIT 960 system) and onto Löwenstein-Jensen (LJ) and Stonebrink slants. Culture positive samples were subjected to line probe assays and ITS gene sequencing for species identification. Cultures of MTBC strains were subjected to phenotypic drug susceptibility testing (pDST), and MTBC DNA was used for genotypic characterisation (genotyping) using spoligotyping, 24-loci MIRU-VNTR typing, and WGS. Culture-negative and contaminated samples were subjected to an in-house real-time PCR (qPCR) for detection of mycobacterial DNA. Finally, to get insights into MTBC L3 global population structure, a data set containing the results of 24-loci MIRU-VNTR genotyping of 1,685 strains from 22 countries across Asia, Africa, and Europe was analyzed. The pDST of 1,070 of the strains was also analyzed. A total of 159 out of the 1,685 L3 strains representing a wide global diversity were selected for a WGS analysis.

This work showed that only 0.1% (2/2304) of the carcasses of ruminants were found with TB-suggestive lesions. None of the MTBC members was isolated from the collected TB-suggestive

lesions but a slow growing *Mycobacterium* species with the closest similarity to *M. terrae* group. Furthermore, growth of organisms was obtained from 51.2% (196/383) of TB patients derived sputum samples from Eastern Sudan to identify the MTBC lineage of the infecting strains, molecular clusters as surrogate marker for recent transmission in the region, and drug resistance profiles. Of the isolated bacteria, 87.2% (171/196) were identified as *M. tuberculosis* and 7.2% (14/196) as *M. intracellulare*. The remaining 5.6% (11/196) positive cultures had more than one species of bacteria. The qPCR showed that 75.1% (127/169) of the culture-negative and contaminated samples were positive for MTBC DNA and 4.7% (8/169) for nontuberculous mycobacteria (NTM) DNA. Results of culture and qPCR together showed a positive predictive value of 90.7% (331/365) for smear microscopy for detection of mycobacteria in Eastern Sudan. Moreover, MTBC molecular analysis revealed that 73.4% (130/177) belong to L3, 22.5% (40/177) to L4, 2.8% (5/177) to L1, and 1.1% (2/177) to L2. pDST data were available for 96.7% (175/181) of the MTBC strains. A percentage of 22.3% (39/175) of the strains with pDST results available showed resistance to at least one of the tested first line antibiotics, of which 10.3% (18/175) were MDR strains. Furthermore, 77.8% (n=14) of the MDR strains were clustered. Overall MTBC L3 strains are divided into six Delhi/CAS Miru Groups (DCMGs) and into eight WGS-based groups. Each of the eight groups has specific mutations. Of the L3 strains with available pDST data, 21.5% were MDR.

This thesis suggests that the overall prevalence of MTBC strains in TB-suggestive lesions is very low in carcasses of ruminants in Kassala and zoonotic transmission of MTBC bacteria is unlikely. It further suggests the need for more specific laboratory tests such as the Xpert MTB/RIF to avoid possible overtreatment as some sputum samples did not contain MTBC bacteria and few others had only NTMs. Moreover, this work confirmed that L3 strains are responsible for causing the majority of TB cases in Eastern Sudan as well as revealed transmission of MDR MTBC strains. Poverty and weaknesses in TB control and diagnostics are potential factors that lead to the emergence of MDR MTBC transmission in the country. This needs to be carefully traced in the following years and implementation of rapid molecular DSTs should be considered. Indeed, the findings of this work provided a robust phylogenetic classification scheme for L3 strains for future studies by defining six distinct DCMGs and eight WGS-based groups. The defined DCMGs have also distinct spatial distribution. It is most likely that L3 strains emerged in South Asia (with L3.2 as the most ancestral strains) and subsequently spread successfully into other parts of the world. Historic migration waves but also recent transmission likely shaped the MTBC L3 phylogeny. However, the effect of antibiotic treatment is moderate, compared to MTBC L2 strains for instance, and it affects all defined DCMGs equally, except DCMG5 which showed a lower MDR MTBC

proportion. Nevertheless, more than half of the MDR MTBC L3 strains were related to recent transmission rendering MTBC L3 strains the possible driver of MDR TB in East Africa.

Jointly, this work showed that zoonotic spillovers of MTBC strains virtually don't occur, and MTBC L3 strains are the main causative agent of TB and are responsible of all MDR TB cases in Eastern Sudan. Moreover, the global genetic analysis of MTBC L3 strains from 22 different countries revealed that the MTBC L3's genetic diversity is higher than currently expected. This finding has significantly improved the classification system and will serve future studies on molecular drug resistance determinants. The finding further provided insights in the evolutionary history of this regionally dominating MTBC lineage.

Keywords: tuberculosis, ruminants, humans, Sudan, lineage 3

5 References

1. **Radostits O. M.*, et al.*** Veterinary Medicine: A text book of disease of cattle, sheep, pig, goat and horses. 10th ed. London, UK: Harcourt Publisher Ltd.; 2007. 2065 p.
2. **WHO.** Definitions and reporting framework for tuberculosis – 2013 revision (updated December 2014). Geneva, Switzerland: World Health Organization; 2014. Report No.: ISBN 978 92 4 150534 5 Contract No.: NLM classification: WF 360. Last Accessed on: 16.03.2018. Available from: http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf.
3. **Brites D.*, et al.*** Co-evolution of *Mycobacterium tuberculosis* and Homo sapiens. *Immunological Reviews* 264(1), 6-24, (2015).
4. **Gagneux S.** Ecology and evolution of *M. tuberculosis*. *Nat Rev Microbiol* 16(4), 202-13, (2018).
5. **Daniel T. M.** The history of tuberculosis. *Respir Med* 100(11),1862-70, (2006).
6. **Comas I.*, et al.*** Out-of-Africa migration & Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet* 45(10), 1176-82, (2013).
7. **Wirth T.*, et al.*** Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS Pathog* 4(9), e1000160, (2008).
8. **Niemann S.*, et al.*** Impact of Genetic Diversity on the Biology of *Mycobacterium tuberculosis* Complex Strains. *Microbiol Spectr* 4(6), 1-18, (2016).
9. **Brites D.*, et al.*** A new phylogenetic framework for the animal-adapted *Mycobacterium tuberculosis* complex. *Front Microbiol* 9, 2820, (2018).
10. **Dheda K.*, et al.*** Tuberculosis. *The Lancet* 387(10024), 1211-26, (2016).
11. **Dheda K.*, et al.*** The epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant, extensively drug-resistant, and incurable tuberculosis. *The Lancet Respiratory Medicine* 5(4), 291-360, (2017).
12. **WHO.** Global Tuberculosis Report. Geneva, Swizerland: WHO; 2017. Report No.: Annual Report I. Last Accessed on: 17.11.2017. Available from: http://apps.who.int/iris/bitstream-10665259366/1/9789241565516-eng.pdf?ua=1.
13. **WHO.** MDR TB Factsheet. Geneva, Swizerland: WHO; 2017. Last Accessed on: 11.01.2018. Available from: http://www.who.int/features/qa/79/en/.
14. **WHO.** Global Tuberculosis Report. Geneva, Swizerland: WHO; 2018. Report No.: Licence: CC BY-NC-SA 3.0 IGO. Last Accessed on: 17.04.2019. Available from: http://www.who.in-t/tb/publications/global_report/en/.
15. **Zumla A.*, et al.*** Tackling the tuberculosis epidemic in sub-Saharan Africa-unique opportunities arising from the second European Developing Countries Clinical Trials Partnership (EDCTP) programme 2015-2024. *Int J Infect Dis* 32, 46-9, (2015).
16. **Zumla A.*, et al.*** Tuberculosis treatment and management—an update on treatment regimens, trials, new drugs, and adjunct therapies. *Lancet Respir Med* 3(3), 220-34, (2015).
17. **Zumla A.*, et al.*** Tuberculosis. *N Engl J Med* 368(8), 745-55, (2013).
18. **Migliori G. B.*, et al.*** Review of MDR and XDR TB: global perspectives with a focus on sub-Saharan Africa. *Trop Med Int Health* 15(9), 1052-66, (2010).
19. **Shah N. S.*, et al.*** Transmission of extensively drug-resistant tuberculosis in south africa. *N Engl J Med* 376(3), 243-53, (2017).
20. **Kapwata T.*, et al.*** Spatial distribution of extensively drug-resistant tuberculosis (XDR TB) patients in KwaZulu-Natal, South Africa. *PLoS One* 12(10), e0181797, (2017).
21. **Gagneux S.*, et al.*** Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis* 7(5), 328-37, (2007).

22. **Niemann S.*, et al.*** Differentiation of clinical *Mycobacterium tuberculosis* complex isolates by *gyrB* DNA sequence polymorphism analysis. *J Clin Microbiol* 38(9), 3231-4, (2000).
23. **Niemann S.*, et al.*** Differentiation among members of the *M. tuberculosis* complex by molecular and biochemical features: evidence for two pyrazinamide-susceptible subtypes of *Mycobacterium bovis*. *J Clin Microbiol* 38(1), 152-7., (2000).
24. **Niemann S.*, et al.*** Biochemical and genetic evidence for the transfer of *M. tuberculosis* subsp. *caprae* Aranaz et al. 1999 to the species *Mycobacterium bovis* Karlson and Lessel 1970 (Approved Lists 1980) as *Mycobacterium bovis* subsp. *capra*e comb. nov. *International Journal of Systematic and Evolutionary Microbiology* 52, 433–6, (2002).
25. **El-Sayed A.*, et al.*** Molecular epidemiology of *M. bovis* in humans and cattle. *Zoonoses Public Health* 63(4), 251-64, (2016).
26. **Gordon S. V.*, et al.*** Microbe profile: *Mycobacterium tuberculosis*: humanity's deadly microbial foe. *Microbiology*, 10.1099/mic.0.000601 (2018).
27. **OIE.** 2009. *Bovine Tuberculosis*. Last Accessed on: 10.01.2018. Paris, France: OIE. Available from: http://www.cfsph.iastate.edu/Factsheets/pdfs/bovine_tuberculosis.pdf
28. **Ahmad S.** Pathogenesis, immunology, and diagnosis of latent *Mycobacterium tuberculosis* infection. *Clin Dev Immunol* 2011,814943, (2011).
29. **Cambier C. J.*, et al.*** Host evasion and exploitation schemes of *Mycobacterium tuberculosis*. *Cell* 159(7), 1497-509, (2014).
30. **Cosma C. L.*, et al.*** The secret lives of mycobacteria. *Annu Rev Microbiol* 57, 641-76, (2003).
31. **Koul A.*, et al.*** The challenge of new drug discovery for tuberculosis. *Nature* 469(7331),483-90, (2011).
32. **WHO.** Global Tuberculosis Report. Geneva, Swizerland: WHO; 2016. Report No.: Annual Report I. Last Accessed on: 14.08.2017. Available from: http://www.who.int/tb/publications/global-report/en/.
33. **CDC.** 2013. *Core Curriculum on Tuberculosis: What the Clinician Should Know*. Last Accessed on: 10.01.2018. Atlanta USA: Centers for Disease Control and Prevention, National Center for HIV STD and TB Prevention, Division of Tuberculosis Elimination. Available from: https://www.cdc.gov/tb/education/corecurr/index.htm
34. **Ryu Y. J.** Diagnosis of pulmonary tuberculosis: recent advances and diagnostic algorithms. *Tuberc Respir Dis (Seoul)* 78(2), 64-71, (2015).
35. **WHO.** Same-Day Diagnosis of Tuberculosis by Microscopy: WHO policy statement. Geneva, Switzerland: World Health Organization; 2011. Contract No.: ISBN: 978 92 4 150160 6. Last Accessed on: 11.01.2018. Available from: http://www.who.int/tb/publications/2011/tb-micros-copy_9789241501606/en/.
36. **Caulfield A. J.*, et al.*** Diagnosis of active tuberculosis disease: from microscopy to molecular techniques. *J Clin Tuberc Other Mycobact Dis* 4, 33-43, (2016).
37. **WHO.** Treatment of tuberculosis: guidelines. Fourth ed. Royce S, Grzemska M, editors. Geneva, Switzerland: WHO Press; 2010.
38. **WHO.** WHO treatment guidelines for drug resistant tuberculosis: 2016 update. Geneva, Switzerland: World Health Organization, strategy TeT; 2016. Contract No.: ISBN: 978 92 4 154963 9. Last Accessed on: 11.01.2018. Available from: http://apps.who.int/iris/handle-10665-250125.

39. **WHO.** Guidelines for treatment of drug-susceptible tuberculosis and patient care: 2017 update. Geneva, Switzerland: World Health Organization, strategy TeT; 2017. Report No.: ISBN: 978-92-4-155000-0. Last Accessed on: 11.01.2018. Available from: http://www.who.int/tb-publications/2017/dstb_guidance_2017/en/.
40. **D'Ambrosio L.*, et al.*** Delamanid and bedaquiline to treat MDR and XDR tuberculosis in children: a systematic review. *J Thorac Dis* 9(7), 2093-101, (2017).
41. **Smith T.*, et al.*** Molecular biology of drug resistance in *Mycobacterium tuberculosis*. *Curr Top Microbiol Immunol* 374, 53-80, (2013).
42. **Almeida P. E.*, et al.*** Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis*: classical and new drugs. *J Antimicrob Chemother* 66(7), (2011).
43. **Feuerriegel S.*, et al.*** PhyResSE: a web tool delineating *M. tuberculosis* antibiotic resistance and lineage from whole-genome sequencing data. *J Clin Microbiol* 53(6), (2015).
44. **Mathema B.*, et al.*** Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 19(4), 658-85, (2006).
45. **Jagielski T.*, et al.*** Current methods in the molecular typing of *Mycobacterium tuberculosis* and other mycobacteria. *Biomed Res Int* 2014, 645802, (2014).
46. **Merker M.*, et al.*** The Evolution of Strain Typing in the *Mycobacterium tuberculosis* Complex. *Adv Exp Med Biol* 1019, 43-78, (2017).
47. **Ei P. W.*, et al.*** Molecular Strain Typing of *Mycobacterium tuberculosis*: a Review of Frequently Used Methods. *J Korean Med Sci* 31(11), 1673-83, (2016).
48. **Vauterin L.*, et al.*** Integrated Databasing and Analysis. In: Stackebrandt E. (eds) Molecular Identification, Systematics, and Population Structure of Prokaryotes. Berlin, Heidelberg: Springer, Berlin, Heidelberg; 2006.
49. **Cabibbe A. M.*, et al.*** WGS of *M. tuberculosis*. *Eur Respir J* 52(5), (2018).
50. **Homolka S.*, et al.*** High resolution discrimination of clinical *M. tuberculosis* complex strains based on single nucleotide polymorphisms. *PLoSOne* 7(7), e39855, (2012).
51. **Walker T. M.*, et al.*** Whole-genome sequencing to delineate *M. tuberculosis* outbreaks: a retrospective observational study. *The Lancet Infectious Diseases* 13(2), 137-46, (2013).
52. **Kohl T. A.*, et al.*** MTBseq: a comprehensive pipeline for whole genome sequence analysis of *Mycobacterium tuberculosis* complex isolates. *PeerJ* 6, e5895, (2018).
53. **Walker T. M.*, et al.*** Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. *The Lancet Infectious Diseases* 15(10), 1193-202, (2015).
54. **Coscolla M.*, et al.*** Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Semin Immunol* 26(6), 431-44, (2014).
55. **Merker M.*, et al.*** Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat Genet* 47(3), 242-9, 2015).
56. **Stucki D.*, et al.*** *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nat Genet* 48(12), 1535-43, (2016).
57. **Coll F.*, et al.*** A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nat Commun* 5, 4812, (2014).
58. **Filliol I.*, et al.*** Global distribution of *Mycobacterium tuberculosis* spoligotypes. *Emerg Infect Dis* 8(11), 1347-9, (2002).

