Heliyon 9 (2023) e22898

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Genetic diversity of *Mycobacterium tuberculosis* isolates from the central, eastern and southeastern Ethiopia

Mulualem Agonafir [a,*], Gurja Belay [a], Nontuthuko E. Maningi [b], Adey Feleke [a], Melese Abate Reta [b,c], Sharon L. Olifant [b], Mohammed Suaudi Hassen [d], Tewodros Girma [e], P. Bernard Fourie [b]

[a] *Department of Microbial, Cellular and Molecular Biology, College of Natural Sciences, Addis Ababa University, Ethiopia*
[b] *Department of Medical Microbiology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa*
[c] *Department of Medical Laboratory Sciences, College of Health Sciences, Woldia University, Woldia, Ethiopia*
[d] *Adama Public Health Research and Referral Laboratory Center, Adama, Ethiopia*
[e] *Harar Health Research and Regional Laboratory, Harar, Ethiopia*

ARTICLE INFO

Keywords:
Spoligotyping
Genetic diversity
Lineage
Sub-lineage
Mycobacterium tuberculosis complex
Ethiopia

ABSTRACT

Introduction: The population structure of *Mycobacterium tuberculosis* complex (MTBC) in Ethiopia is diverse but dominated by Euro-American (Lineage 4) and East-African-Indian (Lineage 3) lineages. The objective of this study was to describe the genetic diversity of MTBC isolates in Central, Eastern and Southeastern Ethiopia.

Methods: A total of 223 MTBC culture isolates obtained from patients referred to Adama and Harar TB reference laboratories were spoligotyped. Demographic and clinical characteristics were collected.

Results: Six major lineages: Euro-American (Lineage 4), East-African-Indian (Lineage 3), East Asian (Lineage 2), Indo-Oceanic (Lineage 1), *Mycobacterium africanum* (Lineage 5 and Lineage 6) and Ethiopian (Lineage 7) were identified. The majority (94.6 %) of the isolates were Euro-American and East-African-Indian, with proportions of 75.3 % and 19.3 %, respectively. Overall, 77 different spoligotype patterns were identified of which 42 were registered in the SITVIT2 database. Of these, 27 spoligotypes were unique, while 15 were clustered with 2–49 isolates. SIT149/T3_ETH (n = 49), SIT53/T1 (n = 33), SIT21/CAS1_Kili (n = 24) and SIT41/Turkey (n = 11) were the dominant spoligotypes. A rare Beijing spoligotype pattern, SIT541, has also been identified in Eastern Ethiopia. The overall clustering rate of sub-lineages with known SIT was 71.3 %. Age group (25–34) was significantly associated with clustering.

Conclusion: We found a heterogeneous population structure of MTBC dominated by T and CAS families, and the Euro-American lineage. The identification of the Beijing strain, particularly the rare SIT541 spoligotype in Eastern Ethiopia, warrants a heightened surveillance plan, as little is known about this genotype. A large-scale investigation utilizing a tool with superior discriminatory power, such as whole genome sequencing, is necessary to gain a thorough understanding of the genetic diversity of MTBC in the nation, which would help direct the overall control efforts.

* Corresponding author. P.O. Box 34738, Addis Ababa, Ethiopia.
E-mail address: macagonafir@gmail.com (M. Agonafir).

https://doi.org/10.1016/j.heliyon.2023.e22898
Received 23 September 2022; Received in revised form 13 November 2023; Accepted 22 November 2023
Available online 29 November 2023

1. Introduction

Tuberculosis (TB) and its consequences afflict millions of individuals worldwide. The World Health Organization (WHO) [1] estimated that TB killed 1.3 million people in 2020. Ethiopia is one of the 30 countries with the highest TB and TB/HIV infection rates, with an estimated 151,000 TB cases in 2020 [1].

Molecular typing of *Mycobacterium tuberculosis* complex (MTBC) has become a vital public health tool helping researchers and TB control programs better understand how specific strains emerge and spread, as well as measure the overall impact of genetic diversity on the outcome of TB infection and disease [2]. Some biological aspects of MTBC strains from distinct genetic lineages show variation, such as in vitro growth rate, pathogenicity in animal models, and the ability to acquire drug resistance [3].

Lineages known to cause TB in humans are divided into seven major TB lineages [4,5]: Indo-Oceanic (Lineage 1), East-Asian (Lineage 2), East-African-Indian (Lineage 3), Euro-American (Lineage 4), West-Africa 1 (Lineage 5), West-Africa 2 (Lineage 6) and Ethiopian (Lineage 7). Most recently, Lineage 8 [6] and Lineage 9 [7] were reported from the Central and Eastern Africa regions, respectively. Among these, the most common lineage on the planet is lineage 4 (L4). A collection of L4 was sequenced and further classified into ten sub-lineages as generalists (globally distributed) and specialists (geographically restricted) based on the ecological niche they occupy [8,9]. A report by Comas et al. (2015) [10] indicated that L4 predominates across Ethiopia, L3 is widespread but more common in the north, and L7 is mostly found in the northern Ethiopian highlands. Numerous studies have identified spoligotypes SIT 149 and SIT 53 as major clades circulating in Ethiopia [11–14].

The number of studies on the genetic diversity of MTBC in Ethiopia has been increasing. However, most of the studies reviewed elsewhere in the literature [15,16] were from Addis Ababa, the Southwestern and the Northwestern parts of Ethiopia. Studies conducted in other regions of the country are scarce or nonexistent. The lack of such information prevents the capacity of the national TB control program to implement targeted interventions, identify transmission patterns, prevent further transmission of the disease, and impede the design of novel control tools. Consequently, the findings of this study not only add to the scientific literature in the field of TB epidemiology, but also hold significant implications for improving TB control efforts in Ethiopia. Therefore, the objective of this study was to describe the population structure of MTBC isolates collected from a large geographical area that mainly included the central, eastern and southeastern parts of Ethiopia.

2. Materials and methods

2.1. Study setting and source of the isolates

This is a health facility-based cross-sectional study conducted at two TB referral diagnostic laboratories found in Harar and Adama, Ethiopia. From August 2018 to January 2019, 232 MTB culture isolates were obtained from pulmonary TB patients referred to Adama and Harar TB regional laboratories located in the premises of Adama Public Health Research and Referral Laboratory Center and Harar Health Research and Regional Laboratory, respectively. Demographic (age, sex, and address) and clinical (history of treatment) characteristics were collected. The regional TB laboratories in Adama and Harar serve as referral centers for *Mycobacteria* culture and drug susceptibility testing. The Adama TB regional laboratory is located in Adama city, which is 99 km southeast of Addis Ababa, on the main trade corridor of the country. It serves nine zones in Oromia, as well as the surrounding Amhara and Afar regions. The Harar TB regional laboratory, located in Harar city 500 km from the capital, provides referral diagnostic services to Diredawa, east Oromia, Somali, and Harari regions. Once the isolates were confirmed to be *Mycobacterium tuberculosis* complex (MTBC) using Capilia TB-Neo (Tauns Laboratories, Japan) and AFB smear staining [17], they were subcultured onto Lowenstein Jensen media and harvested within 3–4 weeks. In duplicate, two loopful colonies of the culture were transferred to 2 ml sterile cryo-vials containing 1 ml 7H9 Middlebrook liquid medium. The isolates were subsequently transported to the University of Pretoria's Medical Microbiology Department, TB laboratory for further testing.

2.2. DNA extraction

Isolates were stored at −80 °C until processed. The stored isolates were then thawed and mixed vigorously. DNA was extracted by heating the isolates at 80 °C for 60 min in a water bath. After centrifugation, the supernatant was collected for further use.

2.3. Isolate characterization

The isolates were characterized using spoligotyping following the procedure described by Kamerbeek et al. [18]. Briefly, the set of spacers in the isolates was amplified by PCR using DRa and DRb primers. The amplicons were then hybridized on a reference set of 43 spacers obtained from *Mycobacterium tuberculosis* H37Rv and *Mycobacterium bovis* BCG impregnated on a membrane (Animal and Plant Health Agency, Great Britain). The presence or absence of spacers was visualized on a film as black and white squares, which were later converted to binary codes (1/0) for analysis. Positive and negative controls were included in each batch of the test run.