59. **Brudey K.*, et al.*** *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol* 6,23, (2006).
60. **Chihota V. N.*, et al.*** Geospatial distribution of *Mycobacterium tuberculosis* genotypes in Africa. *PLoS One* 13(8), e0200632, (2018).
61. **Schiller I.*, et al.*** Bovine tuberculosis: current and emerging diagnostic techniques for disease control and eradication. *Transbound Emerg Dis* 57(4), 205-20, (2010).
62. **Kassa G. M.*, et al.*** Tuberculosis in Goats and Sheep in Afar Pastoral Region of Ethiopia and Isolation of *Mycobacterium tuberculosis* from Goat. *Vet Med Int* 2012, 869146, (2012).
63. **Ehsan A.*, et al.*** Factors contributing to the transmission of bovine tuberculosis caused by *Mycobacterium bovis* and its control status in Sudan. In: Charles OT, James HS, John BK, editors. Zoonotic tuberculosis: *Mycobacterium bovis* and Other Pathogenic Mycobacteria. Third ed. USA: John Wiley & Sons, Inc.; 2014. p. 159–74.
64. **Ramos D. F.*, et al.*** Diagnosis of bovine tuberculosis. *Braz J Biol* 75(4), 830-7, (2015).
65. **Carruth L.*, et al.*** Zoonotic tuberculosis in Africa:. *The Lancet* 388(10059), 2460-1, (2016).
66. **Asil T. A.*, et al.*** Bovine tuberculosis in South darfur: an abattoir study based on microscopy and molecular detection methods. *Trop Anim Health Prod* 45(2), 469-72, (2013).
67. **Aljameel M. A.*, et al.*** Occurrence of bovine tuberculosis at Nyala abattoirs in South Darfur State, Sudan. *Rev Elev Med Vet Pays Trop* 67(2), 61-5, (2014).
68. **Abdallah T. M.*, et al.*** Epidemiology of tuberculosis in Eastern Sudan. *Asian Pacific Journal of Tropical Biomedicine* 2(12), 999-1001, (2012).
69. **Muna O. A.** Determination of the prevalence of tuberculosis With drug-resistant strains of *Mycobacterium tuberculosis* in Khartoum, Gazira and camps for displaced people, Sudan. Oslo, Norway: Faculty of Medicine, University of Oslo; 2002.
70. **Sharaf Eldin G. S.*, et al.*** Tuberculosis in Sudan: a study of *Mycobacterium tuberculosis* strain genotype and susceptibility to anti-tuberculosis drugs. *BMC Infect Dis* 11, 219, (2011).
71. **Hassan S. O.*, et al.*** Drug Resistance in *Mycobacterium tuberculosis* Isolates from Northeastern Sudan. *British J Med Medical Res* 2(3), 424-33, (2012).
72. **Abdul-Aziz A. A.*, et al.*** Multi-drug resistance tuberculosis (MDR-TB) in Kassala State, Eastern Sudan. *Trop Doct* 43(2), 66-70, (2013).
73. **Khalid F. A.*, et al.*** Tuberculosis drug resistance isolates from pulmonary tuberculosis patients, Kassala State, Sudan. *Int J Mycobacteriol* 4(1), 44-7, (2015).
74. **Eman M. M.*, et al.*** Drug Resistance Patterns of *Mycobacterium tuberculosis* Isolates from Patients with Pulmonary Tuberculosis in the Sudan. *IOSR Journal of Dental and Medical Sciences* 14(8), 17-19, (2015).
75. **Eldirdery M. M.*, et al.*** Prevalence of Multidrug-Resistant Tuberculosis among Smear Positive Pulmonary Tuberculosis Patients in Eastern Sudan. *American Journal of Microbiological Research* 5(2), 32-6, (2017).
76. **Maalaoui N.** Strengthening TB Drug Management in the Sudanese National TB Control Program: In-Depth Review of TB Drug Management. Arlington, USA: Management Sciences for Health, U.S. Agency for International Development; 2009. Available from: http://pdf.usaid.gov-pdf_docs/Pdaco436.pdf.
77. **Sharaf Eldin G. S.*, et al.*** Molecular Analysis of Clinical Isolates of *Mycobacterium tuberculosis* Collected from Patients with Persistent Disease in the Khartoum Region of Sudan. *Journal of Infection* (44), 244-51, (2002).

78. **Eldirdery M. M.*, et al.*** Genotyping of Pulmonary *Mycobacterium tuberculosis* Isolates from Sudan Using Spoligotyping. *Am J Microbiol Res* 3(4), 125-8, (2015).
79. **Khalid F. A.*, et al.*** Molecular identification of *Mycobacterium tuberculosis* causing Pulmonary Tuberculosis in Sudan. *European Academic Research* 4(9), 7842-55, (2016).
80. **Devi K. R.*, et al.*** Genetic diversity of *M. tuberculosis* isolates from Assam, India: dominance of Beijing family and discovery of two new clades related to CAS1_Delhi and EAI family based on spoligotyping and MIRU-VNTR typing. *PLoSOne* 10(12), e0145860, (2015).
81. **WHO.** Global Tuberculosis Report. Geneva, Swizerland: WHO; 2012. Report No.: Annual Report I. Last Accessed on: 10.03.2019. Available from: https://www.who.int/tb/publications-global_report/gtbr12_main.pdf.
82. **Shuaib Y. A.*, et al.*** Smear microscopy for diagnosis of pulmonary tuberculosis in Eastern Sudan. *Tuberculosis Research and Treatment* 2018, 1-8, (2018).
83. **WHO.** Systematic screening for active tuberculosis: principles and recommendations. Geneva, Switzerland: World Health Organization; 2013. Contract No.: ISBN 978 92 4 154860 1. Last Accessed on: 11.01.2018. Available from: http://www.who.int/tb/tbscreening/en/.
84. **UN.** 2009. *Migration Review Tribunal*. Sudan. Available from: https://www.justice.gov-sites/default/files/eoir/legacy/2013/06/11/eastern%20sudan-stability.pdf
85. **Abdallah T. M.*, et al.*** Seroprevalence of HIV infection among tuberculosis patients in Kassala, eastern Sudan. *Journal of AIDS and HIV Research* 3(8), 161-3, (2011).
86. **Abdallah T. M.*, et al.*** Epidemiology of extra pulmonary tuberculosis in Eastern Sudan. *Asian Pacific Journal of Tropical Biomedicine* 5(6), 505-8, (2015).
87. **SNAP.** Epidemiology of HIV in Sudan Staging and Analysis Sudan. Khartoum, Sudan: Ministry of Heath; 2013. Last Accessed on: 11:04:2018. Available from: http://www.aidsinfoonline-org/kpatlas/document/SDN/SDN_2013_PSE_FSW_MSM.pdf.
88. **Sulieman H. M.*, et al.*** Assessment of Natural Vegetation Clearing and Re-Growth in Southern Gadarif (Sudan) Using Change Vector Analysis Based on Remote Sensing and Field Data. Proceedings of ISPRS Commission VII Mid-term Symposium "Remote Sensing: From Pixels to Processes"; May 8-11, 2006 Enschede, the Netherlands, 2006. p. PP. 586-91.
89. **Thrusfield M.** Veterinary Epidemiology Third edition ed. London, UK: Wiley-Blackwell Science Ltd.; 2007 January 2007. 624 p.
90. **MARF.** Number of animals in the Sudan. Khartoum, the Sudan: Ministry of Animal Resources and Fisheries (MARF), Statistics Dola; 2012.
91. **Omar H. E.*, et al.*** Constraints in the market chains for export of Sudanese sheep and sheep meat to the Middle East. Addis Ababa, Ethiopia: University of Khartoum, Khartoum, the Sudan, and International Livestock Research Institute, Nairobi, Kenya; 2009. Contract No.: 16. Last Accessed on: 10.10.2017. Available from: https://cgspace.cgiar.org-bitstream/handle/105687/ResearchReport_No16.pdf?sequence=1.
92. **Behnke R.*, et al.*** The Contribution of Livestock to the Sudanese Economy. Nairobi, Kenya: IGAD; 2012. Contract No.: IGAD LPI Working Paper No. 01 - 12. Available from: http://citeseerx-ist.psu.edu/viewdoc/download?doi=10.1.1.447.8147&rep=rep1&type=pdf.
93. **Shuaib Y. A.*, et al.*** Mycobacterial infections in carcasses of ruminants slaughtered at the two slaughterhouses of Kassala, Sudan. *Rev Elev Med Vet Pays Trop* 70(4), 131-136, (2018).
94. **Corner L.*, et al.*** Efficiency of inspection procedures for the detection of tuberculous lesions. *Aust Vet J* 67(11), 389-92, (1990).

95. **Corner L. A.** Post-mortem diagnosis of *Mycobacterium bovis* infection in cattle. *Vet Microbiol* 40(1-2), 53-63, (1994).
96. **DIN.** Medical Microbiology - Diagnosis of tuberculosis. Part 3: Detection of mycobacteria by culture methods. Berlin, Germany: Deutsches Institut für Normung; 2011. Available from: https://www.normen-ticker-kompakt.de/en/standard/din-58943-3/136883527
97. **Siddiqi S.*, et al.*** MGIT TM Procedure Manual. Geneva, Switzerland: Foundation for Innovative New Diagnostics (FIND); 2006. Last Accessed on: 11.01.2018. Available from: https://www.finddx.org/wp-content/uploads/2016/02/mgit_manual_nov2006.pdf.
98. **GLI.** Mycobacteriology Laboratory Manual. Geneva, Swizerland: WHO; 2014. Last Accessed on: 14.08.2017. Available from: http://www.who.int/tb/laboratory/mycobacteriologylaboratory-manual.pdf.
99. **WHO.** Quality assurance of sputum microscopy in DOTS programmes: regional guidelines for countries in the Western Pacific. Manila, Philippines. Geneva, Swizerland; 2003. Last Accessed on: 06.09.2017. Available from: http://www.who.int/ihr/training/laboratoryquality-11-cd_rom_tb_eqa_wpro.pdf.
100. **GLI.** GLI model TB diagnostic algorithms. Geneva, Switzerland: WHO; 2017. Last Accessed on: 10.01.2018. Available from: http://www.stoptb.org/wg/gli/assets/documents-GLI_algorithms.pdf.
101. **Somerville W.*, et al.*** Extraction of *Mycobacterium tuberculosis* DNA: a question of containment. *J Clin Microbiol* 43(6), 2996-7, (2005).
102. **Hillemann D.*, et al.*** Rapid detection of *Mycobacterium tuberculosis* Beijing genotype strains by real-time PCR. *J Clin Microbiol* 44(2), 302-6, (2006).
103. **Richter E.*, et al.*** Identification of *Mycobacterium kansasii* by using a dna probe (AccuProbe) and molecular techniques. *J Clin Microbiol* 37(4), 964-70, (1999).
104. **Richter E.*, et al.*** *Mycobacterium holsaticum* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 52(6), (2002).
105. **Rüsch-Gerdes S.*, et al.*** Multicenter laboratory validation of the BACTEC MGIT 960 technique for testing susceptibilities of *Mycobacterium tuberculosis* to classical second-line drugs and newer antimicrobials. *J Clin Microbiol* 44(3), 688-92, (2006).
106. **Ängeby K.*, et al.*** Challenging a dogma: antimicrobial susceptibility testing breakpoints for *Mycobacterium tuberculosis*. *Bull World Health Organ* 90(9), 693-8, (2012).
107. **Supply P.*, et al.*** Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* 44(12), (2006).
108. **Bolger A. M.*, et al.*** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15), 2114-20, (2014).
109. **Caverly L. J.*, et al.*** Complete Genome Sequences of 17 Rapidly Growing Nontuberculous Mycobacterial Strains. *Genome Announc* 4(5), (2016).
110. **Lagesen K.*, et al.*** RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35(9), 3100-8, (2007).
111. **Anon.** Annual Report for 2014. Kassala, the Sudan: Ministry of Agriculture and Animal Resources, Information DoSa; 2015.
112. **El-Dirani O. H.** Potential of Sudanese Exports of Sheep and Sheep Meat to the Middle East. Khartoum North, the Sudan: University of Khartoum; 2007.

113. **El-Dirani O. H.*, et al.*** Constraints in the market chains for export of Sudanese sheep and sheep meat to the Middle East. Nairobi, Kenya: University of Khartoum, Khartoum, the Sudan, and International Livestock Research Institute (ILRI), Nairobi, Kenya, Economics DoA; 2009. Contract No.: 16. Available from: http://ageconsearch.umn.edu/bitstream-99128/2/2009-Sudan-20sheep%20market%20ILRIRR16.pdf.
114. **Hala M. A.*, et al.*** Determinants of livestock sector return in Kassel State, Sudan. *SUST Journal of Economic Sciences* 17(1),131-47. (Paper in Arabic) (2016).
115. **Richter E.*, et al.*** Evaluation of the GenoType *Mycobacterium* Assay for identification of mycobacterial species from cultures. *J Clin Microbiol* 44(5), 1769-75, (2006).
116. **Weniger T.*, et al.*** MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. *Nucleic Acids Res* 38, (2010).
117. **Heyckendorf J.*, et al.*** What is resistance? Impact of phenotypic versus molecular drug resistance testing on multi- and extensively drug-resistant tuberculosis therapy. *Antimicrobial Agents and Chemotherapy* 62(2),pii: e01550-17, (2018).
118. **Mbugi E. V.*, et al.*** Mapping of *Mycobacterium tuberculosis* Complex Genetic Diversity Profiles in Tanzania and Other African Countries. *PLoS One* 11(5), e0154571, (2016).
119. **Fattorini L.*, et al.*** Tuberculosis in migrants from 106 countries to Italy, 2008-2014. *Eur Respir J* 47(4), 1273-6, (2016).
120. **Jajou R.*, et al.*** A Predominant Variable-Number Tandem-Repeat Cluster of *Mycobacterium tuberculosis* Isolates among Asylum Seekers in the Netherlands and Denmark, Deciphered by Whole-Genome Sequencing. *J Clin Microbiol* 56(2), e01100-17, (2018).
121. **Fitzgibbon M. M.*, et al.*** A snapshot of genetic lineages of *Mycobacterium tuberculosis* in Ireland over a two-year period, 2010 and 2011. *Eurosurveillance* 18(3), 1 - 7, (2013).
122. **Hahn M. Y.*, et al.*** The *Mycobacterium tuberculosis* Extracytoplasmic-Function Sigma Factor SigL Regulates Polyketide Synthases and Secreted or Membrane Proteins and Is Required for Virulence. *Journal of Bacteriology* 187(20), 7062-71, (2005).
123. **Lun S.*, et al.*** Characterization of a novel cell wall-anchored protein with carboxylesterase activity required for virulence in *Mycobacterium tuberculosis*. *J Biol Chem* 282(25), (2007).
124. **Klepp L. I.*, et al.*** Identification of two proteins that interact with the *Erp* virulence factor from *M. tuberculosis* by using the bacterial two-hybrid system. *BMC Mol Biol* 10,3, (2009).
125. **Sheline K. D.*, et al.*** Does the lipR gene of tubercle bacilli have a role in tuberculosis transmission and pathogenesis? *Tuberculosis (Edinb)* 89(2), 114-9, (2009).
126. **Millan-Lou M. I.*, et al.*** Rapid test for identification of a highly transmissible *Mycobacterium tuberculosis* Beijing strain of Sub-Saharan origin. *J Clin Microbiol* 50(2), (2011).
127. **Zacharia V. M.*, et al.*** cor, a novel carbon monoxide resistance gene, is essential for *Mycobacterium tuberculosis* pathogenesis. *MBio* 4(6), e00721-13, (2013).
128. **Nebenzahl-Guimaraes H.*, et al.*** Transmissible *M. tuberculosis* strains share genetic markers and immune phenotypes. *Am J Respir Crit Care Med* 195(11), 1519-27, (2017).
129. **Brynildsrud O. B.*, et al.*** Global expansion of *Mycobacterium tuberculosis* lineage 4 shaped by colonial migration and local adaptation. *Sci Adv* 4(10), eaat58691, (2018).
130. **Rutaihwa L. K.*, et al.*** Multiple Introductions of the *Mycobacterium tuberculosis* Lineage 2 Beijing into Africa over centuries. 10.1101/413039 (2018).
131. **Humblet M. F.*, et al.*** Classification of worldwide bovine tuberculosis risk factors in cattle: a stratified approach. *Vet Res* 40(5), 50, (2009).

132. **Manal H. S.*, et al.*** Bovine Tuberculosis at Omdurman Central Abattoir, Khartoum State and Wau Slaughterhouses (Bahr El-Ghazal) State, Sudan. *Sudan J Vet Res* 25, 1-8, (2010).
133. **Ayman E. A.*, et al.*** Prevalence of Bovine tuberculosis in Kassala state. *Inter J Adv Sci Techn Res* 4(1), 238-49, (2014).
134. **Worku A.*, et al.*** Cross-sectional study and comparison of different diagnostic methods of bovine tuberculosis in Gondar Elfora Abattoir, Ethiopia. *Mycobact Dis* 6(3), (2016).
135. **Nalapa D. P.*, et al.*** Prevalence of tuberculous lesion in cattle slaughtered in Mubende district, uganda. *bmc vet res* 13(1), 73, (2017).
136. **Okeke L. A.*, et al.*** Bovine tuberculosis: a retrospective study at Jos abattoir, Plateau State, Nigeria. *Pan Afr Med J* 25, 202, (2016).
137. **Rieder H. L.*, et al.*** Priorities for tuberculosis bacteriology services in low-income countries. 2nd ed. Paris, France: The Union; 2007. 150 p.
138. **Zhang G.*, et al.*** Screening and Assessing 11 *Mycobacterium tuberculosis* Proteins as Potential Serodiagnostical Markers for Discriminating TB Patients from BCG Vaccinees. *Genomics, Proteomics & Bioinformatics* 7(3), (2009).
139. **Kivihya-Ndugga L. E.*, et al.*** A comprehensive comparison of Ziehl-Neelsen and fluorescence microscopy for the diagnosis of tuberculosis in a resource-poor urban setting. *Int J Tuberc Lung Dis* 7(12), 1163-71, (2003).
140. **George G.*, et al.*** Comparison of the efficacies of loop-mediated isothermal amplification, fluorescence smear microscopy and culture for the diagnosis of tuberculosis. *PLoSOne* 6(6), e21007, (2011).
141. **Bonnet M.*, et al.*** Performance of LED-based fluorescence microscopy to diagnose tuberculosis in a peripheral health centre in Nairobi. *PLoSOne* 6(2), e17214, (2011).
142. **Alfred N.*, et al.*** Optimising *Mycobacterium tuberculosis* detection in resource limited settings. *BMJ Open* 4(3), e004093, (2014).
143. **Gelaw B.*, et al.*** Comparison of loop-mediated isothermal amplification assay and smear microscopy with culture for the diagnostic accuracy of tuberculosis. *BMC Infect Dis* 17(1),79, (2017).
144. **WHO.** Approaches to improve sputum smear microscopy for tuberculosis diagnosis expert group meeting. Geneva, Swizerland: WHO; 2009. Last Accessed on: 14.08.2017. Available from: http://www.who.int/tb/laboratory/egmreport_microscopymethods_nov09-pdf?ua=1.
145. **Baliga S.*, et al.*** Rapid method for detecting and differentiating *Mycobacterium tuberculosis* complex and non-tuberculous mycobacteria in sputum by fluorescence in situ hybridization with DNA probes. *International Journal of Infectious Diseases* 75(2018), 1-7, (2018).
146. **Baldwin S. L.*, et al.*** The complexities and challenges of preventing and treating nontuberculous mycobacterial diseases. *PLoS Negl Trop Dis* 13(2), e0007083, (2019).
147. **Houben R.*, et al.*** What if they don't have tuberculosis? the consequences and trade-offs involved in false-positive diagnoses of tuberculosis. *Clin Infect Dis* 68(1), 150-6, (2019).
148. **Tanimura T.*, et al.*** Financial burden for tuberculosis patients in low- and middle-income countries: a systematic review. *Eur Respir J* 43(6), 1763-75, (2014).
149. **Merker M.*, et al.*** Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia. *Elife* 7, (2018).

150. **Varma-Basil M.*, et al.*** Comparison of spoligotyping, MIRU-VNTR typing and IS6110-RFLP in a study of genotypic diversity of *Mycobacterium tuberculosis* in Delhi, North India. *Mem Inst Oswaldo Cruz, Rio de Janeiro* 106(5), 524-35, (2011).
151. **Ali A.*, et al.*** *M. tuberculosis* Central Asian Strain (CAS) lineage strains in Pakistan reveal lower diversity of MIRU loci than other strains. *Int J Mycobacteriol* 3(2), 108-16, (2014).
152. **Ali R.*, et al.*** Molecular epidemiology and genotyping of *Mycobacterium tuberculosis* isolated in Baghdad. *Biomed Res Int* 2014, (2014).
153. **Al-Hajoj S.*, et al.*** Tuberculosis in Saudi Arabia: the journey across time. *J Infect Dev Ctries* 9(3),222-31, (2015).
154. **Biadglegne F.*, et al.*** Tuberculous Lymphadenitis in Ethiopia Predominantly Caused by Strains Belonging to the Delhi/CAS Lineage and Newly Identified Ethiopian Clades of the *Mycobacterium tuberculosis* Complex. *PLoS One* 10(9), e0137865, (2015).
155. **Mekonnen A.*, et al.*** Molecular epidemiology and drug resistance patterns of *Mycobacterium tuberculosis* complex isolates from university students and the local community in Eastern Ethiopia. *PLoSOne* 13(9), e0198054, (2018).
156. **Tanveer M.*, et al.*** Genotyping and drug resistance patterns of *Mycobacterium tuberculosis* strains in Pakistan. *BMC Infect Dis* 8, 171, (2008).
157. **Varghese B.*, et al.*** Tuberculosis transmission among immigrants and autochthonous populations of the eastern province of Saudi Arabia. *PLoS One* 8(10), e77635, (2013).
158. **Comas I.*, et al.*** Population genomics of *M. tuberculosis* in Ethiopia contradicts the virgin soil hypothesis for human tuberculosis in Sub-Saharan Africa. *Curr Biol* 25(24), 3260-6, (2015).
159. **IOM.** Migration in Sudan: A Country Profile 2011. Khartoum, the Sudan; 2011. Last Accessed on: 12.03.2018. Available from: https://reliefweb.int/sites/reliefweb.int/filesresourcesD2-ECC0D6226DF41DC1257842003973F1-Full_Report.pdf.
160. **Amal H. F.** Eritrea: The Rashaida People. Asmara, Eritrea: Madote; 2018 [Available from: http://www.madote.com/2010/02/eritrea-rashaida-people.html.
161. **Joshua P.** Rashaida in Sudan. Pasadena, USA: Ministry of Frontier Ventures, USA; 2018 Available from: https://joshuaproject.net/people_groups/14523/SU.
162. **Zaki A. M.*, et al.*** History of TB in the Sudan. *Sudan J Medical Sci* 4(2), 179-88, (2009).
163. **MEA. 2014.** *India-Sudan Relations*.Last Accessed on: 12.03.2018. Khartoum, the Sudan: Embassy of India, Khartoum, the Sudan. Available from: https://www.mea.gov.in/Portal-Foreign/Relation/Sudan_July_2014__2_.pdf
164. **MEA. 2017.** *India-Sudan Relations*.Last Accessed on: 12.03.2018. Khartoum, the Sudan: Embassy of India, Khartoum, the Sudan. Available from: http://www.mea.gov.in/Portal-Foreign/Relation/Sudan_Nov_2017.pdf
165. **Campbell M. C.*, et al.*** African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annu Rev Genomics Hum Genet* 9, 403-33, (2008).
166. **Diab H. M.*, et al.*** First insight into the genetic population structure of *Mycobacterium tuberculosis* isolated from pulmonary tuberculosis patients in Egypt. *Tuberculosis (Edinb)* (96), 13-20, (2016).
167. **Tagliani E.*, et al.*** Culture and Next-generation sequencing-based drug susceptibility testing unveil high levels of drug-resistant-TB in Djibouti: results from the first national survey. *Sci Rep* 7(1), (2017).

168. **Tadesse M.*, et al.*** The predominance of Ethiopian specific *Mycobacterium tuberculosis* families and minimal contribution of Mycobacterium bovis in tuberculous lymphadenitis patients in Southwest Ethiopia. *Infect Genet Evol* 55, 251-9, (2017).
169. **Walker T. M.*, et al.*** A cluster of multidrug-resistant *Mycobacterium tuberculosis* among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. *The Lancet Infectious Diseases*, (2018).
170. **Migliori G. B.*, et al.*** Diagnosis of multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis: Current standards and challenges. *Can J Infect Dis Med Microbiol* 19(2), 169–72, (2008).
171. **Abu Shanab M. E.** Defaulting to anti-tuberculosis treatment: proportional and associated factors among internally displaced people around khartoum state. Khartoum, the Sudan: Faculty of Public and Environmental Health, University of Khartoum; 2003.
172. **Ali A. O.*, et al.*** Patient non adherence to tuberculosis treatment in Sudan: socio demographic factors influencing non adherence to tuberculosis therapy in Khartoum State. *Pan Afr Med J* 25, 80, (2016).
173. **Esmael A.*, et al.*** Drug Resistance Pattern of *Mycobacterium tuberculosis* in Eastern Amhara Regional State, Ethiopia. *JMBT* 6(2), 75-9, (2014).
174. **Abdella K.*, et al.*** Drug resistance patterns of *Mycobacterium tuberculosis* complex and associated factors among retreatment cases around Jimma, Southwest Ethiopia. *BMC Public Health* 15, 599, (2015).
175. **Bedewi Z.*, et al.*** *Mycobacterium tuberculosis* in central Ethiopia: drug sensitivity patterns and association with genotype. *New Microbes New Infect* 17, 69-74, (2017).
176. **Kerubo G.*, et al.*** Drug susceptibility profiles of pulmonary *M. tuberculosis* isolates from patients in informal urban settlements in Nairobi, Kenya. *BMC Infect Dis* 16(1), 583, (2016).
177. **Yonge S. A.*, et al.*** Drug Susceptibility Patterns of *Mycobacterium tuberculosis* Isolates from Tuberculosis Patients in Coastal Kenya. *J Tuberc Res* 05(04), 201-19, (2017).
178. **Lukoye D.*, et al.*** Variation and risk factors of drug resistant tuberculosis in sub-Saharan Africa: a systematic review and meta-analysis. *BMC Public Health* 15, 291, (2015).
179. **Lukoye D.*, et al.*** Anti-tuberculosis drug resistance among new and previously treated sputum smear-positive tuberculosis patients in Uganda: results of the first national survey. *PLoS One* 8(8), e70763, (2013).
180. **UNAIDS.** UNAIDS report on the global AIDS epidemic. Geneva, Switzerland: WHO; 2017. Report No.: UNAIDS/JC2910E. Last Accessed on: 12.03.2018. Available from: http://www.unaids-org/en/resources/documents/2017/2017_data_book.
181. **Gröschel M. I.*, et al.*** Pathogen-based precision medicine for drug-resistant tuberculosis. *PLoS Pathog* 14(10), e1007297, (2018).
182. **Pankhurst L. J.*, et al.*** Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. *Lancet Respir Med T* 4(1), 49-58, (2016).
183. **Schleusener V.*, et al.*** *M. tuberculosis* resistance prediction and lineage classification from genome sequencing: comparison of automated analysis tools. *Sci Rep* 7, 46327, (2017).
184. **Kibiki G. S.*, et al.*** *M. tuberculosis* genotypic diversity and drug susceptibility pattern in HIV-infected and non-HIV-infected patients in northern Tanzania. *BMC Microbiol* 7, 51, (2007).
185. **Tessema B.*, et al.*** Analysis of gene mutations associated with isoniazid, rifampicin and ethambutol resistance among *Mycobacterium tuberculosis* isolates from Ethiopia. *BMC Infect Dis* 10(12), (2012).