2.4. Spoligotype assignment and database comparison

The spoligotype patterns entered into the MS Excel spreadsheet were converted into binary and octal formats using the SITVITWEB website (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/tools.jsp#) which were later compared with online international

databases such as SITVIT2 [19], SpolLineages [5] and RUN TB-lineage [20,21]. Sub-lineages and International Shared Types (SIT) were obtained using the updated version of SPOLDB4 and SITVITWEB (http://www.pasteur-guadeloupe.fr:8081/SITVIT2/batch.jsp). Conformal Bayesian Network (CBN) major lineages, and SNP-based lineages were determined using online tools RUN TB-lineage (https://tbinsight.cs.rpi.edu/run_tb_lineage.html) and SpolLineages (http://www.pasteur-guadeloupe.fr:8081/SpolLineages/spol.jsp), respectively. CBN does not distinguish between the West African (L6 and L5) and Ethiopian (L7) lineages, classifying them as *Mycobacterium africanum*. We used SpolLineages to avoid this, further verify the CBN classification of major lineages and explore the classification of lineages as generalists and specialists [8,9].

A dendrogram was constructed based on the UPGMA algorithm using the MIRU-VNTRplus identification database to determine the molecular clustering of the isolates. The UPGMA tree was then retouched using FigTree v1.4.4 (Fig. 1). A cluster was defined as two or more isolates with similar spoligotype patterns. A spoligotype pattern that was not previously reported in an international database was defined as orphan or new whereas a spoligotype with SIT reported once in the study was defined as unique.

2.5. Spatial distribution of lineages and sub-lineages analyses

Geographic mapping of MTBC lineage and sub-lineages was performed using QGIS v3.22.6. The shape files of study sites from where the isolates collected were obtained from UNOCHA website (https://data.humdata.org/dataset/cod-ab-eth).

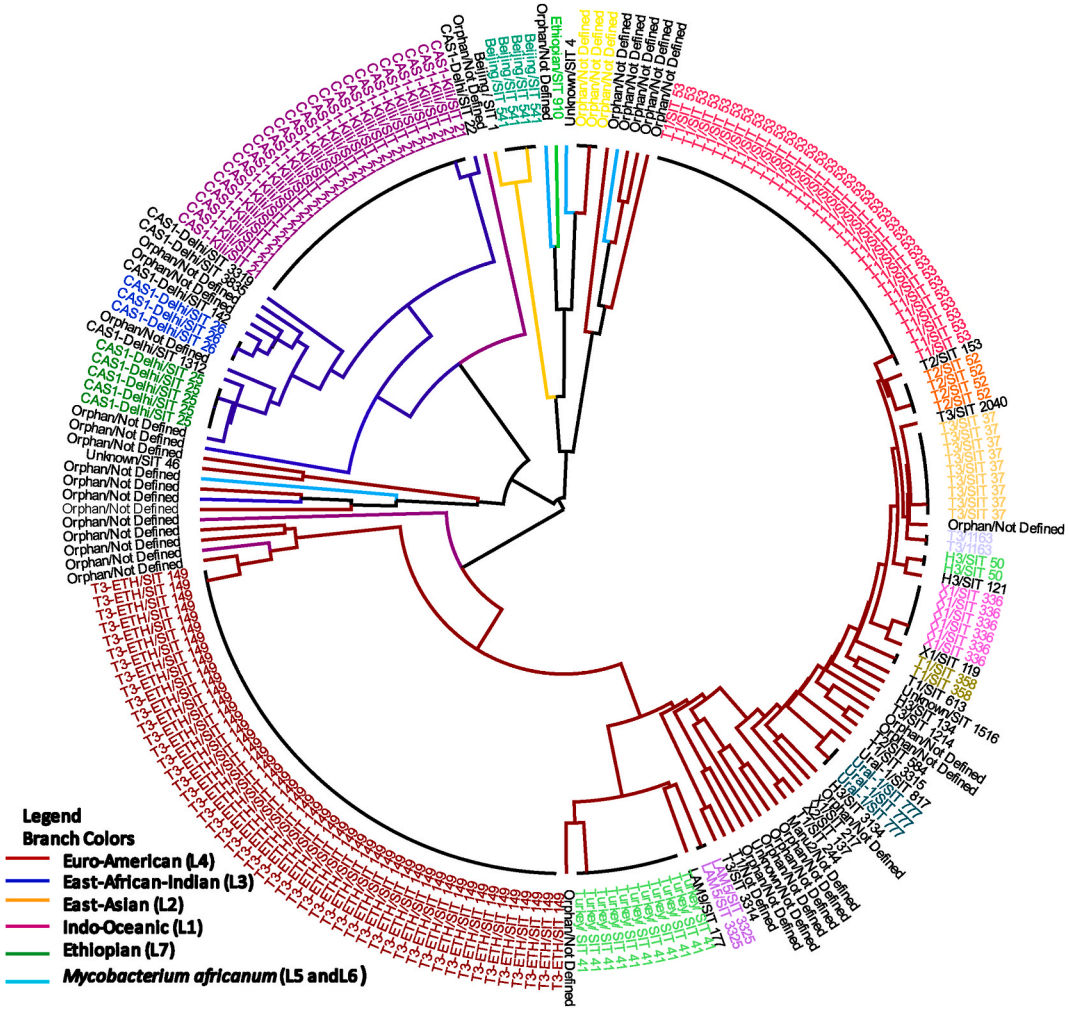

Fig. 1. Radial UPGMA tree based on spoligotyping data of 223 MTB isolates from central, eastern and southeastern Ethiopia. Annotations with similar colored fonts in the outer ring represent clustered sub-lineages/spoligotypes. The ones written in black fonts are unique/orphan isolates. The colors of branches in the inner circle indicate the major lineages. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

2.6. Statistical analysis

Data were entered into an Excel spreadsheet; cleared, and analyzed using the SPSS statistical software package, V20 (SPSS Inc., Chicago, IL, USA). Sociodemographic variables were depicted using descriptive statistics. A logistic regression model was used to assess variables associated with clustering in terms of the odds ratio and 95 % confidence interval (CI). The chi-squared test was applied to compare categorical data. A p-value less than or equal to 0.05 was considered significant.

2.7. Ethical consideration

The Ethical Review Board of Natural Sciences at Addis Ababa University granted ethical approval for this study. The Ethiopian Food, Medicine and Health Care Administration and Control Authority (now known as the Ethiopian Food and Drug Administration) and the Health Department of South Africa approved the transfer of mycobacterial isolates to South Africa for further molecular analysis. Permission to perform the research was also acquired from the Harari and Adama Public Health Research and Referral Laboratories. Only isolates routinely obtained from patients for diagnostic and therapeutic purposes were used in the study. Patients' personal information was not collected.

3. Result

3.1. Characteristics of the study population

Of the 232 isolates collected, 9 were either lost or did not have successful spoligotyping result. Hence, this study examined a total of 223 isolates from TB patients who were presumed to have DR TB, with a mean age of 30.4 years. Most of the study subjects (58.3 %) were males and in the age group-15-34 years (58.3 %). The majority (80.8 %) were from the Central (Arsi; East, North, Southeast and West Shewa) and Eastern (Diredawa, Harar, Jigjiga, East and West Hararge) parts of Ethiopia. The remaining were from the South-eastern part of Ethiopia, specifically from Bale, Borena, Guji, West Arsi and West Guji zones. There was an equal proportion of patients in relation to treatment history (Table 1).

3.2. Genetic diversity of Mycobacterium tuberculosis lineages/sub-lineages

Spoligotyping was used to genotype 223 isolates, and 77 different spoligotype patterns were identified, of which 42 spoligotype patterns comprising 186 isolates were registered in the SITVIT2 database. Of these, 27 were unique while 15 were clustered with 2–49 isolates that accounted for 85.5 % (159/186) of all isolates with known SIT. The remaining 35 orphan patterns, representing 16.6 % (37/223) of the total isolates, were not found in the SITVIT2 database (Table 2, Table 3). The overall clustering rate of sub-lineages with known SIT was 71.3 %. The proportion of clustered isolates was higher in each geographic region: Central (72.3 %), Eastern (72.1 %) and Southeastern (67.4 %). There were five dominant spoligotypes with SIT, which accounted for over half of the genotyped isolates 117 (52.5 %): SIT149/T3_ETH (n = 49), SIT53/T1 (n = 33), SIT21/CAS1_Kili (n = 24) and SIT41/Turkey (n = 11). T and CAS were the dominant families, accounting for 48.9 % and 16.6 % of the isolates, respectively. Of the 49 isolates with SIT149/T3_ETH sub-lineage, 8 (16.3 %) were children of 15 years or younger. According to the CBN analysis, 94.6 % of the total 223 isolates belonged to two major lineages: EA/L4 (75.3 %) and EAI/L3 (19.3 %). The remaining 12/223 (5.4 %) isolates were represented by EAS/L2, IO/L1, ETH/L7 and MA (L5 and L6) which were represented by five, three, one and three isolates, respectively. Of those classified by CBN as EA and EAI, 8.3 % (EA) and 4.7 % (EAI) were not known by SNP-based lineage analysis of the SpolLineages online tool. Additionally,

Table 1
Characteristics of study subjects.