186. **Ssengooba W.*, et al.*** Whole genome sequencing to complement tuberculosis drug resistance surveys in Uganda. *Infect Genet Evol* 40,8-16, (2016).
187. **Sabeel S. M.*, et al.*** Phenotypic and genotypic analysis of multidrug-resistant *Mycobacterium tuberculosis* isolates from sudanese patients. *Tuberc Res Treat* 2017, 8340746, (2017).
188. **Kompala T.*, et al.*** Transmission of tuberculosis in resource-limited settings. *Curr HIV/AIDS Rep* 10(3), 264-72, (2013).
189. **Oren E.*, et al.*** Neighborhood socioeconomic position and tuberculosis transmission: a retrospective cohort study. *BMC Infect Dis* 27(14), 227, (2014).
190. **Yimer S. A.*, et al.*** *Mycobacterium tuberculosis* lineage 7 strains are associated with prolonged patient delay in seeking treatment for pulmonary tuberculosis in Amhara Region, Ethiopia. *J Clin Microbiol* 53(4), (2015).
191. **Middelkoop K.*, et al.*** Transmission of tuberculosis in a South African community with a high prevalence of HIV infection. *J Infect Dis* 211(1), 53-61, (2015).
192. **Al-Hajoj S. A.** Tuberculosis in Saudi Arabia: can we change the way we deal with the disease? *J Infect Public Health* 3(1), 17-24, (2010).
193. **Belay T.*, et al.*** Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia. *BMC Infectious Diseases* 13(131), 1-11, (2013).
194. **Hamed F. A.*, et al.*** Poverty among Sudanese communities along the eastern borders: A case study from the Kassala and Gedarif States. Bergen, Norway: Chr. Michelsen Institute, University of Bergen; 2017. Last Accessed on: 11.12.2018. Available from: https://www-cmi.no-publications/file/6259-poverty-among-sudanese-communities-along-the.pdf.
195. **Cohen K. A.*, et al.*** Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of *Mycobacterium tuberculosis* Isolates from KwaZulu-Natal. *PLoS Med* 12(9), e1001880, (2015).
196. **Auld S. C.*, et al.*** Extensively drug-resistant tuberculosis in South Africa: genomic evidence supporting transmission in communities. *Eur Respir J* 52(4), (2018).
197. **WHO.** Multidrug and extensively drug-resistant tuberculosis (M/XDR-TB): 2010 global report on surveillance and response. Geneva, Switzerland: World Health Organization; 2010. Available from: http://apps.who.int/iris/bitstream/handle/10665/44286/9789241599191-eng.pdf=jsessionid=9596D72EC9E6C0470C481B6EFA976B72?sequence=1.
198. **WHO.** Systematic screening for active tuberculosis: principles and recommendations. Geneva, Switzerland: WHO; 2013. Last Accessed on: 17.11.2017. Available from: http://w-ww.who.int/tb/publications/Final_TB_Screening_guidelines.pdf.
199. **Schierloh P.*, et al.*** Differential expression of immunogenic proteins on virulent *Mycobacterium tuberculosis* clinical isolates. *Biomed Res Int* 2014, 741309, (2014).
200. **Satta G.*, et al.*** Genetic variation in *Mycobacterium tuberculosis* isolates from a London outbreak associated with isoniazid resistance. *BMC Med* 14(1), 117, (2016).
201. **Dasgupta A.*, et al.*** An oligopeptide transporter of *Mycobacterium tuberculosis* regulates cytokine release and apoptosis of infected macrophages. *PLoS One* 5(8), e12225, (2010).
202. **Nisa S.*, et al.*** Targeting the chromosome partitioning protein *ParA* in tuberculosis drug discovery. *J Antimicrob Chemother* 65(11), 2347-58, (2010).
203. **Tullius M. V.*, et al.*** Discovery and characterization of a unique mycobacterial heme acquisition system. *Proc Natl Acad Sci U S A* 108(12), 5051-6, (2011).

204. **Manganelli R.*, et al.*** Identification of *M. tuberculosis* Rv3441c and *M. smegmatis* MSMEG 1556 and Essentiality of *M. smegmatis* MSMEG_1556. *PLoS ONE* 7(8), e42769, (2012).
205. **Lingaraju S.*, et al.*** Geographic Differences in the Contribution of ubiA Mutations to High-Level Ethambutol Resistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 60(7), 4101-5, (2016).
206. **Yu X.*, et al.*** Wild-Type and Non-Wild-Type *Mycobacterium tuberculosis* MIC Distributions for the Novel Fluoroquinolone Antofloxacin Compared with Those for Ofloxacin, Levofloxacin, and Moxifloxacin. *Antimicrob Agents Chemother* 60(9), 5232-7, (2016).
207. **Shi W.*, et al.*** Identification of Novel Mutations in *LprG* (rv1411c), rv0521, rv3630, rv0010c, *ppsC*, and *cyp128* Associated with Pyrazinoic Acid/Pyrazinamide Resistance in *Mycobacterium tuberculosis*. *Antimicro Agents Chemother* 62(7), pii: e00430-18. (2018).
208. **Xie L.*, et al.*** First succinyl-proteome profiling of extensively drug-resistant *Mycobacterium tuberculosis* revealed involvement of succinylation in cellular physiology. *J Proteome Res* 14(1), 107-19, (2015).
209. **Varghese B.*, et al.*** Impact of *Mycobacterium tuberculosis* complex lineages as a determinant of disease phenotypes from an immigrant rich moderate tuberculosis burden country. *Respir Res* 19(1), 259, (2018).
210. **Stavrum R.*, et al.*** In-depth molecular characterization of *Mycobacterium tuberculosis* from New Delhi--predominance of drug resistant isolates of the 'modern' (TbD1) type. *PLoSOne* 4(2), e4540, (2009).
211. **Vadwai V.*, et al.*** Do diverse *Mycobacterium tuberculosis* strains circulate in Mumbai? *Asia-Pac J Trop Dis* 2(2012), S563-S6, (2012).
212. **Chatterjee A.*, et al.*** MIRU-VNTR profiles of three major *Mycobacterium tuberculosis* spoligotypes found in western India. *Tuberculosis (Edinb)* 93(2), 250-6, (2013).
213. **Casali N.*, et al.*** Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat Genet* 46(3), 279-86, (2014).
214. **Sharkey H. J.** The Sudan of the Three Niles: The Funj Chronicle. *International Journal Middle East Studies* 32(2), 288-90, (2000).
215. **Collins R.** The wealth of Africa: the kingdom of Aksum. London, UK: British Museum, London, UK; 2010. Last Accessed on: 22.02.2019. Available from: https://www.british-museum-org/pdf/KingdomOfAksum_TeachersNotes.pdf.
216. **Selassie Y. G.** Plague as a Possible Factor for the Decline and Collapse of the Aksumite Empire: a New Interpretation. *ITYOPIS* 1, 36-61, (2011).
217. **MEA.** 2017. *India-Ethiopia Relations*.Last Accessed on: 11.03.2018. Addis Ababa, Ethiopia: Embassy of India, Addis Ababa, Ethiopia. Available from: https://www.mea.gov.in/Portal-ForeignRelation/Ethiopia_Sept_2017.pdf
218. **Mosher M. S.** The architecture of Mohenjo-Daro as evidence for organization of Indus Valley Civilization urban neighbourhood. Toronto, Canada: University of Toronto; 2017.
219. **Gensheimer T. R.** The Role of shell in Mesopotamia : evidence for trade exchange with Oman and the Indus Valley. *Paléorient* 10(1), 65-73, (1984).
220. **Prabhakar V. N.** Harappans and their Mesopotamian contacts. New Delhi, India: the India International Centre; 2013. Report No.: 48. Available from: http://www.iicdelhi.nic.in/Content/Attachments/Publications/DiaryFiles/145615June132013_IIC_48_13_06_2013.pdf.
221. **Al-Hajoj S. A.*, et al.*** First insight into the population structure of *Mycobacterium tuberculosis* in Saudi Arabia. *J Clin Microbiol* 45(8), 2467-73, (2007).

222. **Varghese B.*, et al.*** Admixed phylogenetic distribution of drug resistant *Mycobacterium tuberculosis* in Saudi Arabia. *PLoSOne* 8(2), e55598, (2013).
223. **Al-Ghafli H.*, et al.*** Drug-resistance profiling and transmission dynamics of multidrug-resistant *Mycobacterium tuberculosis* in Saudi Arabia revealed by whole genome sequencing. *Infect Drug Resist* 11, 2219-29, (2018).
224. **Abouzeid M. S.*, et al.*** Tuberculosis trends in Saudis and non-Saudis in the Kingdom of Saudi Arabia--a 10 year retrospective study (2000-2009). *PLoS One* 7(6), e39478, (2012).
225. **Hasan Z.*, et al.*** Spoligotyping of *Mycobacterium tuberculosis* isolates from Pakistan reveals predominance of CAS 1 and Beijing isolates. *J Clin Microbiol* 44(5), 1763-8, (2006).
226. **Agonafir M.*, et al.*** Phenotypic and genotypic analysis of multidrug resistant tuberculosis in Ethiopia. *Int J Tuberc Lung Dis* 14(10), 1259 - 65, (2010).
227. **Ayaz A.*, et al.*** Characterizing *Mycobacterium tuberculosis* isolates from Karachi, Pakistan: drug resistance and genotypes. *Int J Infect Dis* 16(4), e303-9, (2012).
228. **Dhatwalia S. K.*, et al.*** Distribution and association of *M. tuberculosis* CAS lineage with multidrug resistance in North India. *Int J Tuberc Lung Dis* 20(6), 806-11, (2016).
229. **Ndungu P. W.*, et al.*** Mycobacteria Interspersed Repetitive Units-Variable Number of Tandem Repeat, Spoligotyping and Drug Resistance of Isolates from Pulmonary Tuberculosois Patients in Kenya. *Advances in Microbiology* 07(03), 205-16, (2017).
230. **Maharjan B.*, et al.*** Genetic diversity and distribution dynamics of multidrug-resistant *Mycobacterium tuberculosis* isolates in Nepal. *Sci Rep* 8(1), 16634, (2018).
231. **Cain K. P.*, et al.*** The movement of multidrug-resistant tuberculosis across borders in East Africa needs a regional and global solution. *PLoS Med* 12(2), e1001791, (2015).
232. **Lawrence D. B.*, et al.*** **I**nterval Estimation for a Binomial Proportion. *Statistical Science* 16(2), 101-33, (2001).

6 Annexes

Annex 1: Materials for sputum sample collection

Material	Manufacturer	Location
Sputum collection container	Thermo Fischer Scientific	Langenselbold, Germany
50 ml centrifuge tubes	Sarstedt	Nümbrecht, Germany
Eppendorf tubes	Sarstedt	Nümbrecht, Germany
Isoamylalcohol	Merck KGaA	Darmstadt, Germany
Biosafety container (BIOTAINER)	E3 Cortex	Thieux, France

Annex 2: Materials for decontamination of collected sputum samples using the N-acetyl-L-cysteine sodium hydroxide (NALC-NaOH) method

Material	Manufacturer	Location
NALC	Merck KGaA	Darmstadt, Germany
NaOH	Merck KGaA	Darmstadt, Germany
50 ml centrifuge tubes	Sarstedt	Nümbrecht, Germany
Orbital shaker standard analog	VWR	Darmstadt, Germany
Phosphate buffer	Merck KGaA	Darmstadt, Germany
Tween 80	Merck KGaA	Darmstadt, Germany
Heraeus Multifuge X3R centrifuge	Thermo Fischer Scientific	Langenselbold, Germany
Mischgefäß (mixing vessel)	IKA	Staufen, Germany
Tube drive control	IKA	Staufen, Germany
Vortex 1	IKA	Staufen, Germany
Petri dish	Sarstedt	Nümbrecht, Germany

Annex 3: Materials for smear microscopy of sputum samples collection

Material	Manufacturer	Location
Microscope slides	Paul Marienfeld	Königshofen, Germany
DEPC treated H_2O	Merck KGaA	Darmstadt, Germany
Carbol Fuchsin stain	RAL diagnostics	Schiltigheim, France
Acid alcohol 3%	RAL diagnostics	Schiltigheim, France
Methylene blue stain 0.3%	RAL diagnostics	Schiltigheim, France
Brilliant green 1%	RAL diagnostics	Schiltigheim, France
ZN Aerospray stainer	ELITechGroup	Logan, USA
Light microscope Leitz Laborlux D	Lapequip LTD.	Ontario, Canada
Light microscope DM 2500	Leica	Wetzlar, Germany
Auramine O staining	RAL diagnostics	Schiltigheim, France
Potassium permanganate	RAL diagnostics	Schiltigheim, France
LED microscope	Leica	Wetzlar, Germany
Microscope immersion oil	Cargille Labs	New York, USA

Annex 4: Materials for culturing and DST of sputum samples collection

Material	Manufacturer	Location
Class 2 cabinet	ScanLaf	Lynge, Denmark
Löwenstein-Jensen media	Artelt-Enclit	Leipzig, Germany

Annex 4 continued

Material	Manufacturer	Location
Stonebrink media	Artelt-Enclit	Leipzig, Germany
Middlebrook 7H10 media	Artelt-Enclit	Leipzig, Germany
MGIT media (Middelbrook 7H9)	Becton-Dickinson	Heidelberg, Germany
MGIT growth supplement	Becton-Dickinson	Heidelberg, Germany
MGIT PANTA	Becton-Dickinson	Heidelberg, Germany
BACTEC MGIT 960 system	Becton-Dickinson	Heidelberg, Germany
Columbia blood agar	Becton-Dickinson	Heidelberg, Germany
Inoculating loops	Thomas Scientific	Swedesboro, USA
Disposable loops	Sarstedt	Nümbrecht, Germany
Pasteur pipette	Sarstedt	Nümbrecht, Germany
Incubator	Hitachi	Tokyo, Japan
MGIT SIRE and PZA kit	Becton-Dickinson	Heidelberg, Germany
Ethanol	Merck KGaA	Darmstadt, Germany
HCl (1M)	Merck KGaA	Darmstadt, Germany
Glycerin	Merck KGaA	Darmstadt, Germany

Annex 5: Materials for DNA extraction, PCR, and LPA

Material	Manufacturer	Location
Tris-EDTA	Carl Roth	Karlsruhe, Germany
Tween 80	Merck KGaA	Darmstadt, Germany
Disinfectants	Laboratories Anios	Lille- Hellemmes, France
Elution-buffer	Macherey-Nagel	Düren, Germany
Ethanol	Scharlab SL	Barcelona, Spain
HCl (1M)	Merck KGaA	Darmstadt, Germany
QIAamp® DNA Mini kit	Qiagen	Hilden, Germany
CTAB DNA extraction kit	Qiagen	Hilden, Germany
DNA amplification kit	Qiagen	Hilden, Germany
Rotor-Gene® Q	Qiagen	Hilden, Germany
PCR pure kit	Analytik Jena AG	Jena, Germany
HAIN GenoType CM kit	HAIN Lifescience GmbH	Nehren, Germany
HAIN GenoType MTBC kit	HAIN Lifescience GmbH	Nehren, Germany
TwinCubator	HAIN Lifescience GmbH	Nehren, Germany
Eppendorf tubes	Sarstedt	Nümbrecht, Germany
Water heater	DuX Prodigy	Milan, Italy
Sonicator	Branson	Danbury, USA
Vortex VF2 Janke and Kunkel	IKA	Staufen, Germany
Heraeus Megafuge 8 centrifuge	Thermo Fischer Scientific	Langenselbold, Germany
Eppendorf centrifuge 5418	Eppendorf	Hamburg, Germany
Reax vortex mixer	Heidolph	Schwabach, Germany
Microtiter pipette	Nerbe plus GmbH	Winsen, Germany
Mini-PROTEAN® tetra electrode	Bio-Rad	München, Germany
Ultra-Pure™ Agarose	Thermo Fischer Scientific	Langenselbold, Germany
Ethidium bromide	Merck KGaA	Darmstadt, Germany
C1000 Touch™ Thermal Cycler	Bio-Rad	München, Germany

Annex 5 continued

Material	Manufacturer	Location
DNA/RNA UV-cleaner box	Biosan	Riga, Latvia
S-Monovette® rack D12	Sarstedt	Nümbrecht, Germany
Water bath WNB / WNE / WPE	Memmert GmbH & Co. KG	Schwabach, Germany