Variable		Frequency (Percent)
Age, years		
	≤15	19 (8.5)
	16–24	56 (25.1)
	25–34	74(33.2)
	35–44	38(17.0)
	>45	34(15.2)
	Missing	2(.9)
Sex		
	Male	130 (58.3)
	Female	93 (41.7)
History of anti-TB drug treatment		
	New	112 (50.2)
	Previously treated	111 (49.8)
Region		
	Central	94 (42.2)
	Southeastern	43 (19.3)
	Eastern	86 (38.6)

Table 2
Description of 35 orphan strains (n = 37) and the corresponding spoligotyping defined lineages/sub-lineages recorded among MTB strains isolated from Central, Eastern and Southeastern Ethiopia.

Octal code	Binary code	SITVIT2	CBN	SNP-based lineage
525252525240521	■□■□■□■□■□■□■□■□■□■□■□■□■□■□■□□□□□■□■□■□■	ND	**EA**	EA (L4.1.1)
007777777760771	□□□□□□■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	T1	**EA**	EA (L4)
500177740003171	■□■□□□□□□□□■■■■■■■■■■■□□□□□□□□□□□□□■■□□■■■■■	ND	**EAI**	EAI (L3)
501252500002121	■□■□□□□□■□■□■□■□■□■□■□□□□□□□□□□□□□■□□■□■□■	ND	**EAI**	EAI (L3)
777000377777771	■■■■■■■■■□□□□□□□□□□■■■■■■■■■■■■■■■■■■■■■■■■	ND	**IO**	UNK
700006037177661	■■■□□□□□□□□□□□□■■□□□□□□■■■■■□□■■■■■■■■■■■□■■□■	ND	**MA**	WA (L5 and L6)
703757740003171	■■■□□□□■■■■■■■□■■■■■■■■■□□□□□□□□□□□□□■■□□■■■■■	ND	**EAI**	EAI (L3)
713247664001661	■■■□□■□■■□■□■□□■■■■■□■■□■□□□□□□□□□□■■■□■■□■	ND	**IO**	UNK
777777777000171	■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□□□□□□■■■■■	UKN	**EA**	UNK
771046024100261	■■■■■■□□■□□□■□□■■□□□□□■□■□□■□□□□□□□■□■■■□■	ND	**EA**	UNK
773777776000771	■■■■■■□■■■■■■■■■■■■■■■■■■■□□□□□□□□□■■■■■■■	ND	**EA**	UNK
774242525042000	■■■■■■■□□□■□■□□□■□■□■□■□■□■□□□■□□□■□□□□□□□□	ND	**MA**	UNK
703357740003071	■■■□□□□■■□■■■□■■■■■■■■■□□□□□□□□□□□□□■■□□□■■■■	ND	**EAI**	EAI (L3)
703777710003771	■■■□□□□■■■■■■■■■■■■■■■□□■□□□□□□□□□□□■■■■■■■■■	ND	**EAI**	UNK
777737777360771	■■■■■■■■■■■■□■■■■■■■■■■■■■■□■■■■□□□□□■■■■■■■	ND	**EA**	EA (L4)
777775477760731	■■■■■■■■■■■■■■■■□■■□□■■■■■■■■□□□□■■■□■■□□■■	ND	**EA**	EA (L4)
777777404060771	■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□□□■■□□□□■■■■■■■	ND	**EA**	EA (L4.3)
477777376413771	■□□■■■■■■■■■■■■■■■□■■■■■■■□■□□□□□■■□■■■■■■■■	ND	**IO**	IO (L1)
777737377750771	■■■■■■■■■■■■□■■■■■□■■■■■■■■■■■■□■□□□□■■■■■■■	ND	**EA**	UNK
510047236343661	■□■□□■□□□□□□■□□■■■□■□□■■■■□□■■■□□■■■■■□■■□■	ND	**MA**	WA (L5 and L6)
525252500000000	■□■□■□■□■□■□■□■□■□■□■□□□□□□□□□□□□□□□□□□□□□□	ND	**EA**	UNK
551246236340261	■□■■□■□□■□■□■□□■■□□■□□■■■■□□■■■□□□□□□□■□■■□■	ND	**EA**	EA (L4)
511047236340261	■□■□□■□□■□□□■□□■■■□■□□■■■■□□■■■□□□□□□□■□■■□■	ND	**EA**	EA (L4)
777347777763771	■■■■■■■■■□■■■□□■■■■■■■■■■■■■■■■■■□□■■■■■■■■■	Manu2	**EA**	UNK
511046637561671	■□■□□■□□■□□□■□□■■□■■□□■■■■■■□■■■□□□■■■□■■■■■	ND	**EA**	UNK
603777700003771	■■□□□□□■■■■■■■■■■■■■■■□□□□□□□□□□□□□■■■■■■■■■	ND	**EAI**	UNK
311000377760771	□■■□□■□□■□□□□□□□□□□■■■■■■■■■■■■■■□□□□■■■■■■■	ND	**EA**	EA (L4)
777000377730771	■■■■■■■■■□□□□□□□□□□■■■■■■■■■■■□■■□□□■■■■■■■	ND	**EA**	UNK
713357776363771	■■■□□■□■■□■■■□■■■■■■■■■■■■□□■■■■■□□■■■■■■■■■	ND	**EA**	UNK
777347636361771	■■■■■■■■■□■■■□□■■■■■□□■■■■□□■■■■■□□□■■■■■■■■	ND	**EA**	UNK
777775777760721	■■■■■■■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■□■□■	ND	**EA**	EA (L4)
774177770000021	■■■■■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□□□■□■	ND	**EA**	UNK
520000377760771	■□■□■□□□□□□□□□□□□□□■■■■■■■■■■■■■■□□□□■■■■■■■	ND	**EA**	EA (L4)
700000017760771*	■■■□□□□□□□□□□□□□□□□□□□□■■■■■■■■■■■□□□□■■■■■■■	ND	**EA**	EA (L4)
777357677761771	■■■■■■■■■□■■■□■■■■■■■□■■■■■■■■■■■■■□□□■■■■■■■	ND	**EA**	UNK

EA = Euro-American, EA = East-Asian, EAI = East-African-Indian, IO=Indo-Oceanic, UKN=Unknown, WA=West Africa, MA = *Mycobacterium Africanum,* ND=Not defined*, *n=3*.

we found two generalist sub-lineages [8,9], L4.1.2/Haarlem (n = 5) and L4.3/LAM (n = 5), and most (9/10) were isolated from the central and eastern regions (Table 2).

3.3. Geographical distribution of the lineages/sub-lineages

EA (L4) contributed a significantly high proportion to the lineage distribution across the study sites: 83.7 % (Southeastern), 81.9 % (Central) and 64.0 % (Eastern) of the isolates collected in the respective parts of Ethiopia (Supplementary material 1). EAI (L3) was reported mainly in the eastern regions, followed by the central regions, with proportions of 62.8 % and 27.9 % of the total isolates, respectively. With the exception of Guji/West Guji, where no EAI was recorded, EA and EAI were identified in all zones from where the isolates were acquired. MTBC isolates obtained from Arsi (n = 2) and West Guji (n = 1) zones were identified as *Mycobacterium africanum* (Fig. 2, Supplementary material 2). The SIT910 spoligotype (ETH/L7), which is confined to Ethiopia, was isolated from West Arsi. There were five isolates with Beijing sub-lineage (EAS/L2), four from the eastern (three from Diredawa and one from Harar) and one from the central part (Southwest Shewa) of Ethiopia. Of note, 10/11 (90.9 %) of EA (L4) lineage, SIT41 spoligotype/Turkey sub-lineage were reported from southeastern Ethiopia (Borena and Guji). On the contrary, the SIT149 spoligotype/T3-ETH sub-lineage and SIT53 spoligotype/T1 sub-lineages were reported from most of the sites where the isolates were collected (Fig. 3, Supplementary materials 3 and 4).

3.4. Comparison of factors associated with clustering

In a bivariate logistic regression analysis, factors such as age, treatment history, and lineage were statistically associated with isolate clustering at p-values 0.2. However, there was no statistically significant association (p-value>0.05) of sex and region with clustering. Age group was the only variable statistically associated with clustering of MTBC isolates (p-value<0.05) in multivariable analysis (Table 4).

Table 3
Description of 42 shared types (SITs; n = 186 isolates) registered in the SITVIT2 or SpolDB4 database and the corresponding spoligotyping defined lineages/sub-lineages isolated from the Central, Eastern and Southeastern Ethiopia.

Octal code	Binary code	SITVIT2	CBN	SNP-based lineage	SIT	Number of isolates
777000377760771		T3-ETH	EA	EA (L4)	149	49
777777777760771		T1	EA	EA (L4)	53	33
703377400001771		CAS1-Kili	EAI	EAI (L3)	21	24
777777404760771		Turkey	EA	EA (L4.2.2.1)	41	11
777737777760771		T3	EA	EA (L4)	37	9
777776777760731		X1	EA	EA (L4.1.1)	336	6
703777740003171		CAS1-Delhi	EAI	EAI (L3)	25	5
777777777760731		T2	EA	EA (L4)	52	4
000000000003711		Beijing	EAS	EAS (L2)	541	4
703777740003771		CAS1-Delhi	EAI	EAI (L3)	26	3
777777777420771		Ural-1	EA	EA (L4.2.1)	777	3
677737777760771		T3	EA	EA (L4)	1163	2
776737407760771		LAM5	EA	EA (L4.3)	3325	2
717777777760771		T1	EA	EA (L4)	358	2
777777777720771		H3	EA	EA (L4.1.2)	50	2
000000000003771		Beijing	EAS	EAS (L2)	1	1
777776777760771		X1	EA	EA (L4.1.1)	119	1
777777775720771		H3	EA	EA (L4.1.2)	121	1
777617777760771		T3	EA	EA (L4)	1214	1
703777740003131		CAS1-Delhi	EAI	EAI (L3)	1312	1
777777777720631		H3	EA	EA (L4.1.2)	134	1
777776777760601		X2	EA	EA (L4.1.1)	137	1
703777700003771		CAS1-Delhi	EAI	EAI (L3)	142	1
777377777761771		UKN	EA	UNK	1516	1
757777777760731		T2	EA	EA (L4)	153	1
377777607760771		LAM9	EA	EA (L4.3)	177	1
777737777760371		T3	EA	EA (L4)	2040	1
777736777760771		X1	EA	EA (L4.1.1)	217	1
703777400001771		CAS1_Dehi	EAI	EAI (L3)	22	1
777777777760601		T1	EA	EA (L4)	244	1
777737377720771		H3	EA	EA (L4.1.2)	3134	1
177000377760771		T3-ETH	EA	EA (L4)	3141	1
006737777760771		T3	EA	EA (L4)	3314	1
376777737760771		T1	EA	EA (L4)	3315	1
703417740003771		CAS1-Delhi	EAI	EAI (L3)	3319	1
703177740003771		CAS1-Delhi	EAI	EAI (L3)	3835	1
000000007760771		UKN	EA	EA (L4.3)	4	1
777777770000000		UKN	EA	UKN	46	1
777775777760731		T2	EA	EA (L4)	584	1
777177777760771		T1	EA	EA (L4)	613	1
777777777420731		Ural-1	EA	EA (L4.2.1)	817	1
700000007177771		ETH	MA	ETH (L7)	910	1

EA = Euro-American, EA = East-Asian, EAI = East-African-Indian, IO=Indo-Oceanic, ETH = Ethiopian, UKN=Unknown, WA=West Africa, MA = *Mycobacterium Africanum,* ND=Not-defined.

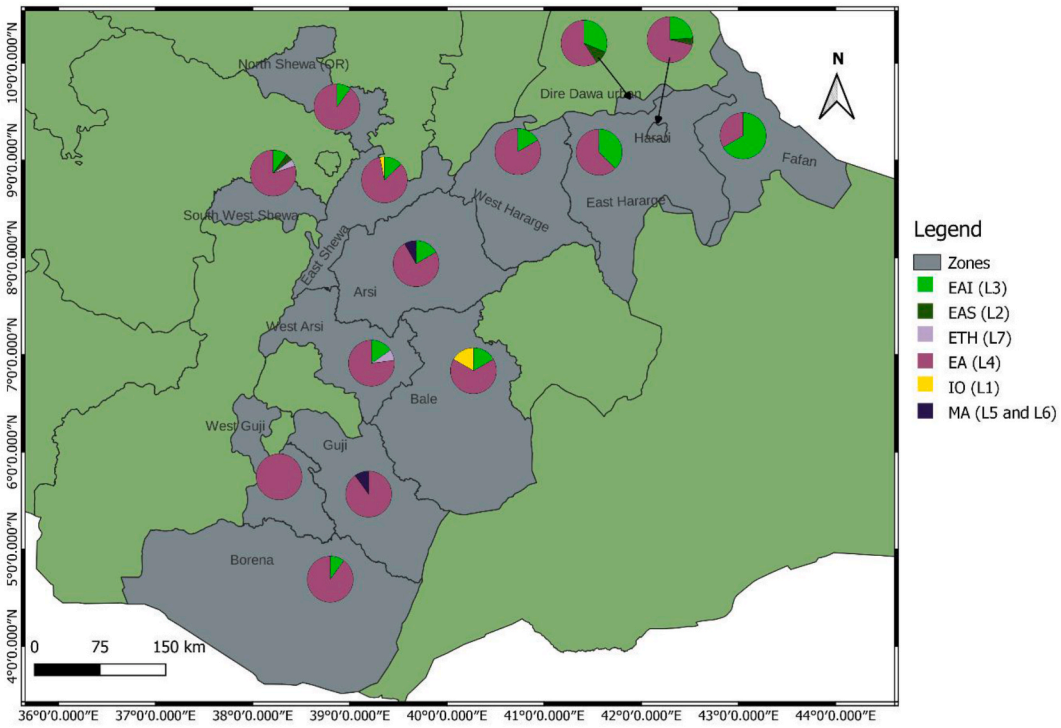

Fig. 2. Distribution of major lineages in central, eastern and southeastern Ethiopia. Pie charts show the proportions of the six lineages among MTBC isolates in each zone from where isolates were obtained. The size of the circle does not correspond to the number of isolates analyzed. The actual numbers are shown in Supplementary material 2: color codes are as in Fig. 2. A total of 223 MTBC isolates were included. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4. Discussion

A review on molecular epidemiology of MTBC in Ethiopia [16] highlighted the scarcity of data from certain regions such as Harari and pastoralist communities of Oromia which include the Borena and Guji zones. The low number of isolates identified from these regions and other less investigated areas would restrict the scientific community's ability to observe the full picture of the transmission dynamics and population structure of MTBC in Ethiopia. In line with this, the current study examined the genomic diversity of MTBC isolates obtained from patients in a large geographical area spanning from the central part of Ethiopia to the far Southeast (Moyale) bordering the northern part of Kenya and the East (Diredawa, Harar, and Jigjiga) doorway to Djibouti and Somalia.

Supporting previous studies summarized in two recent reviews [15,16], a heterogeneous MTBC population structure dominated by the T and CAS families with the corresponding EA (L4) and EAI (L3) lineages was shown in the current study. Three-quarters (75.3 %) of the isolates in this study were members of EA (L4). A similar or higher proportion of EA (L4) has been reported from different parts of Ethiopia, including the current study sites [14,22–27]. However, a lower proportion (40.1 %) of EA (L4) has been reported in northwest Ethiopia [28]. We also identified the globally predominant sub-lineages L4.1.2/Haarlem (n = 5) and L.4.3/LAM (n = 5). L4.1.2/Haarlem was previously shown to be connected to a high rate of transmission clusters [29] in Ethiopia. It is also the third major sub-lineage in Ethiopia [15,16], albeit isolated in modest proportions in the current study.

Various studies conducted in different parts of Ethiopia reported an overall low prevalence of IO (L1) and ETH (L7) [15,16]. Accordingly, the current study identified three IO (L1) isolates from Bale, East, and Southwest Shewa and one ETH (L7) isolate from West Arsi. ETH (L7), first reported from Woldia [30], is predominant in the northern highlands of Ethiopia [10,31]. It has also been identified in different parts of Ethiopia [11,14,32–35]. Notably, this lineage is known to progress toward disease at a slower rate than other lineages [36] and is mainly limited to Ethiopia and Ethiopian immigrants [37,38]. Consequently, extensive investigation is required to comprehend why it is unique to Ethiopia and Ethiopians.

Three isolates of the West African-bound MA (L5 and L6) lineages were also identified from Arsi and Guji. When verified using the SNP-based lineage identification tool [5], one isolate was identified as unknown, whereas the other two were confirmed as MA (L5 and L6). There is an evidence that the MA (L5 and L6) are less virulent and are being replaced by EA (L4) in West Africa [10]. Conversely, in Ethiopia where EA (L4) is predominant, MA (L5 and L6) lineages have been isolated sporadically [22,23,32,39,40]. Therefore, it is crucial to carefully examine the impact of these lineages on the epidemiology of TB in Ethiopia, as this may have implications for TB control strategies and treatment approaches.