Annex 6: Materials for spoligotyping, MIRU-VNTR, and Sanger sequencing

Material	Manufacturer	Location
DEPC treated H_2O	Merck KGaA	Darmstadt, Germany
DMSO	Merck KGaA	Darmstadt, Germany
Glycerol for analysis EMSURE® ACS	Merck KGaA	Darmstadt, Germany
Ethidium bromide	Merck KGaA	Darmstadt, Germany
Multiscreen®-HV filtration plate	Merck KGaA	Darmstadt, Germany
Sephadex™ G50	Merck KGaA	Darmstadt, Germany
HotStarTaq DNA Polymerase (5U/µL)	Qiagen	Hilden, Germany
HotStarTaq Puffer (10x)	Qiagen	Hilden, Germany
$MgCl_2$ (25mM)	Qiagen	Hilden, Germany
GelPilot 100 bp DNA Ladder (100)	Qiagen	Hilden, Germany
Roti®-Stock TE-Puffer (100x)	Carl Roth	Karlsruhe, Germany
dNTPs (25mM)	Carl Roth	Karlsruhe, Germany
Rotiphorese® 50x TAE Puffer	Carl Roth	Karlsruhe, Germany
Streptavidin-peroxidase	Roche	Mannheim, Germany
Detection reagent	Roche	Mannheim, Germany
MIRU-VNTR Triplex PCR-Kit	Genoscreen	Lille, France
MicroAmp® 96-well plates	Bio-Rad	München, Germany
C1000 Touch Thermal Cyler	Bio-Rad	München, Germany
T100 Thermal Cycler	Bio-Rad	München, Germany
DRa and DRb primers	TIB Molbiol	Berlin, Germany
Hi-Di™ Formamide	Thermo Fischer Scientific	Langenselbold, Germany
Ultra-Pure™ Agarose	Thermo Fischer Scientific	Langenselbold, Germany
Exonuclease I (20U/µL)	Thermo Fischer Scientific	Langenselbold, Germany
FastAP Alkaline phosphatase (1U/µL)	Thermo Fischer Scientific	Langenselbold, Germany
BigDye® Terminator v3.1	Thermo Fischer Scientific	Langenselbold, Germany
Sequencing buffer (5x)	Thermo Fischer Scientific	Langenselbold, Germany
3130xl Genetic Analyzer	Thermo Fischer Scientific	Langenselbold, Germany
3500xl Genetic Analyzer	Thermo Fischer Scientific	Langenselbold, Germany
96 well Septa	Thermo Fischer Scientific	Langenselbold, Germany
Flex Cycler	Analytik Jena	Jena, Deutschland
Universal 32R Zentrifuge	Hettich Zentrifugen	Tuttlingen, Deutschland
Gel iX Imager	Intas	Göttingen, Deutschland

Annex 7: Materials for whole genome sequencing

Material	Manufacturer	Location
MiSeq	Illumina	San Diego, USA
Nextera™ (XT) Sample Preparation Kit	Illumina	San Diego, USA

Annex 5 continued

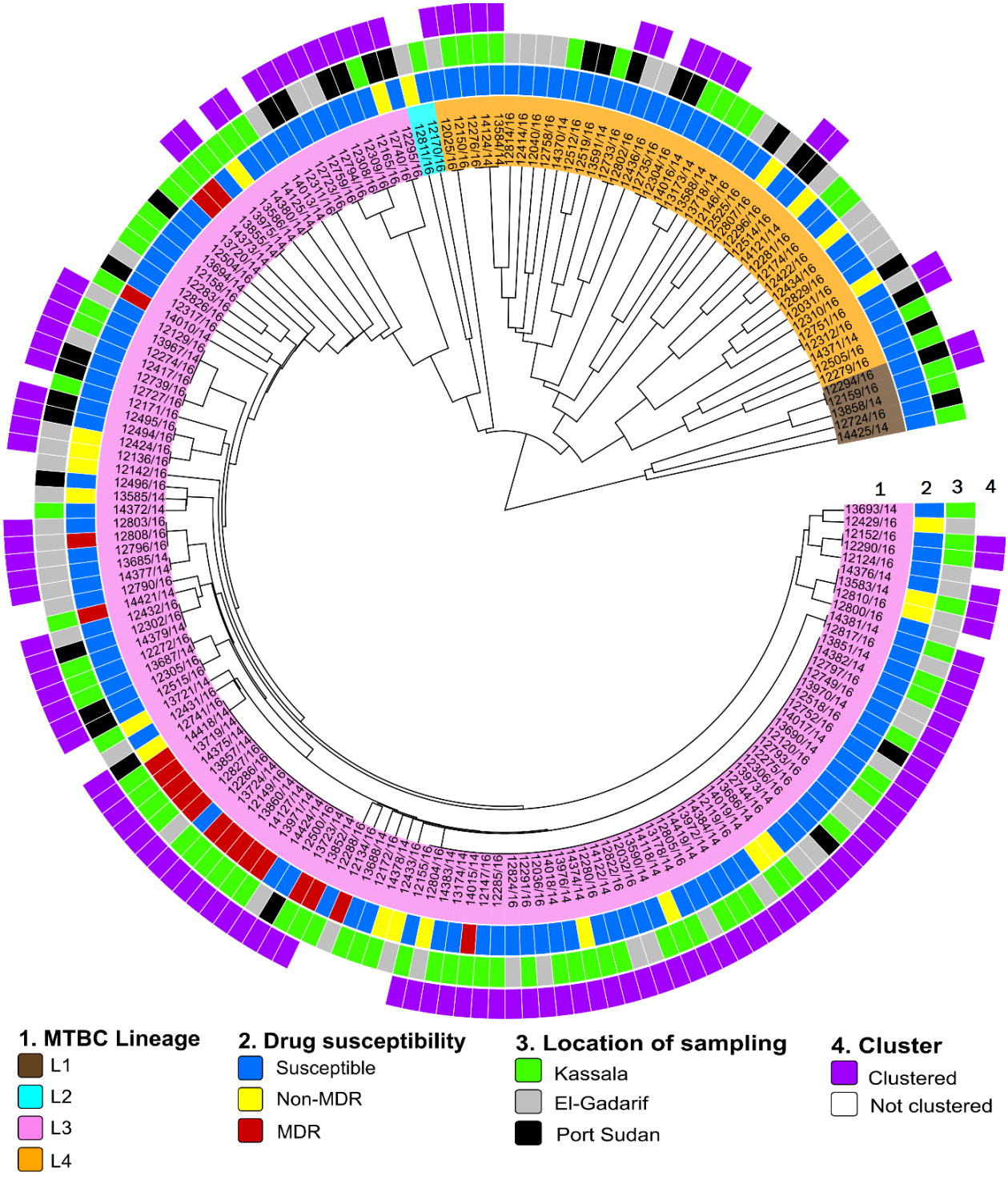
Material	Manufacturer	Location
Nextera™ Index Kit (24 indices)	Illumina	San Diego, USA
Quant-it™ DNA assay kit	Life Technologies	Darmstadt, Germany
Qubit®	Life Technologies	Darmstadt, Germany
NanoPhotometer® P330	Implen	München, Germany
High Sensivity DNA Kit	Agilent Technologies	Santa Clara, USA
MiSeq Reagent Kit v1, v2, v3	Illumina	San Diego, USA
Agencourt® AMPure® XP Kit	Beckmann Coulter	Krefeld, Germany
Ethanol (abs.)	Merck KGaA	Darmstadt, Germany
NaOH (1M)	Merck KGaA	Darmstadt, Germany
Elution-Buffer (5 nM Tris/HCl, pH8,5)	Macherey-Nagel	Düren, Germany
HCl (1M)	Merck KGaA	Darmstadt, Germany
Tween® 20	Merck KGaA	Darmstadt, Germany
MicroAmp® 8-Cap Strip	Life Technologies	Darmstadt, Germany
2100 Bioanalyzer	Agilent Technologies	Santa Clara, USA

Annex 8: Software and databases for the analyses of the generated data

Material	Manufacturer/reference	Location
MIRU-VNTR*plus*	Weniger et al.[116]	Münster, Germany
Genemapper v4.1	Life Technologies	Darmstadt, Germany
BioNumerics v7.5	Applied Maths	Gent, Belgium
PhyResSE v1.0	Feuerriegel et al.[43]	Borstel, Germany
FigTree v1.4.2	University of Edinburgh	Edinburgh, Scotland
KaKs Calculator v1.2	Beijing Institute of Genomics	Beijing, China
MLVA Compare v1.03	Ridom	Münster, Germany
Prism v5	GraphPad	La Jolla, USA
IBM® SPSS	IBM Analytics	Chicago, Illinois, USA
Epi Tools	Lawrence et al.[232]	

Annex 9: Number of smear, culture, and PCR positive sputum samples collected from Eastern Sudan and the positive predictive value of smear microscopy

Smear score at NRL	Total	Culture positive for Mycobacteria		Total	PCR positive for Mycobacteria		Total	Culture or PCR positive for Mycobacteria		Total	Culture or PCR positive for MTBC	
		No.	% (95% CI)		No.	% (95% CI)		No.	% (95% CI)		No.	% (95% CI)
Negative	123	31	25.2 (17.5-32.9)	82	49	59.8 (48.8-70.8)	113	80	71.0 (63.4-79.4)	113	70	62.0 (53.0-71.0)
Scanty	110	68	62.0 (52.9-71.1)	38	37	97.4 (95.8-99.8)	106	105	99.1 (98.3-99.9)	106	98	92.5 (87.5-97.5)
1+	77	43	55.8 (44.8-66.8)	34	34	100 (0.00-0.00)	77	77	100 (0.00-0.00)	77	74	96.0 (92.0-100)
2+	49	34	69.4 (56.4-82.4)	12	12	100 (0.00-0.00)	46	46	100 (0.00-0.00)	46	44	96.0 (93.2-99.2)
3+	24	20	83.3 (68.3-98.3)	3	3	100 (0.00-0.00)	23	23	100 (0.00-0.00)	23	22	96.0 (93.8-99.8)
Total	383	196	51.2 (46.2-56.2)	169	135	80.0 (74.0-86.0)	365	331	91.0 (87.4-95.4)	365	308	84.4 (80.7-88.1)



Annex 10: The phylogenetic diversity of MTBC strains. A UPGMA tree was calculated based on 24-loci MIRU-VNTR and spoligotyping patterns of 177 strains. **(1)** The tree shows that the strains belong to 4 different lineages from L 1 to 4. **(2)** Phenotypic DST to first line drugs. **(3)** Geographical location of sample collection.

Annex 11: Overview of the 177 MTBC strains identified in this study

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
1	12504/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
2	12142/16	S	WT	R	katG Ser315Asn (agc/aAc)	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
3	13851/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
4	12494/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
5	14382/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
6	14379/14	nd	WT	nd	WT	nd	WT	nd	WT	nd	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
7	12136/16	R	gidB Ser136_ (tca/tAa)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
8	13852/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
9	12417/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
10	12414/16	S	WT	S	WT	S	WT	S	WT	S	WT	Sudan like-H37Rv	Sudan H37Rv-like	Not CL	Not Gr
11	13723/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (ggc/gAc), embB Gly406Asp (ggc/gAc)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
12	12152/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
13	12797/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
14	14376/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
15	14371/14	S	WT	S	WT	S	WT	S	WT	S	WT	Haarlem	Haarlem	Not CL	Not Gr
16	12134/16	R	gidB Ala138Val (gcg/gTg)	R	katG Ser315An (agc/aCc)	R	rpoB His445Asn (cac/Cac)	R	embB Gln497Arg (cag/cGg)	S		Delhi/CAS	Delhi/CAS	Not CL	Not Gr
17	12749/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
18	13970/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
19	12283/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
20	12500/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
21	13591/14	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	Not CL	Not Gr
22	12317/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
23	12290/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
24	13687/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
25	12518/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
26	12172/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr

Annex 11 continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
27	14016/14	S	WT	S	WT	S	WT	S	WT	S	WT	LAM	LAM	CL	Gr
28	12752/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
29	14424/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
30	14017/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
31	12308/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
32	14372/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
33	13690/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
34	12120/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
35	12525/16	S	WT	S	WT	S	WT	S	WT	S	WT	UgandaII	Uganda	Not CL	Not Gr
36	12793/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
37	12274/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
38	13971/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
39	14013/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
40	12275/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
41	12796/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
42	12512/16	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	Not CL	Not Gr
43	12031/16	S	WT	S	WT	S	WT	S	WT	S	WT	TUR	Euro-American	Not CL	Not Gr
44	13688/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Gr
45	14377/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
46	12302/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
47	12807/16	R	gidB (4408017_GAP)	S	WT	S	WT	S	WT	S	WT	UgandaII	Uganda	Not CL	Not Gr
48	12306/16	S		S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
49	14127/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
50	13585/14	R	WT	R	WT	S	WT	R	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	
51	12124/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
52	12496/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
53	12759/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
54	12294/16	S	WT	S	WT	S	WT	S	WT	S	WT	EAI	EAI	CL	Gr
55	13860/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
56	13973/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
57	12171/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
58	12505/16	S	WT	S	WT	S	WT	S	WT	S	WT	Haarlem	Haarlem	Not CL	Not Gr
59	12744/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr

Annex 11 continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
60	12159/16	S	WT	S	WT	S	WT	S	WT	S	WT	EAI	EAI	CL	Gr
61	13718/14	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	Not CL	Not Gr
62	13855/14	R	rpsL Lys43Arg (aag/aGg)	R	fabG1 Thr4I so (aca/aTa)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
63	12735/16	S	WT	S	WT	S	WT	S	WT	S	WT	S-type	S-type	CL	Gr
64	12309/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
65	12149/16	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
66	14373/14	S	WT	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
67	13720/14	nd	WT	nd	WT	nd	WT	nd	WT	nd	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
68	13686/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
69	14019/14	R	rrs (---) (a/C)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
70	12119/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
71	12296/16	S	WT	S	WT	S	WT	S	WT	S	WT	UgandaII	Uganda	CL	Gr
72	12515/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
73	12811/16	R	rpsL Lys43Arg (aag/aGg)	S	WT	S	WT	S	WT	S	WT	Beijing	Beijing	Not CL	Not Gr
74	12794/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
75	12432/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
76	12276/16	S	WT	S	WT	S	WT	S	WT	S	WT	Sudan like H37Rv	Sudan H37Rv-like	CL	Gr
77	14384/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
78	13975/14	S	WT	S	WT	S	WT	S		S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
79	13586/14	R	rpsL Lys88Met (aag/aTg)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
80	13972/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
81	14419/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
82	12174/16	S	WT	S	WT	S	WT	S	WT	S	WT	X-type	Euro-American	Not CL	Not Gr
83	12434/16	S	WT	S	WT	S	WT	S	WT	S	WT	TUR	Euro-American	Not CL	Not Gr
84	12790/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr

Annex 11 continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
85	12808/16	S	WT	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Asp (cac/Gac)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
86	12724/16	S	WT	S	WT	S	WT	S	WT	S	WT	EAI	EAI	Not CL	Not Gr
87	12288/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
88	12040/16	S		S	WT	S	WT	S	WT	S	WT	Cameroon	Cameroon	Not CL	Not Gr
89	14121/14	R	gidB (4408116_ GAP)	S	WT	S	WT	S	WT	S	WT	X-type	X-type	Not CL	Not Gr
90	13724/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
91	13693/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
92	14124/14	S	WT	S	WT	S	WT	S	WT	S	WT	Sudan like H37Rv	Sudan H37Rv-like	CL	Gr
93	12146/16	S	WT	S	WT	S	WT	S	WT	S	WT	LAM	LAM	Not CL	Not Gr
94	12286/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
95	12805/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
96	12827/16	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	R	embB Met306Val (atg/Gtg)	R	pncA Gln10Arg (cag/cGg)	Delhi/CAS	Delhi/CAS	CL	Gr
97	12733/16	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	CL	Not Gr
98	12422/16	R	rpsL Lys43Arg (aag/aGg)	S	WT	S	WT	S	WT	S	WT	X-type	Euro-American	Not CL	Not Gr
99	12158/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
100	14378/14	R	gidB Iso114Ser (atc/aGc)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
101	13857/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
102	13178/14	S		S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
103	14118/14	R	rrs (---) (a/C)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
104	12025/16	S	WT	S	WT	S	WT	S	WT	S	WT	LAM	LAM	CL	Not Gr
105	12281/16	S	WT	S	WT	S	WT	S	WT	S	WT	X-type	X-type	Not CL	Not Gr
106	12429/16	R	gidB (4408101_ GAP)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr

Annex 11 continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
107	12803/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
108	12814/16	S	WT	S	WT	S	WT	S	WT	S	WT	Sudan like H37Rv	Euro-American	Not CL	Not Gr
109	12740/16	R	gidB Phe12Ser (ttc/tCc)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
110	12311/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
111	14375/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
112	14010/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser450Leu (tcg/tTg)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
113	13590/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
114	13719/14	R	WT	R	katG Ser315Thr (agc/aCc)	R	rpoB Ser431Thr (agc/aCc)	S	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	un
115	14418/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
116	12032/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
117	12279/16	S	WT	S	WT	S	WT	S	WT	S	WT	Haarlem	Haarlem	Not CL	Not Gr
118	12436/16	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	Not CL	Not Gr
119	12829/16	S	WT	S	WT	S	WT	S	WT	S	WT	TUR	Euro-American	Not CL	Not Gr
120	12822/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
121	12727/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
122	12170/16	S	WT	S	WT	S	WT	S	WT	S	WT	Beijing	Beijing	Not CL	Not Gr
123	14381/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
124	14122/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
125	13583/14	R	rrs (---) (a/C)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
126	13173/14	S	WT	S	WT	S	WT	S	WT	S	WT	LAM	LAM	CL	Gr
127	12802/16	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	CL	Not Gr
128	12739/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
129	12495/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
130	12310/16	R	gidB Arg39Pro (cgc/cCc)	S	WT	S	WT	S	WT	S	WT	TUR	Euro-American	Not CL	Not Gr
131	12280/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr

Annex 11 continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
137	13584/14	S	WT	S	WT	S	WT	S	WT	S	WT	Sudan H37Rv-like	Sudan H37Rv-like	CL	Gr
132	12150/16	S	WT	S	WT	S	WT	S	WT	S	WT	LAM	LAM	CL	Not Gr
133	14425/14	nd	WT	nd	WT	S	WT	nd	WT	S	WT	EAI	EAI	Not CL	
134	14374/14	R	rrs (---) (a/C)	S	WT	R	rpoB His445Tyr (cac/Tac)	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
135	13976/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
136	13721/14	R	gidB Try45_ (tgg/tAg)	R	fabG1 (---) (c/T)	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
138	13685/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
139	13858/14	S	WT	S	WT	S	WT	S	WT	S	WT	EAI	EAI	Not CL	Not Gr
140	14018/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
141	14380/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
142	12036/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
143	12272/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
144	12291/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
145	12424/16	R	gidB Ser136_ (tca/tAa)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
146	12514/16	S	WT	S	WT	S	WT	S	WT	S	WT	UgandaII	Uganda	CL	Gr
147	12758/16	S	WT	S	WT	S	WT	S	WT	S	WT	Cameroon	Cameroon	Not CL	Not Gr
148	12810/16	R	rpsL Lys88Arg (aag/aGg)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
149	12824/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
150	12800/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
151	12723/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
152	12433/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
153	12305/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
154	12285/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
155	12147/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
156	14421/14	R	rpsL Lys43Arg (aag/aGg)	R	katG Ser315Thr (agc/aCc)	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
157	14370/14	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	Not CL	Not Gr
158	14015/14	R	rpsL Lys43Arg (aag/aGg)	S	WT	R	rpoB His445Tyr (cac/Tac)	R	embB Met306Ile (atg/atA)	S	WT	Delhi/CAS	Delhi-CAS	CL	Gr
159	13588/14	S	WT	S	WT	S	WT	S	WT	S	WT	LAM	LAM	Not CL	Not Gr

Annex 11 continued

No	Key	SM		INH		RMP		EMB		PZA		Genotype	Coll lineage	MIRU cluster	SNP group
		pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST	pDST	gDST				
160	13174/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
161	13694/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
162	13967/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
163	14125/14	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
164	14383/14	nd	WT	nd	WT	nd	WT	nd	WT	nd	WT	Delhi/CAS	Delhi/CAS	CL	Gr
165	12129/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
166	12165/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
167	12304/16	S	WT	S	WT	S	WT	S	WT	S	WT	S-type	S-type	CL	Gr
168	12431/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	Not CL	Not Gr
169	12519/16	S	WT	S	WT	S	WT	S	WT	S	WT	Not defined	Euro-American	Not CL	Not Gr
170	12751/16	S	WT	S	WT	S	WT	S	WT	S	WT	Haarlem	Haarlem	CL	Group 6
171	12817/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
172	12826/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
173	12804/16	R	rpsL Lys43Arg (aag/aGg)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
174	12741/16	R	rpsL Lys43Arg (aag/aGg)	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Not Gr
175	12312/16	S	WT	S	WT	S	WT	S	WT	S	WT	Haarlem	Haarlem	CL	Gr
176	12295/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr
177	12155/16	S	WT	S	WT	S	WT	S	WT	S	WT	Delhi/CAS	Delhi/CAS	CL	Gr

SM = streptomycin, INH = isoniazid, RMP = rifampicin, EMB = ethambutol, PZA = pyrazinamide, pDST = phenotypic drug susceptibility testing, gDST = genotypic drug susceptibility testing, R = resistant, S = susceptible, WT = wild-type, nd = not done, CL = clustered, and Gr = grouped

Annex 12: Number of mutations conferring drug resistance among the different identified genotypes by whole genome sequencing

Genotype	Total	Number of resistant strains					No. of mutations per genotype	Non-MDR per genotype	MDR per genotype
		SM	INH	RMP	EMB	PZA			
Delhi/CAS	129	31	20	20	14	1	**86**	16	18
EAI	4	0	0	0	0	0	**0**	0	0
LAM	6	0	0	0	0	0	**0**	0	0
Uganda	4	1	0	0	0	0	**1**	1	0
S-type	2	0	0	0	0	0	**0**	0	0
Haarlem	5	0	0	0	0	0	**0**	0	0
Sudan H37Rv-like	4	0	0	0	0	0	**0**	0	0
X-type	2	1	0	0	0	0	**1**	1	0
Cameroon	2	0	0	0	0	0	**0**	0	0
Euro-American	15	2	0	0	0	0	**2**	2	0
Beijing	1	1	0	0	0	0	**1**	1	0
Total	**175**	**36**	**20**	**20**	**14**	**1**	**91**	**21**	**18**

SM = streptomycin, INH = isoniazid, RMP = rifampicin, EMB = ethambutol, PZA = pyrazinamide No. = number, CAS = Central Asian Strain, EAI = East-African-Indian, LAM = Latin American Mediterranean, Non-MDR = resistance to any anti-TB drug other than INH and RMP at once, and MDR = resistance to at least INH and RMP.

Annex 13: List of the studies and contacts included in the Delhi/CAS meta-analysis

No.	Country of study	Sampling period	No. of strains	MIRU done in	DST data	Institute/Researcher	Sampling method/ Study type	Email address
1.	Cameroon	2010-2011	3	Borstel	available	V. Penlap	Population based	v.penlap@yahoo.fr
2.	Congo	2011	2	Borstel	available	V. Penlap	Convenient sampling	v.penlap@yahoo.fr
3.	Ethiopia	2009	97	Borstel	available	B. Tessema	Cross sectional	bt1488@yahoo.com
4.	Ethiopia	2013	19	Borstel	available	S. Ali	Cross sectional	solomon.ali@ju.edu.et
5.	Ethiopia	2012	78	Borstel	unavailable	B. Fantahun	Cross sectional	fantahun.degeneh@gmail.com
6.	Ethiopia	2016	50	Borstel	available	E. Abera	MDR cohort	eyob2001@gmail.com
7.	Baden-Württemberg	2008-2011	30	Genoscreen	unavailable	E. Göhring-Zwacka	Population based	elke.goehring-zwacka@rps.bwl.de
8.	Frankfurt	2008-2013	59	NA	unavailable	DBgermanbase	Population based	udo.goetseh@stadt-frankfurt.de
9.	Hamburg	1998-2015	64	Borstel	available	DBgermanbase	Population based	roltatand.1480@hotmail.com
10.	Hannover	2001-2014	42	NA	unavailable	DBgermanbase	Population based	ange.hanke-lensing@region-hannover.de
11.	MUT	2003-2007	28	Genoscreen	unavailable	E. Göhring-Zwacka	Population based	elke.goehring-zwacka@rps.bwl.de
12.	Schleswig-Holstein	2010-2013	9	Borstel	unavailable	S. Schuback	Population based	Schuback@steinburg.de
13.	Germany	1998-2014	169	Borstel	not for all	DBgermanbase and DBresbase	MDR cohort	NA
14.	Ghana	2001-2014	23	Genoscreen	available	R. Horstmann	Cross sectional	horstmann@bnitm.de
15.	Iran	2002	1	Borstel	available	DBgermanbase	MDR cohort	issam788@yahoo.com
16.	Iraq	2010-2012	64	NA	available	A. Issam	NA	issam788@yahoo.com
17.	Kenya	2010	59	Borstel	available	P. Ndung'u	Cross sectional	perpetualndungu@yahoo.com
18.	Kenya	2014	50	Borstel	unavailable	G. Kerubo	Cross sectional	glenakerubo@gmail.com
19.	Mozambique	2014	4	NA	unavailable	A. Rachow	Cross sectional	rachow@lrz.uni-muenchen.de
20.	Pakistan	2012	122	Borstel	not for all	S. Noor	Convenient sampling	khattak799@gmail.com
21.	Senegal	2013	10	Borstel	unavailable	A. Diallo	Convenient sampling	awa1.diallo@ucad.edu.sn
22.	Sudan	2014-2016	130	Borstel	available	Y. A. Shuaib	Cross sectional	vet.aboamar@gmail.com

Annex 13 continued

No.	Country of study	Sampling period	No. of strains	MIRU done in	DST data	Institute/Researcher	Sampling method/ Study type	Email address
23.	Swaziland	2009	3	Borstel	available	E. Sanchez	Cross sectional	elisabeth.sanchez@epicentre.msf.org
24.	Sweden	NA	96	NA	NA	NA	NA	NA
25.	Tanzania	2008-2015	86	NA	NA	M. Hoelscher	Cross sectional	hoelscher@lrz.uni-muenchen.de
26.	Turkmenistan	2001-2008	18	Genoscreen	not for all	DBostbase	Cross sectional	helen.cox@uct.ac.za kai.braker@berlin.msf.org
27.	Uzbekistan	2004-2006	3	Genoscreen	available	H. Cox	MDR cohort	helen.cox@uct.ac.za kai.braker@berlin.msf.org
28.	Zambia	2014	1	NA	NA	A. Rachow	Cross sectional	rachow@lrz.uni-muenchen.de
29.	India	2010-2011	73	Borstel	available	M. Basil	Convenient sampling	mandirav2001@yahoo.com
30.	India	2006	1	NA	available	DBgermanbase	NA	NA
31.	India	2010-2013	30	India	available	K. Narain	Cross sectional	kanwarnarain@hotmail.com
32.	Saudi Arabia	2009-2010	61	SA	available	S. Al-Hajoj	National drug surveillance	hajoj@kfshrc.edu.sa
33.	Saudi Arabia	NA	69	SA	available	S. Al-Hajoj	NA	hajoj@kfshrc.edu.sa
34.	Djibouti	2009	16	Djibouti	unavailable	S. Godreuil	Cross sectional	godreuil@yahoo.fr
35.	Ireland	2010-2011	42	Ireland	available	M. Fitzgibbon	Population based	mfitzgibbon@stjames.ie
36.	No data	NA	92	NA	unavailable	- P. Supply 1 sample - DBgermanbase 4 samples - DBresbase 1 sample	NA	NA

No. = number and NA = not available

Annex 14: List of the publications included in the Delhi/CAS meta-analysis

Study no.	Reference	doi or PMID
1.	Not published	NA
2.	Malm S., Linguissi L. S., Tekwu E. M., Vouvoungui J. C., Kohl T. A., Beckert P., Sidibe A., Rüsch-Gerdes S., Madzou-Laboum I. K., Kwedi S., Penlap Beng V., Frank M., Ntoumi F., Niemann S. (2017): New *Mycobacterium tuberculosis* Complex Sublineage, Brazzaville, Congo. *Emerging infectious diseases*, 23(3), 423-429.	10.3201/eid2303.160679
3.	Tessema B., J. Beer, M. Merker, F. Emmrich, U. Sack, A. C. Rodloff and S. Niemann (2013): Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia. *BMC Infectious Diseases*, 13, 131.	10.1186/1471-2334-13-131
4.	Ali S., P. Beckert, A. Haileamlak, A. Wieser, M. Pritsch, N. Heinrich, T. Löscher, M. Hoelscher, S. Niemann and A. Rachow (2016): Drug resistance and population structure of *Mycobacterium tuberculosis* strains from prisons and communities in Ethiopia. BMC Infectious Diseases, 16, 687.	https://doi.org/10.1186/s12879-016-2041-x
5.	Biadglegne F., Merker M., Sack U., Rodloff A. C., Niemann S. (2015): Tuberculous Lymphadenitis in Ethiopia Predominantly Caused by Strains Belonging to the Delhi/CAS Lineage and Newly Identified Ethiopian Clades of the *Mycobacterium tuberculosis* Complex. PLoSOne, 10(9): e0137865	10.1371/journal.pone.0137865
6.	Not published	NA
7.	Not published	NA
8.	NA	NA
9.	NA	NA
10.	NA	NA
11.	Not published	NA
12.	NA	NA
13.	NA	NA
14.	NA	NA
15.	Not published	NA
16.	NA	NA
17.	Ndung'u P.W., Kariuki S., Ng'ang'a Z., and Revathi G. (2012): Resistance patterns of *Mycobacterium tuberculosis* strains from pulmonary tuberculosis patients in Nairobi. J. Infect. Dev. Ctries., 6(1): 33-39.	22240426
18.	Not published	NA
19.	NA	NA
20.	NA	NA
21.	NA	NA