Consistent with other investigations carried out in Ethiopia [11–13], the T3-ETH sub-lineage/SIT149 spoligotype, reported to be more likely in a cluster [36,41,42], was the most frequent in our study, followed by the T1 sub-lineage/SIT53 spoligotype. A study

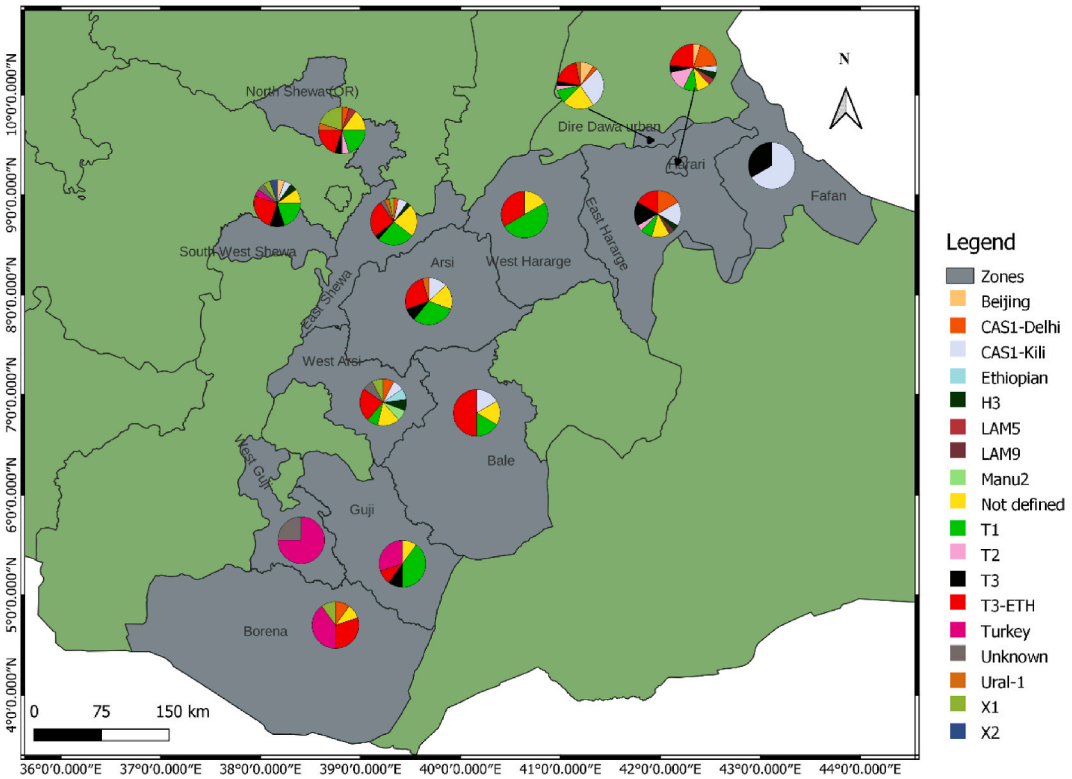

Fig. 3. Distribution of MTBC sub-lineages in central, eastern and southeastern Ethiopia. Pie charts show the proportions of the sub-lineages among MTBC isolates in each zone from where isolates were obtained. The size of the circle does not correspond to the number of isolates analyzed. The actual numbers are shown in Supplementary material 4: color codes are as shown in Fig. 3. A total of 223 MTBC isolates were included. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Comparison of patient characteristics with clustering of MTBC isolates from the central, eastern and southeastern Ethiopia.

Variable	Category	Clustered (N/%)	Unique (N/%)	COR (95 % CI)	P-value	AOR (95 % CI)	P-value
Age	≤15	17 (89.5)	4 (10.5)	2.975 (0.823–10.760)	0.096	3.055 (0.831–11.230)	0.093
	16–24	36 (64.3)	20 (35.7)	1.260 (0.525–3.022)	0.605	1.393 (0.572–3.390)	0.465
	25–34	61 (82.4)	13 (17.6)	3.285 (1.324–8.146)	0.01	3.728 (1.467–9.470)	0.006
	35–44	25 (65.8)	13 (34.2)	1.346 (0.517–3.505)	0.543	1.363 (0.517–3.593)	0.531
	≥45	20 (58.8)	14 (41.2)	1		1	
Sex	Male	95 (73.1)	35 (26.9)	1.230 (0.685–2.209)	0.488		
	Female	64 (68.8)	29 (31.2)	1			
Region	Central	68 (72.3)	26 (27.9)	1.012 (0.527–1.945)	0.97		
	Southeast	29 (67.4)	14 (32.6)	0.802 (0.363–1.772)	0.585		
	East	62 (72.1)	24 (27.9)	1			
Treatment	New	74 (66.7)	37 (33.3)	0.635 (0.354–1.141)	0.129	0.644 (0.350–1.185)	0.158
	Previously treated	85 (75.9)	27 (24.1)	1		1	
Lineage	EA (L4)	123 (73.2)	45 (26.8)	1.443 (0.751–2.770)	0.201	1.842 (0.919–3.689)	0.061
	Others	36 (65.5)	19 (34.5)	1			

conducted in Oromia region [43], which included the zones from which two-thirds of our isolates were obtained, found a high proportion of clustered ST149 spoligotype followed by orphan spoligotypes rather than the commonly reported T or CAS sub-lineages. This could be attributed to several factors including differences in sampling techniques, the broad geographic area and high population of the region and the relatively small sample size used in the current study.

The high proportion of clustered spoligotypes, particularly SIT149, identified in different parts of the country magnifies its key role in TB transmission dynamics and disease burden in Ethiopia. Furthermore, the finding that 16.3 % of the 49 isolates with spoligotype SIT149 in the current study were from children aged ≤15 years may indicate an ongoing TB transmission from adults to children. A similar proportion (16.6 %) of the total isolates were not registered in the online SITVIT2/SPOLDB4 database suggesting that there is a need further clarify the population structure of MTBC causing TB in Ethiopia.

A relatively large cluster (n = 11) of the SIT41 spoligotype, which is phylogeographically specific to Turkey [44], was reported in

this study with the majority (10/11) being from patients in southeastern (Borena and Guji) Ethiopia. This sub-lineage, although in small numbers (n ≤ 5), was previously reported [23,28,30,32,43,45–48] in different parts of Ethiopia. Similarly, it was reported to be less common in the African continent [49]. Clusters of larger size (n = 6) from Guji and Liben area pastoralists [50] and most recently from Arsi zone (n = 7) [29] were reported. The findings of our study strengthen the increasing importance of this clade in the pastoralist communities of the southeastern Ethiopia and call for further investigation to assess its role in TB disease burden of these regions and to implement appropriate interventions.

The widely studied Beijing sub-lineage of the (EAS) L2 lineage characterized by its worldwide geographical distribution and virulent properties [38] has previously been documented in some parts of Ethiopia [25,48,51,52]. In this study, in addition to the classical Beijing strain-SIT1, we identified a rare spoligotype cluster, SIT541, characterized by the absence of two more spacers (40 and 41) compared to the predominant SIT1. At the time of our analysis, only 16 isolates were reported worldwide, representing 0.1 % of the total registered Beijing lineage, as per the SITVIT2 database. Interestingly, in the present study, all SIT541 isolates were detected in eastern Ethiopia, specifically in Diredawa (three isolates) and Harar (one isolate). The exclusive detection of this spoligotype in eastern Ethiopia and its limited global distribution may suggest a localized emergence and transmission of this strain in the region. Furthermore, the findings of the current study, together with a previous study reporting a similar frequency of SIT1 from Diredawa [24], may indicate a unique epidemiological pattern and stress the evolving nature of the Beijing strain in the region, highlighting the need for enhanced surveillance and monitoring. A large scale longitudinal study is required to better understand the pressure that this lineage might put on the community in the region and on the overall TB control effort of the region.

A high rate (71.3 %) of clustered sub-lineages with known SIT was recorded in the current study. Besides, the proportion of clustered isolates was higher in each geographic region: Central (72.3 %), Eastern (72.1 %) and Southeastern (67.4 %). A comparable rate of clustering has been reported from different regions across Ethiopia [11,13,23,53,54]. The high rate of clustering rate, although not with epidemiological link information, could indicate an ongoing transmission of TB in the respective regions. Variables, such as sex, geographic region, treatment history, and lineage were not significantly associated with an isolate being in a cluster. However, patients aged 25–34 years (AOR = 3.728, 95 % CI (1.467–9.470)) were more likely to be part of a cluster compared to those aged >45 years. This may indicate that individuals in this age group had a higher risk of recent TB and therefore, could be an entry point for targeted interventions.

Convenient selection of isolates from specific diagnostic centers may pose selection bias and the relatively small number of isolates could affect the representativeness of the study result. Besides, the low discriminatory power of spoligotyping limits our conclusion regarding MTBC transmission dynamics. In contrast, whole-genome sequencing (WGS) provides a more comprehensive understanding of TB transmission patterns and strain diversity owing to its higher resolution, single nucleotide polymorphisms identification capabilities, better discrimination and provision of additional information compared to traditional typing methods such as spoligotyping [55]. Additionally, it prevents unnecessary false cluster investigations and allows more targeted interventions [56]. Additionally, it enables more focused actions to avoid pointless cluster investigations.

5. Conclusion

Though dominated by T (L4) and CAS (L3) families, our findings indicate that the genetic diversity of MTBC in the study area is diverse. The predominance of EA (L4) in the study area, comprising three-fourths of the isolates analyzed, shows its impact on the TB landscape of the country. Moreover, the identification of the Beijing strain, particularly the rare SIT541 spoligotype cluster in Eastern Ethiopia, calls for enhanced surveillance. Understanding the genetic diversity and transmission dynamics of EA (L4) could reveal the factors contributing to its predominance in Ethiopia. Hence, it is imperative to undertake a large scale study using advanced tools such as WGS to gather a comprehensive information and guide the national TB control program in implementing targeted interventions to combat TB in Ethiopia.

Data availability statement

Data related to the current study are all included in the results section.

CRediT authorship contribution statement

Mulualem Agonafir: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Gurja Belay:** Methodology, Supervision, Validation, Writing - review & editing. **Nontuthuko E. Maningi:** Investigation, Methodology, Validation, Writing - review & editing. **Adey Feleke:** Methodology, Validation, Writing - review & editing. **Melese Abate Reta:** Investigation, Validation, Writing - review & editing. **Sharon L. Olifant:** Formal analysis, Investigation, Validation, Writing - review & editing. **Mohammed Suaudi Hassen:** Investigation, Resources, Validation, Writing - review & editing. **Tewodros Girma:** Formal analysis, Investigation, Resources, Validation, Writing - review & editing. **P. Bernard Fourie:** Investigation, Methodology, Resources, Supervision, Validation, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to

influence the work reported in this paper.

Acknowledgements

We would like to thank University of Pretoria, Medical Microbiology Department for supporting the laboratory work. We are grateful to Dr. Giorgio Roscigno for facilitating the process of collaborative work and Dr. Abel Gizaw for technical support during the manuscript preparation. We acknowledge the comprehensive assistance provided by the Department of Microbial, Cellular, and Molecular Biology, Addis Ababa University. We also express our gratitude to the regional TB reference laboratories at Adama and Harar.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e22898.

References

[1] World Health Organization, Global Tuberculosis Report 2021, World Health Organization, Geneva, 2021. https://apps.who.int/iris/handle/10665/346387. (Accessed 18 May 2022).
[2] I. Comas, S. Gagneux, The past and future of tuberculosis research, PLoS Pathog. 5 (2009), e1000600, https://doi.org/10.1371/journal.ppat.1000600.
[3] I. Mokrousov, O. Pasechnik, A. Vyazovaya, I. Yarusova, A. Gerasimova, A. Blokh, V. Zhuravlev, Impact of pathobiological diversity of Mycobacterium tuberculosis on clinical features and lethal outcome of tuberculosis, BMC Microbiol. 22 (2022) 50, https://doi.org/10.1186/s12866-022-02461-w.
[4] D. Brites, C. Loiseau, F. Menardo, S. Borrell, M.B. Boniotti, R. Warren, A. Dippenaar, S.D.C. Parsons, C. Beisel, M.A. Behr, J.A. Fyfe, M. Coscolla, S. Gagneux, A new phylogenetic framework for the animal-adapted Mycobacterium tuberculosis complex, Front. Microbiol. 9 (2018) 2820, https://doi.org/10.3389/fmicb.2018.02820.
[5] D. Couvin, W. Segretier, E. Stattner, N. Rastogi, Novel Methods Included in SpolLineages Tool for Fast and Precise Prediction of *Mycobacterium tuberculosis* Complex Spoligotype Families, Database. 2020, 2020, https://doi.org/10.1093/database/baaa108 baaa108.
[6] J.C.S. Ngabonziza, C. Loiseau, M. Marceau, A. Jouet, F. Menardo, O. Tzfadia, R. Antoine, E.B. Niyigena, W. Mulders, K. Fissette, M. Diels, C. Gaudin, S. Duthoy, W. Ssengooba, E. André, M.K. Kaswa, Y.M. Habimana, D. Brites, D. Affolabi, J.B. Mazarati, B.C. de Jong, L. Rigouts, S. Gagneux, C.J. Meehan, P. Supply, A sister lineage of the Mycobacterium tuberculosis complex discovered in the African Great Lakes region, Nat. Commun. 11 (2020) 2917, https://doi.org/10.1038/s41467-020-16626-6.
[7] M. Coscolla, S. Gagneux, F. Menardo, C. Loiseau, P. Ruiz-Rodriguez, S. Borrell, I.D. Otchere, A. Asante-Poku, P. Asare, L. Sánchez-Busó, F. Gehre, C.N. Sanoussi, M. Antonio, D. Affolabi, J. Fyfe, P. Beckert, S. Niemann, A.S. Alabi, M.P. Grobusch, R. Kobbe, J. Parkhill, C. Beisel, L. Fenner, E.C. Böttger, C.J. Meehan, S. R. Harris, B.C. de Jong, D. Yeboah-Manu, D. Brites, Phylogenomics of Mycobacterium africanum reveals a new lineage and a complex evolutionary history, Microb. Genom. 7 (2021), 000477, https://doi.org/10.1099/mgen.0.000477.
[8] F. Coll, R. McNerney, J.A. Guerra-Assunção, J.R. Glynn, J. Perdigão, M. Viveiros, I. Portugal, A. Pain, N. Martin, T.G. Clark, A robust SNP barcode for typing Mycobacterium tuberculosis complex strains, Nat. Commun. 5 (2014) 4812, https://doi.org/10.1038/ncomms5812.
[9] D. Stucki, D. Brites, L. Jeljeli, M. Coscolla, Q. Liu, A. Trauner, L. Fenner, L. Rutaihwa, S. Borrell, T. Luo, Q. Gao, M. Kato-Maeda, M. Ballif, M. Egger, R. Macedo, H. Mardassi, M. Moreno, G. Tudo Vilanova, J. Fyfe, M. Globan, J. Thomas, F. Jamieson, J.L. Guthrie, A. Asante-Poku, D. Yeboah-Manu, E. Wampande, W. Ssengooba, M. Joloba, W. Henry Boom, I. Basu, J. Bower, M. Saraiva, S.E.G. Vaconcellos, P. Suffys, A. Koch, R. Wilkinson, L. Gail-Bekker, B. Malla, S.D. Ley, H.-P. Beck, B.C. de Jong, K. Toit, E. Sanchez-Padilla, M. Bonnet, A. Gil-Brusola, M. Frank, V.N. Penlap Beng, K. Eisenach, I. Alani, P. Wangui Ndung'u, G. Revathi, F. Gehre, S. Akter, F. Ntoumi, L. Stewart-Isherwood, N.E. Ntinginya, A. Rachow, M. Hoelscher, D.M. Cirillo, G. Skenders, S. Hoffner, D. Bakonyte, P. Stakenas, R. Diel, V. Crudu, O. Moldovan, S. Al-Hajoj, L. Otero, F. Barletta, E. Jane Carter, L. Diero, P. Supply, I. Comas, S. Niemann, S. Gagneux, Mycobacterium tuberculosis Lineage 4 comprises globally distributed and geographically restricted sublineages, Nat. Genet. 48 (2016) 1535–1543, https://doi.org/10.1038/ng.3704.
[10] I. Comas, E. Hailu, T. Kiros, S. Bekele, W. Mekonnen, B. Gumi, R. Tschopp, G. Ameni, R.G. Hewinson, B.D. Robertson, G.A. Goig, D. Stucki, S. Gagneux, A. Aseffa, D. Young, S. Berg, Population genomics of Mycobacterium tuberculosis in Ethiopia contradicts the virgin soil hypothesis for human tuberculosis in sub-saharan Africa, Curr. Biol. 25 (2015) 3260–3266, https://doi.org/10.1016/j.cub.2015.10.061.
[11] G. Ameni, K. Tadesse, E. Hailu, Y. Deresse, G. Medhin, A. Aseffa, G. Hewinson, M. Vordermeier, S. Berg, Transmission of Mycobacterium tuberculosis between farmers and cattle in Central Ethiopia, PLoS One 8 (2013), e76891, https://doi.org/10.1371/journal.pone.0076891.
[12] L. Garedew, A. Mihret, G. Ameni, Molecular typing of mycobacteria isolated from extrapulmonary tuberculosis patients at Debre Birhan Referral Hospital, central Ethiopia, Scand. J. Infect. Dis. 45 (2013) 512–518, https://doi.org/10.3109/00365548.2013.773068.
[13] O. Zewdie, A. Mihret, G. Ameni, A. Worku, T. Gemechu, T. Abebe, Molecular typing of mycobacteria isolated from tuberculous lymphadenitis cases in Addis Ababa, Ethiopia, Int. J. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis. 20 (2016) 1529–1534, https://doi.org/10.5588/ijtld.15.1023.
[14] Y. Merid, E. Hailu, G. Habtamu, M. Tilahun, M. Abebe, M. Hailu, T. Hailu, D.G. Datiko, Y. Woldeamanuel, A. Aseffa, Molecular Epidemiology of Mycobacterium tuberculosis strains isolated from pulmonary tuberculosis patients in south Ethiopia, J. Infect. Dev. Ctries. 15 (2021) 1299–1307, https://doi.org/10.3855/jidc.14742.
[15] D. Mekonnen, A. Derbie, A. Chanie, A. Shumet, F. Biadglegne, Y. Kassahun, K. Bobosha, A. Mihret, L. Wassie, A. Munshea, E. Nibret, S.A. Yimer, T. Tønjum, A. Aseffa, Molecular epidemiology of M. tuberculosis in Ethiopia: a systematic review and meta-analysis, Tuberculosis 118 (2019), 101858, https://doi.org/10.1016/j.tube.2019.101858.
[16] B. Tulu, G. Ameni, Spoligotyping based genetic diversity of Mycobacterium tuberculosis in Ethiopia: a systematic review, BMC Infect. Dis. 18 (2018) 140, https://doi.org/10.1186/s12879-018-3046-4.
[17] G.-H. Shen, C.-H. Chen, C.-H. Hung, K.-M. Wu, C.-F. Lin, Y.-W. Sun, J.-H. Chen, Combining the Capilia TB assay with smear morphology for the identification of Mycobacterium tuberculosis complex, Int. J. Tubercul. Lung Dis. 13 (2009) 371–376.
[18] J. Kamerbeek, L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, J. van Embden, Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology, J. Clin. Microbiol. 35 (1997) 907–914, https://doi.org/10.1128/jcm.35.4.907-914.1997.
[19] D. Couvin, A. David, T. Zozio, N. Rastogi, Macro-geographical specificities of the prevailing tuberculosis epidemic as seen through SITVIT2, an updated version of the Mycobacterium tuberculosis genotyping database, Infect. Genet. Evol. 72 (2019) 31–43, https://doi.org/10.1016/j.meegid.2018.12.030.
[20] M. Aminian, A. Shabbeer, K.P. Bennett, A conformal Bayesian network for classification of Mycobacterium tuberculosis complex lineages, BMC Bioinf. 11 (2010) S4, https://doi.org/10.1186/1471-2105-11-S3-S4.
[21] M. Aminian, D. Couvin, A. Shabbeer, K. Hadley, S. Vandenberg, N. Rastogi, K.P. Bennett, Predicting *Mycobacterium tuberculosis* complex clades using knowledge-based bayesian networks, BioMed Res. Int. 2014 (2014) 1–11, https://doi.org/10.1155/2014/398484.