Annex 14 continued

Study no.	Reference	doi or PMID
22.	Shuaib Y. A., Merker M., Khalil E. A., Ulrich E. S., Wieler L. H., Bakheit M. A., Mohamed-Noor S. E., Abdalla M. A., Andres A., Hillemann D., Richter E., Kranzer K., and Niemann S. (2019): *Mycobacterium tuberculosis* complex lineage 3 strains drive the pulmonary tuberculosis epidemic in Eastern Sudan. In preparation.	NA
23.	Not published	NA
24.	NA	NA
25.	Not published	NA
26.	Not published	NA
27.	Not published	NA
28.	NA	NA
29.	Varma-Basil M., Narang A., Chakravorty S., Garima K., Gupta S., Kumar Sharma N., Giri A., Zozio T., Couvin D., Hanif M., Bhatnagar A., Menon B., Niemann S., Rastogi N., Alland D., and Bose M. (2016): A snapshot of the predominant single nucleotide polymorphism cluster groups of *Mycobacterium tuberculosis* clinical strains in Delhi, India. *Tuberculosis (Edinb)*, 100: 72-81.	10.1016/j.tube.2016.07.007
30.	Varghese B., Supply P., Allix-Béguec C., Shoukri M., Al-Omari R., Herbawi M., and Al-Hajoj, S. (2013): Admixed Phylogenetic Distribution of Drug Resistant *Mycobacterium tuberculosis* in Saudi Arabia. *PLosOne*, *8*(2), e55598.	10.1371/journal.pone.0055598
31.	Devi K. R., Bhutia R., Bhowmick S., Mukherjee K., and M. J., Narain K (2015) Genetic Diversity of *Mycobacterium tuberculosis* Strains from Assam, India: Dominance of Beijing Family and Discovery of Two New Clades Related to CAS1_Delhi and EAI Family Based on Spoligotyping and MIRU-VNTR Typing. *PLosOne*, 10(12): e0145860.	10.1371/journal.pone.0145860
32.	Al-Hajoj, S., Varghese, B., Shoukri, M. M., Al-Omari, R., Al-Herbwai, M., Alrabiah, F., Alrajhi, A. A., Abuljadayel, N., Al-Thawadi, S., Zumla, A., Zignol, M., Raviglione, M. C., and Memish, Z. (2013). Epidemiology of antituberculosis drug resistance in Saudi Arabia: findings of the first national survey. *Antimicrobial agents and chemotherapy*, *57*(5), 2161-2166.	10.1128/AAC.02403-12
33.	NA	NA
34.	Godreuil S., F. Renaud, M. Choisy, J. J. Depina, E. Garnotel, M. Morillon, P. Van de Perre and A. L. Banuls (2010): Highly structured genetic diversity of the *Mycobacterium tuberculosis* population in Djibouti. *Clin. Microbiol. Infect.*, 16, 1023–1026.	10.1111/j.14690691.2009.03025.x
35.	Fitzgibbon M. M., Gibbons N., Roycroft E., Jackson S., O'Donnell J., O'Flanagan D., Rogers T. R. (2013): A snapshot of genetic lineages of *Mycobacterium tuberculosis* in Ireland over a two-year period, 2010 and 2011. *Euro. Surveill.*, 18(3), pii=20367.	23351653

No. = number and NA = not available

Annex 15: The identified MLVA 15-9 types of the global collection of clinical Delhi/CAS strains

MLVA type	No.	MLVA type	No.	MLVA type	No.	MLVA type	No.	MLVA type	No.
1064-32	81	2218-15	5	1422-25	3	12564-88	2	17366-32	2
1557-32	68	317-32	5	1449-145	3	12569-32	2	17403-32	2
1061-32	33	4536-32	5	1457-32	3	12784-15	2	18482-32	2
1220-15	27	7122-15	5	15106-32	3	12786-32	2	18969-32	2
4534-32	22	9048-332	5	17165-32	3	12788-32	2	2213-15	2
1212-32	21	**?-387**	**4**	17280-32	3	12796-32	2	2215-32	2
1449-32	20	**?-914**	**4**	17363-257	3	12855-25	2	3397-32	2
?-15	**20**	12525-32	4	17651-32	3	12928-32	2	345-32	2
2021-32	18	12560-32	4	17938-32	3	13016-32	2	3526-32	2
1422-32	16	1320-32	4	18513-15	3	13018-88	2	3612-145	2
1212-25	14	1491-25	4	2136-32	3	13024-288	2	3612-193	2
1295-32	13	3354-15	4	3633-32	3	13027-32	2	3625-32	2
?-122	**10**	3612-32	4	3644-32	3	13031-419	2	3627-32	2
?-189	**9**	3630-32	4	367-32	3	13084-32	2	3635-32	2
12760-32	9	3667-15	4	4538-32	3	13113-32	2	3649-32	2
1557-419	9	5085-32	4	4539-32	3	1344-229	2	3668-15	2
3549-145	9	6755-32	4	5070-25	3	1422-493	2	4535-32	2
1649-32	8	8966-32	4	7186-32	3	1451-32	2	5358-32	2
4439-32	8	9064-32	4	8738-32	3	14680-32	2	5847-32	2
12794-32	7	9072-32	4	9063-32	3	14887-32	2	6690-32	2
145-62	7	9928-32	4	9139-32	3	1557-229	2	6974-32	2
9049-32	7	**?-135**	**3**	9259-32	3	1557-25	2	7210-229	2
?-25	**6**	**?-63**	**3**	9714-32	3	15862-32	2	8301-34	2
?-82	**6**	10425-32	3	**?-145**	**2**	15863-32	2	9045-175	2
12306-25	6	1064-145	3	**?-288**	**2**	1643-32	2	9046-32	2
12792-32	6	1064-15	3	**?-394**	**2**	16716-32	2	9060-32	2
12904-32	6	1064-25	3	**?-88**	**2**	16723-32	2	9061-32	2
4727-32	6	11529-32	3	10024-32	2	16730-25	2	9070-32	2
7785-32	6	12306-32	3	1062-32	2	16731-32	2	9075-32	2
?-1039	**5**	12535-32	3	11471-32	2	17104-32	2	9635-257	2
1231-32	5	12557-32	3	12087-32	2	1719-15	2	9685-32	2
13028-32	5	13018-32	3	1219-32	2	17228-32	2	9762-32	2
13036-32	5	13025-32	3	12410-15	2	17286-15	2	9773-32	2
1534-145	5	13208-32	3	12530-32	2	17328-34	2	9812-32	2

Annex 16: Mutations of Delhi/CAS SNP group 3.2 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
pykA	Rv1617	essential	1816370	Probable pyruvate kinase PykA	gcg/gTg	A61V
nrp	Rv0101	essential	113872	Probable peptide synthetase Nrp (peptide synthase)	ccc/cTc	P1291L
rpsA	Rv1630	essential	1834159	30S ribosomal protein S1 RpsA	ttg/ttA	L206L
dlaT	Rv2215	essential	2483579	DlaT, dihydrolipoamide acyltransferase, E2 component of pyruvate dehydrogenase	aag/Gag	K539E
ptrBa	Rv0781	essential	874323	Probable protease II PtrBa [first part] (oligopeptidase B)	tat/Aat	Y31N
Rv2038c	Rv2038c	essential	2283881	Probable sugar-transport ATP-binding protein ABC transporter	tat/Cat	Y306H
ppnK	Rv1695	essential	1919627	Inorganic polyphosphate/ATP-NAD kinase PpnK (poly(P)/ATP NAD kinase)	gtg/gtA	V294V
rplN	Rv0714	essential	811621	50S ribosomal protein L14 RplN	gcc/gcT	A83A
pks7	Rv1661	essential	1881051	Probable polyketide synthase Pks7	gtg/gtC	V1916V
aceE	Rv2241	essential	2512641	Pyruvate dehydrogenase E1 component AceE (pyruvate decarboxylase) (pyruvate dehydrogenase) (pyruvic dehydrogenase)	ctg/Ttg	L35L
embB	Rv3795	essential	4249195	Integral membrane indolylacetylinositol arabinosyltransferase EmbB (arabinosylindolylacetylinositol synthase)	ccc/ccA	P894P
Rv3649	Rv3649	essential	4090890	Probable helicase	gcc/Tcc	A704S
purC	Rv0780	essential	873464	Phosphoribosylaminoimidazole-succinocarboxamide synthase PurC (SAICAR synthetase)	gtc/gCc	V41A
Rv1272c	Rv1272c	essential	1421398	Probable drugs-transport transmembrane ATP-binding protein ABC transporter	tct/tTt	S303F
aftD	Rv0236c	essential	283496	Possible arabinofuranosyltransferase AftD	tcg/tTg	S1119L
Rv1422	Rv1422	essential	1597190	hypothetical protein	ctg/Ttg	L104L
lprK	Rv0173	essential	204866	Possible Mce-family lipoprotein LprK (Mce-family lipoprotein Mce1E)	tat/Gat	Y268D

Annex 17: Mutations of Delhi/CAS SNP group 3.3 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
fadB3	Rv1715	essential	1943460	Probable 3-hydroxybutyryl-CoA dehydrogenase FadB3 (beta-hydroxybutyryl-CoA dehydrogenase) (BHBD)	acc/Gcc	T268A
mrsA	Rv3441c	essential	3860889	Probable phospho-sugar mutase / MrsA protein homolog	cgc/cAc	R161H
Rv1332	Rv1332	essential	1501153	Probable transcriptional regulatory protein	acc/acT	T76T
nusA	Rv2841c	essential	3149111	Probable N utilization substance protein A NusA	gag/gaC	E106D
parA	Rv3918c	essential	4407292	Probable chromosome partitioning protein ParA	cgg/cgA	R80R
recN	Rv1696	essential	1919901	Probable DNA repair protein RecN (recombination protein N)	ctc/ctT	L73L
lysX	Rv1640c	essential	1851932	Lysyl-tRNA synthetase 2 LysX	tgc/tTc	C35F
rimM	Rv2907c	essential	3216584	Probable 16S rRNA processing protein RimM	tat/tGt	Y103C
nrp	Rv0101	essential	110567	Probable peptide synthetase Nrp (peptide synthase)	gac/gaT	D189D

Annex 17 continued

Gene	Rv number	Category	Position	Annotation	SNP	AA
accA2	Rv0973c	essential	1084567	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (alpha subunit) AccA2: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	tac/tGc	Y395C
ugpC	Rv2832c	essential	3138742	Probable Sn-glycerol-3-phosphate transport ATP-binding protein ABC transporter UgpC	ggc/gTc	G147V
folD	Rv3356c	essential	3770067	Probable bifunctional protein FolD: methylenetetrahydrofolate dehydrogenase + methenyltetrahydrofolate cyclohydrolase	cgc/Tgc	R195C
eccD3	Rv0290	essential	353833	ESX conserved component EccD3 ESX-3 type VII secretion system protein Probable transmembrane protein	gcg/Acg	A251T
dnaE1	Rv1547	essential	1748125	Probable DNA polymerase III (alpha chain) DnaE1 (DNA nucleotidyltransferase)	ctc/ctT	L144L
canA	Rv1284	essential	1437497	Beta-carbonic anhydrase	aac/aaT	N58N
dnaE1	Rv1547	essential	1750602	Probable DNA polymerase III (alpha chain) DnaE1 (DNA nucleotidyltransferase)	gat/gGt	D970G
espI	Rv3876	essential	4354256	ESX-1 secretion-associated protein EspI Conserved proline and alanine rich protein	tcg/tTg	S416L
pks13	Rv3800c	essential	4260939	Polyketide synthase Pks13	ggg/Agg	G70R
caeA	Rv2224c	essential	2496797	Probable carboxylesterase CaeA	ccg/cAg	P76Q
proS	Rv2845c	essential	3151376	Probable prolyl-tRNA synthetase ProS (proline--tRNA ligase) (PRORS) (global RNA synthesis factor) (proline translase)	tcg/tcA	S525S
ugpC	Rv2832c	essential	3138256	Probable Sn-glycerol-3-phosphate transport ATP-binding protein ABC transporter UgpC	acg/aTg	T309M
pks13	Rv3800c	essential	4258023	Polyketide synthase Pks13	cgc/Tgc	R1042C
ugpC	Rv2832c	essential	3139011	Probable Sn-glycerol-3-phosphate transport ATP-binding protein ABC transporter UgpC	att/atG	I57M
trpD	Rv2192c	essential	2455785	Probable anthranilate phosphoribosyltransferase TrpD	gcc/gTc	A320V
fadB3	Rv1715	essential	1943454	Probable 3-hydroxybutyryl-CoA dehydrogenase FadB3 (beta-hydroxybutyryl-CoA dehydrogenase) (BHBD)	gcg/Tcg	A266S
fadB3	Rv1715	essential	1943459	Probable 3-hydroxybutyryl-CoA dehydrogenase FadB3 (beta-hydroxybutyryl-CoA dehydrogenase) (BHBD)	deletion	-

Annex 18: Mutations of Delhi/CAS SNP group 3.4 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
gltD	Rv3858c	essential	4330870	Probable NADH-dependent glutamate synthase (small subunit) GltD (L-glutamate synthase) (L-glutamate synthetase) (NADH-glutamate synthase) (glutamate synthase (NADH)) (GLTS beta chain) (NADPH-GOGAT)	cgg/cgT	R212R
infA	Rv3462c	essential	3880462	Probable translation initiation factor if-1 InfA	tcc/tcT	S64S

Annex 18 continued

Gene	Rv number	Category	Position	Annotation	SNP	AA
mrsA	Rv3441c	essential	3861095	Probable phospho-sugar mutase / MrsA protein homolog	gac/gaT	D92D
Rv1254	Rv1254	essential	1401782	Probable acyltransferase	gcc/gGc	A42G
wag31	Rv2145c	essential	2405010	Diviva family protein Wag31	gcc/gAc	A130D
Rv1204c	Rv1204c	essential	1347615	hypothetical protein	gcc/gTc	A337V
ribG	Rv1409	essential	1585248	Probable bifunctional riboflavin biosynthesis protein RibG : diaminohydroxyphosphoribosylaminopyrimidine deaminase (riboflavin-specific deaminase) + 5-amino-6-(5-phosphoribosylamino) uracil reductase (HTP reductase)	cac/Tac	H19Y
fadE23	Rv3140	essential	3507594	Probable acyl-CoA dehydrogenase FadE23	gca/Aca	A269T
rpoC	Rv0668	essential	764181	DNA-directed RNA polymerase (beta' chain) RpoC (transcriptase beta' chain) (RNA polymerase beta' subunit)	gac/gGc	D271G
rpsB	Rv2890c	essential	3199764	30S ribosomal protein S2 RpsB	aag/aaA	K73K
mprB	Rv0982	essential	1098984	Two component sensor kinase MprB	gtt/Att	V493I
Rv0226c	Rv0226c	essential	270889	Probable conserved transmembrane protein	gcc/Tcc	A226S
otsB2	Rv3372	essential	3787421	Trehalose 6-phosphate phosphatase OtsB2 (trehalose-phosphatase) (TPP)	gca/Cca	A370P