[22] G. Abebe, K. Abdissa, K. Abdella, M. Tadesse, A. Worku, G. Ameni, Spoligotype-based population structure of Mycobacterium tuberculosis in the Jimma Zone, southwest Ethiopia, MicrobiologyOpen 8 (2019), e00744, https://doi.org/10.1002/mbo3.744.

[23] Z. Bedewi, A. Worku, Y. Mekonnen, G. Yimer, G. Medhin, G. Mamo, R. Pieper, G. Ameni, Molecular typing of Mycobacterium tuberculosis complex isolated from pulmonary tuberculosis patients in central Ethiopia, BMC Infect. Dis. 17 (2017) 184, https://doi.org/10.1186/s12879-017-2267-2.

[24] A. Mekonnen, M. Merker, J.M. Collins, D. Addise, A. Aseffa, B. Petros, G. Ameni, S. Niemann, Molecular epidemiology and drug resistance patterns of Mycobacterium tuberculosis complex isolates from university students and the local community in Eastern Ethiopia, PLoS One 13 (2018), e0198054, https://doi.org/10.1371/journal.pone.0198054.

[25] T.A. Lobie, Y. Woldeamanuel, D. Asrat, D. Beyene, M. Bjørås, A. Aseffa, Genetic diversity and drug resistance pattern of Mycobacterium tuberculosis strains isolated from pulmonary tuberculosis patients in the Benishangul Gumuz region and its surroundings, Northwest Ethiopia, PLoS One 15 (2020), e0231320, https://doi.org/10.1371/journal.pone.0231320.

[26] B. Haile, K. Tafess, A. Zewude, B. Yenew, G. Siu, G. Ameni, Spoligotyping and drug sensitivity of Mycobacterium tuberculosis isolated from pulmonary tuberculosis patients in the Arsi Zone of southeastern Ethiopia, New Microbes New Infect 33 (2020), 100620, https://doi.org/10.1016/j.nmni.2019.100620.

[27] L.N. Welekidan, S.A. Yimer, E. Skjerve, T.A. Dejene, H. Homberset, T. Tønjum, O. Brynildsrud, Whole genome sequencing of drug resistant and drug susceptible Mycobacterium tuberculosis isolates from tigray region, Ethiopia, Front. Microbiol. 12 (2021), 743198, https://doi.org/10.3389/fmicb.2021.743198.

[28] M. Ejo, G. Torrea, C. Uwizeye, M. Kassa, Y. Girma, T. Bekele, Y. Ademe, E. Diro, F. Gehre, L. Rigouts, B.C. de Jong, Genetic diversity of the Mycobacterium tuberculosis complex strains from newly diagnosed tuberculosis patients in Northwest Ethiopia reveals a predominance of East-African-Indian and Euro-American lineages, Int. J. Infect. Dis. 103 (2021) 72–80, https://doi.org/10.1016/j.ijid.2020.11.129.

[29] K. Tafess, T.K. Beyen, S. Girma, A. Girma, G. Siu, Spatial clustering and genetic diversity of Mycobacterium tuberculosis isolate among pulmonary tuberculosis suspected patients, Arsi Zone, Ethiopia, BMC Pulm. Med. 21 (2021) 206, https://doi.org/10.1186/s12890-021-01567-7.

[30] R. Firdessa, S. Berg, E. Hailu, E. Schelling, B. Gumi, G. Erenso, E. Gadisa, T. Kiros, M. Habtamu, J. Hussein, J. Zinsstag, B.D. Robertson, G. Ameni, A.J. Lohan, B. Loftus, I. Comas, S. Gagneux, R. Tschopp, L. Yamuah, G. Hewinson, S.V. Gordon, D.B. Young, A. Aseffa, Mycobacterial lineages causing pulmonary and extrapulmonary tuberculosis, Ethiopia, Emerg. Infect. Dis. 19 (2013) 460–463, https://doi.org/10.3201/eid1903.120256.

[31] H. Nebenzahl-Guimaraes, S.A. Yimer, C. Holm-Hansen, J. de Beer, R. Brosch, D. van Soolingen, Genomic characterization of Mycobacterium tuberculosis lineage 7 and a proposed name: 'Aethiops vetus, Microb. Genom. 2 (2016), e000063, https://doi.org/10.1099/mgen.0.000063.

[32] B. Wondale, K. Keehwan, G. Medhin, T. Teklu, T. Mohammed, S. Tolosa, A. Zewude, F. Amsalu, R. Pieper, G. Ameni, Molecular epidemiology of clinical Mycobacterium tuberculosis complex isolates in South Omo, Southern Ethiopia, BMC Infect. Dis. 20 (2020) 750, https://doi.org/10.1186/s12879-020-05394-9.

[33] S. Ayalew, T. Wegayehu, H. Taye, L. Wassie, S. Girma, S. Berg, A. Mihret, Drug resistance conferring mutation and genetic diversity of Mycobacterium tuberculosis isolates in tuberculosis lymphadenitis patients; Ethiopia, Infect. Drug Resist. 14 (2021) 575–584, https://doi.org/10.2147/IDR.S298683.

[34] A. Mengistu, F. Enquselassie, E. Hailu, A. Aseffa, D. Beyene, Identification and characterization of Mycobacterium tuberculosis isolates from cattle owners in north western and north eastern parts of rural Ethiopia, Int. J. Innovat. Appl. Stud. 10 (2015) 85–94.

[35] A. Mihret, Y. Bekele, A.G. Loxton, A.M. Jordan, L. Yamuah, A. Aseffa, R. Howe, G. Walzl, Diversity of *Mycobacterium tuberculosis* isolates from new pulmonary tuberculosis cases in Addis Ababa, Ethiopia, Tuberc. Res. Treat. 2012 (2012) 1–7, https://doi.org/10.1155/2012/892079.

[36] S.A. Yimer, G. Norheim, A. Namouchi, E.D. Zegeye, W. Kinander, T. Tønjum, S. Bekele, T. Mannsåker, G. Bjune, A. Aseffa, C. Holm-Hansen, 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 (2015) 1301–1309, https://doi.org/10.1128/JCM.03566-14.

[37] Y. Blouin, Y. Hauck, C. Soler, M. Fabre, R. Vong, C. Dehan, G. Cazajous, P.-L. Massoure, P. Kraemer, A. Jenkins, E. Garnotel, C. Pourcel, G. Vergnaud, Significance of the identification in the horn of Africa of an exceptionally deep branching Mycobacterium tuberculosis clade, PLoS One 7 (2012), e52841, https://doi.org/10.1371/journal.pone.0052841.

[38] M. Coscolla, S. Gagneux, Consequences of genomic diversity in Mycobacterium tuberculosis, Semin. Immunol. 26 (2014) 431–444, https://doi.org/10.1016/j.smim.2014.09.012.

[39] B. Haile, K. Tafess, A. Zewude, B. Yenew, G. Siu, G. Ameni, Spoligotyping and drug sensitivity of Mycobacterium tuberculosis isolated from pulmonary tuberculosis patients in the Arsi Zone of southeastern Ethiopia, New Microbes New Infect 33 (2020), 100620, https://doi.org/10.1016/j.nmni.2019.100620.

[40] A. Nuru, G. Mamo, A. Worku, A. Admasu, G. Medhin, R. Pieper, G. Ameni, Genetic diversity of *Mycobacterium tuberculosis* complex isolated from tuberculosis patients in bahir dar city and its surroundings, northwest Ethiopia, BioMed Res. Int. 2015 (2015) 1–9, https://doi.org/10.1155/2015/174732.

[41] B. Tessema, J. Beer, M. Merker, F. Emmrich, U. Sack, A.C. Rodloff, S. Niemann, Molecular epidemiology and transmission dynamics of Mycobacterium tuberculosis in Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia, BMC Infect. Dis. 13 (2013) 131, https://doi.org/10.1186/1471-2334-13-131.

[42] M. Tadesse, G. Abebe, A. Bekele, M. Bezabih, P. de Rijk, C.J. Meehan, B.C. de Jong, L. Rigouts, 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 (2017) 251–259, https://doi.org/10.1016/j.meegid.2017.09.016.

[43] B. Hussien, A. Zewude, B. Wondale, A. Hailu, G. Ameni, Spoligotyping of clinical isolates of Mycobacterium tuberculosis complex species in the Oromia region of Ethiopia, Front. Public Health 10 (2022), 808626, https://doi.org/10.3389/fpubh.2022.808626.

[44] O. Kisa, G. Tarhan, S. Gunal, A. Albay, R. Durmaz, Z. Saribas, T. Zozio, A. Alp, I. Ceyhan, A. Tombak, N. Rastogi, Distribution of spoligotyping defined genotypic lineages among drug-resistant Mycobacterium tuberculosis complex clinical isolates in ankara, Turkey, PLoS One 7 (2012), e30331, https://doi.org/10.1371/journal.pone.0030331.

[45] M. Agonafir, E. Lemma, D. Wolde-Meskel, S. Goshu, A. Santhanam, F. Girmachew, D. Demissie, M. Getahun, M. Gebeyehu, D. van Soolingen, Phenotypic and genotypic analysis of multidrug-resistant tuberculosis in Ethiopia, Int. J. Tubercul. Lung Dis. 14 (2010) 1259–1265.

[46] G. Diriba, A. Kebede, H.H. Tola, B. Yenew, S. Moga, D. Addise, A. Alemu, Z. Mohammed, M. Getahun, M. Fantahun, M. Tadesse, B. Dagne, M. Amare, G. Assefa, D. Abera, K. Desta, Molecular characterization and drug resistance patterns of Mycobacterium tuberculosis complex in extrapulmonary tuberculosis patients in Addis Ababa, Ethiopia, PLoS One 15 (2020), e0243493, https://doi.org/10.1371/journal.pone.0243493.

[47] L. Garedew, A. Mihret, G. Mamo, T. Abebe, R. Firdessa, Y. Bekele, G. Ameni, Strain diversity of mycobacteria isolated from pulmonary tuberculosis patients at Debre Birhan Hospital, Ethiopia, Int. J. Tubercul. Lung Dis. 17 (2013) 1076–1081, https://doi.org/10.5588/ijtld.12.0854.

[48] H. Taye, K. Alemu, A. Mihret, S. Ayalew, E. Hailu, J.L.N. Wood, Z. Shkedy, S. Berg, A. Aseffa, The ETHICOBOTS consortium, Epidemiology of Mycobacterium tuberculosis lineages and strain clustering within urban and peri-urban settings in Ethiopia, PLoS One 16 (2021), e0253480, https://doi.org/10.1371/journal.pone.0253480.

[49] V.N. Chihota, A. Niehaus, E.M. Streicher, X. Wang, S.L. Sampson, P. Mason, G. Källenius, S.G. Mfinanga, M. Pillay, M. Klopper, W. Kasongo, M.A. Behr, N.C. Gey van Pittius, P.D. van Helden, D. Couvin, N. Rastogi, R.M. Warren, Geospatial distribution of Mycobacterium tuberculosis genotypes in Africa, PLoS One 13 (2018), e0200632, https://doi.org/10.1371/journal.pone.0200632.

[50] B. Gumi, E. Schelling, S. Berg, R. Firdessa, G. Erenso, W. Mekonnen, E. Hailu, E. Melese, J. Hussein, A. Aseffa, J. Zinsstag, Zoonotic transmission of tuberculosis between pastoralists and their livestock in south-east Ethiopia, EcoHealth 9 (2012) 139–149, https://doi.org/10.1007/s10393-012-0754-x.

[51] S. Ali, P. Beckert, A. Haileamlak, A. Wieser, M. Pritsch, N. Heinrich, T. Löscher, M. Hoelscher, S. Niemann, A. Rachow, Drug resistance and population structure of M.tuberculosis isolates from prisons and communities in Ethiopia, BMC Infect. Dis. 16 (2016) 687, https://doi.org/10.1186/s12879-016-2041-x.

[52] F. Biadglegne, M. Merker, U. Sack, A.C. Rodloff, S. Niemann, 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 (2015), e0137865, https://doi.org/10.1371/journal.pone.0137865.

[53] M. Belay, G. Ameni, G. Bjune, D. Couvin, N. Rastogi, F. Abebe, Strain diversity of *Mycobacterium tuberculosis* isolates from pulmonary tuberculosis patients in Afar pastoral region of Ethiopia, BioMed Res. Int. 2014 (2014) 1–12, https://doi.org/10.1155/2014/238532.

[54] M. Maru, S.H. Mariam, T. Airgecho, E. Gadissa, A. Aseffa, Prevalence of tuberculosis, drug susceptibility testing, and genotyping of mycobacterial isolates from pulmonary tuberculosis patients in dessie, Ethiopia, Tuberc. Res. Treat. 2015 (2015) 1–10, https://doi.org/10.1155/2015/215015.
[55] G. Satta, M. Lipman, G.P. Smith, C. Arnold, O.M. Kon, T.D. McHugh, Mycobacterium tuberculosis and whole-genome sequencing: how close are we to unleashing its full potential? Clin. Microbiol. Infect. 24 (2018) 604–609, https://doi.org/10.1016/j.cmi.2017.10.030.
[56] V. Nikolayevskyy, S. Niemann, R. Anthony, D. van Soolingen, E. Tagliani, C. Ködmön, M.J. van der Werf, D.M. Cirillo, Role and value of whole genome sequencing in studying tuberculosis transmission, Clin. Microbiol. Infect. 25 (2019) 1377–1382, https://doi.org/10.1016/j.cmi.2019.03.022.