Annex 19: Mutations of Delhi/CAS SNP group 3.5 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
pgi	Rv0946c	essential	1056469	Probable glucose-6-phosphate isomerase Pgi (GPI) (phosphoglucose isomerase) (phosphohexose isomerase) (phi)	cgg/Tgg	R73W
pks13	Rv3800c	essential	4258550	Polyketide synthase Pks13	gag/gCg	E866A
accD1	Rv2502c	essential	2818085	Probable acetyl-/propionyl-CoA carboxylase (beta subunit) AccD1	atg/atC	M130I
Rv0175	Rv0175	essential	207341	Probable conserved Mce associated membrane protein	gtc/gtA	V176V
lysX	Rv1640c	essential	1851835	Lysyl-tRNA synthetase 2 LysX	cct/ccG	P67P
thrC	Rv1295	essential	1451491	Threonine synthase ThrC (ts)	cag/caA	Q265Q
Rv3400	Rv3400	essential	3817449	Probable hydrolase	gcg/Acg	A71T
hadA	Rv0635	essential	732145	(3R)-hydroxyacyl-ACP dehydratase subunit HadA	ttc/ttT	F72F
cobL	Rv2072c	essential	2329554	Precorrin-6Y C(5,15)-methyltransferase (decarboxylating) CobL	gtg/gCg	V198A
ctpH	Rv0425c	essential	512610	Possible metal cation transporting P-type ATPase CtpH	cgg/cgA	R904R
ispF	Rv3581c	essential	4023949	Probable 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase IspF (MECPS)	acc/acG	T133T
Rv0060	Rv0060	essential	64444	hypothetical protein	ctc/cAc	L179H
Rv3910	Rv3910	essential	4398122	Probable conserved transmembrane protein	cac/cCc	H509P
mihF	Rv1388	essential	1563784	Putative integration host factor MihF	cgc/Tgc	R31C

Annex 20: Mutations of Delhi/CAS SNP group 3.1.1 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
accA2	Rv0973c	essential	1084911	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (alpha subunit) AccA2: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	tac/taT	Y280Y
Rv1466	Rv1466	essential	1653374	hypothetical protein	ggg/ggC	G48G
Rv3660c	Rv3660c	essential	4098522	hypothetical protein	ggg/ggC	G209G
murF	Rv2157c	essential	2418560	Probable UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate-D-alanyl-D-alanyl ligase MurF	gcc/Acc	A149T
Rv2897c	Rv2897c	essential	3207373	hypothetical protein	gac/gaA	D190E
Rv0347	Rv0347	essential	417404	Probable conserved membrane protein	cgt/cTt	R34L

Annex 21: Mutations of Delhi/CAS SNP group 3.1.2 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
tyrS	Rv1689	essential	1914217	Probable tyrosyl-tRNA synthase TyrS (TYRRS)	cgg/Agg	R206R
trpS	Rv3336c	essential	3722702	Probable tryptophanyl-tRNA synthetase TrpS (tryptophan--tRNA ligase) (TRPRS) (tryptophan translase)	ctc/ctG	L310L

Annex 22: Mutations of Delhi/CAS SNP group 3.6 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
mesJ	Rv3625c	essential	4064291	Possible cell cycle protein MesJ	gaa/gaC	E194D
papA3	Rv1182	essential	1321096	Probable conserved polyketide synthase associated protein PapA3	tcg/tcA	S354S
lpqZ	Rv1244	essential	1387478	Probable lipoprotein LpqZ	gac/Cac	D208H

Annex 23: Mutations of Delhi/CAS SNP group 3.2 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
gdh	Rv2476c	epitope	2779824	Probable NAD-dependent glutamate dehydrogenase Gdh (NAD-Gdh) (NAD-dependent glutamic dehydrogenase)	cgc/cgT	R813R
Rv0203	Rv0203	epitope	241758	Possible exported protein	ggg/gAg	G82E
Rv2823c	Rv2823c	epitope	3131177	hypothetical protein	ctc/ctT	L199L
Rv2223c	Rv2223c	epitope	2494494	Probable exported protease	gtc/gtT	V302V

Annex 24: Mutations of Delhi/CAS SNP group 3.5 in comparison to H37Rv genome

Gene	Rv number	Category	Position	Annotation	SNP	AA
gdh	Rv2476c	epitope	2781290	Probable NAD-dependent glutamate dehydrogenase Gdh (NAD-Gdh) (NAD-dependent glutamic dehydrogenase)	gtg/Atg	V325M
Rv1945	Rv1945	epitope	2196405	hypothetical protein	deletion	-

7 List of publications

1. Arede M., …, **Shuaib Y. A.** (2019): Combating Vaccine Hesitancy: Teaching the Next Generation to Navigate Through the Post Truth Era. *Front. Public Health,* **6(381)**.

2. **Shuaib Y. A.** *et al.* (2018): Smear microscopy for diagnosis of pulmonary tuberculosis in Eastern Sudan. *Tuberculosis Research and Treatment*, **2018**, Article ID 8038137.

3. Eman M. A., …, **Shuaib Y. A.** (2018): Seroprevalence and risk factors of caprine brucellosis in Khartoum state, the Sudan. *Veterinary World*, **11(4)**, 511-518.

4. Khuzaima B. A., …, **Shuaib Y. A.** *et al.* (2018): Sero-prevalence of anti-brucella antibodies in goats in El-Gedarif state, Eastern Sudan. *ARC J. of Anim. Vet. Sci.*, **4(1)**, 1-8.

5. **Shuaib Y. A.** *et al.* (2017): Mycobacterial infections in carcasses of ruminants slaughtered at the two slaughterhouses of Kassala, Sudan. *Rev. Elev. Med. Vet. Pays Trop.*, **7**, 131-136.

6. Abu-Samra M. T. and **Shuaib Y. A.** (2017a): Ultrastructure, life-cycle, survival and motility of *Demodex bovis* isolated from cattle in the Sudan. *ARC J. of Anim. Vet. Sci.*, **3(4)**, 11-18.

7. Abu-Samra M. T. and **Shuaib Y. A.** (2017b): Morphometric and morphologic characteristics of *Demodex bovis* and *Demodex ghanensis* isolated from cattle in the Sudan. *ARC J. of Anim. Vet. Sci.*, **3(4)**, 25-33.

8. Abu-Samra M. T. and **Shuaib Y. A.** (2017): High mortalities among one-humped camels due to salinomycin poisoning in the Kingdom of Saudi Arabia. *J. Adv. Vet. Anim. Res*., **4(2)**, 214-221.

9. Atail H. B., …, **Shuaib Y. A.** *et al.* (2017). Sero-prevalence of toxoplasmosis in sheep and goats in El-Gadarif state. *J. Adv. Vet. Anim. Res*., **4(2)**, 207-213.

10. Hamza I. I., **Shuaib Y. A.** *et al.* (2017): Aerobic bacteria isolated from internal lesions of camels at Tambool slaughterhouse. *J. Adv. Vet. Anim. Res*., **4(1)**, 22-31.

11. Ibrahim W. H., …, **Shuaib Y. A.** (2016): Isolated aerobic bacteria and fungi from skin lesions of fish in Khartoum state. *J. Adv. Vet. Anim. Res*., **3(4)**, 375-385.

12. Salih M. E., **Shuaib Y. A.** *et al.* (2016): Seroprevalence and risk factors of brucellosis in camels in and around Alzulfi, Saudi Arabia. *J. Camel Pract. Res.*, **23(2)**, 265-275.

13. **Shuaib Y. A.** *et al.* (2016): Bacteria associated with skin wounds of equines and their antimicrobial susceptibility in Sharg Alneel, Khartoum State. *Sch. J. Agric. Vet. Sci*., **3(2)**, 140-146.

14. Abu-Samra M. T., **Shuaib Y. A.** *et al.* (2016): Infectious bovine keratoconjunctivitis in Al-silaite area, Khartoum state. *Sch. J. Agric. Vet. Sci*., **3(3)**, 227-233.

15. Wegdan O. M., …, **Shuaib Y. A.** (2016). Seroprevalence and risk factors of anti-brucella antibodies in cattle in Khartoum State, the Sudan. *J. Adv. Vet. Anim. Res*., **3(2)**, 134-144.

16. **Shuaib Y. A.** *et al.* (2015): Seroprevalence of *Babesia Bigemina* Antibodies in Cattle in North Kordofan State, the Sudan. *ARC J. Anim. Vet. Sci.*, **1(2)**, 1-11.

17. **Shuaib Y. A.** *et al.* (2015): Evaluation of the Safety Status of Beef for Domestic Consumption during Processing at Three Slaughterhouses in Omdurman, the Sudan. *IOSR-JAVS*, **8(4)**, 22-27.

18. Abu-Samra M. T. and **Shuaib Y. A.** (2014): Pathology and pathogenesis of bovine skin and meibomian gland demodicosis. *Rev. Elev. Med. Vet. Pays Trop.*, **67(2)**, 77-85.

19. **Shuaib Y. A.** *et al.* (2014): Seroprevalence and risk factors of *Peste des petits ruminants* in sheep in the Kassala and North Kordofan States of the Sudan. *Inter. J. Vet. Sci.*, **3(1)**, 18-28.

Teile dieser Dissertation sind veröffentlicht in oder werden veröffentlicht

1. Shuaib Y. A., Niemann S., Khalil E. A., Schaible U., Wieler L. H., Bakheit M. A., Mohamed-Noor S. E., Abdalla M. A., and Richter E. (2017): Mycobacterial infections in carcasses of ruminants slaughtered at the two slaughterhouses of Kassala, Sudan. *Rev. Elev. Med. Vet. Pays Trop.*, **70 (4)**: 131-136. doi:10.19182/remvt.31530

2. Shuaib Y. A., Khalil E. A., Ulrich E. S., Wieler L. H., Bakheit M. A., Mohamed-Noor S. E., Abdalla M. A., Homolka S., Andres A., Hillemann D., Lonnroth K., Richter E., Niemann S., and Kranzer K. (2018): Smear Microscopy for Diagnosis of Pulmonary Tuberculosis in Eastern Sudan," *Tuberculosis Research and Treatment*, vol. 2018, Article ID 8038137, 8 pages, 2018. https://doi.org/10.1155/2018/8038137.

3. Shuaib Y. A., Khalil E. A. G., Ulrich E. S., Wieler L. H., Bakheit M. A., Mohamed-Noor S. E., Abdalla M. A., Andres A., Hillemann D., Richter E., Kranzer K., Niemann S., and Merker M. (2020): *Mycobacterium tuberculosis complex* lineage 3 drives the pulmonary tuberculosis epidemic in Eastern Sudan. ***Emerg. Infect. Dis.***, 26(3), 427-436. https://doi.org/10.3201/eid2603.191145

4. Shuaib Y. A., Merker M., Wieler L. H., …, and Niemann S. (2020): The global population structure of *Mycobacterium tuberculosis complex* lineage 3. **In preparation.**

8 Acknowledgements

Above all, I wish to thank Allah, as well as my parents and siblings, for their abundant blessings and graces throughout my life and during my PhD work.

I would like to express my sincerest appreciation and deepest thank to the Deutscher Akademischer Austausch Dienst (DAAD) and the Research Center Borstel in Germany, for their significant contribution and for providing the scholarship without which it would not have been possible to accomplish this PhD work. The contribution of the Dahlem Research School, Freie Universität Berlin, Germany, in accomplishing this work is appreciated.

I would like to express my sincere gratitude and appreciation to Herr Prof. Dr. S. Niemann for giving me the opportunity to pursue my PhD work under his supervision. I am very thankful and value his excellent guidance, support, and encouragement very much. I would also like to thank Frau PD Dr. E. Richter for her impactful guidance and support during the first 18 months of my PhD. I appreciate the continuous support of my co-supervisor Herr Prof. Dr. U. E. Schaible. I also thank Herr Prof. Dr. L. H. Wieler, my supervisor at Freie Universität Berlin, for his high interest in this work and for his continuous support.

I am indebted to Herr Prof. Dr. M. Merker and Herr Dr. S. Malm for their endless assistance and guidance through out my PhD work, during writing up of the thesis and for their time to proofreading the whole document. I deeply appreciate their effort and would like to thank them very much. I am thankful to my dear friends Herr Dr. M. Gröschel (my great and smart friend and office partner) and Herr Dr. P. Beckert for proofreading parts of my thesis and for the very valuable suggestions that made the thesis better and for all of the nice chats, laughter, and memories.

The help provided by the late Frau Prof. Dr. U. Seitzer and the late Frau Dr. S. Rüsch-Gerdes before my arrival and during my first few months in Germany is highly acknowledged and valued.

I offer my sincere gratitude with special thanks to the staff of the NRL, Borstel, including Frau G. Heinonen, Frau I. Radzio, Frau B. Schlüter, Frau P. Vock, Frau K. Ott, Frau A.-K. Witt, Frau D. Sievert, Frau S. Höllger, Frau K. Beuck, Frau B. Voß, Frau M. Kernbach, Frau S. Höllger, Frau K. Klein, Frau Dr. K. Kranzer, Herr Dr. S. Andres, and Frau Dr. D. Hillemann for their very useful and everlasting assistance during my PhD.

I would like to show my sincere gratitude and special thanks to the folks of the Molecular and Experimental Mycobacteriology research Group (FG Niemann) in Borstel, including Frau Dr. C. Gerlach, Frau Dr. S. Homolka, Herr Dr. T. Kohl, Frau Dr. S. Feuerriegel, Frau Dr. V. Dreyer, Frau Dr. J. Petersen, Herr Dr. C. Utpatel, Frau L. Sonnenkalb, Frau A. Lüdemann, Frau J. P. Ilanos,

Frau D. Beyer, Frau T. Struve-Sonnenschein, Frau T. Ubben, Frau S. Maaß, Frau V. Mohr, Frau J. Zallet, Frau C. Hahn, Frau F. Boysen, Frau H. Blankson, Herr I. Barilar, and Frau M. Weimer for their very useful and endless assistance during my PhD.

I gratefully acknowledge the extraordinary cooperation of all TB patients who participated in this PhD work and provided samples without which this work would not have been possible. Moreover, I would like to thank the physicians, technicians, and the working staff at the outpatient departments of public hospitals and local TB laboratories in Eastern Sudan for their very kind and significant assistance. The contribution of the Ministry of Health in Kassala, in El-Gadarif, and in Port Sudan, Eastern Sudan, in accomplishing this work is appreciated.

I wish to thank the Ministry of Animal Resources, Kassala, the Sudan, for its contribution in accomplishing this work. The thank also goes out to Dr. Mohamed Alhafiz Altejani, Dr. Nihal Mohamed-Noor, Dr. Malaz Isaam-Al-din Izz-Al-din, Dr. Mona Hassan, and all of the technicians at the two slaughterhouses of Kassala, and all colleagues at the Ministry of Animal Resources for their significant help and support during this work.

I am grateful to my supervisors in the Sudan, Prof. Dr. E. A. Khalil, Prof. Dr. M. A. Bakheit, and Prof. Dr. M. A. Abdalla for their kind help and for facilitating field missions and sampling. Moreover, I deeply thank the contributors (and their institutions) from Africa, Asia, and Europe for sharing their data to make the Delhi/CAS study possible and meaningfull.

Finally, I am deeply thankful to all people who helped me in one way or another to accomplish my PhD work successfully.

9 Selbstständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen in Anspruch genommen habe.

Berlin, den 18.10.2019

Yassir Adam Shuaib

mbvberlin mensch und buch verlag

49,90 Euro | ISBN: 978-3-96729-037-0