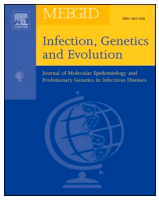
Infection, Genetics and Evolution 116 (2023) 105530

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid



Genomic transmission clusters and circulating lineages of *Mycobacterium tuberculosis* among refugees residing in refugee camps in Ethiopia

Abyot Meaza [a,b,*], Emmanuel Riviere [c], Zegeye Bonsa [d], Vincent Rennie [c], Gebremedhin Gebremicael [b], Miguel de Diego-Fuertes [c], Conor J. Meehan [e], Girmay Medhin [a], Gemeda Abebe [d,g], Gobena Ameni [a,f], Annelies Van Rie [c], Balako Gumi [a]

[a] *Aklilu Lemma Institute of Pathobiology (ALIPB), Addis Ababa University (AAU), PO Box 1176, Sefere Selam campus, Addis Ababa, Ethiopia*
[b] *Ethiopian Public Health Institute (EPHI), PO Box 1242, Swaziland Street, Addis Ababa, Ethiopia*
[c] *Family Medicine and Population Health (FAMPOP), Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium*
[d] *Mycobacteriology Research Center, Jimma University, Jimma, Ethiopia*
[e] *Department of Biosciences, Nottingham Trent University, Nottingham, UK*
[f] *Department of Veterinary Medicine, College of Agriculture and Veterinary Medicine, United Arab Emirates University, PO Box 15551, Al Ain, United Arab Emirates*
[g] *Department of Medical Laboratory Sciences, Jimma University, Jimma, Ethiopia*

ARTICLE INFO

Keywords:
Mycobacterium tuberculosis
Mtb
WGS
L9
Transmission clusters
Refugees

ABSTRACT

Background: Understanding the transmission dynamics of *Mycobacterium tuberculosis (Mtb)* could benefit the design of tuberculosis (TB) prevention and control strategies for refugee populations. Whole Genome Sequencing (WGS) has not yet been used to document the *Mtb* transmission dynamics among refugees in Ethiopia. We applied WGS to accurately identify transmission clusters and *Mtb* lineages among TB cases in refugee camps in Ethiopia.

Method and design: We conducted a cross-sectional study of 610 refugees in refugee camps in Ethiopia presenting with symptoms of TB. WGS data of 67 isolates was analyzed using the Maximum Accessible Genome for *Mtb* Analysis (MAGMA) pipeline; iTol and FigTree were used to visualize phylogenetic trees, lineages, and the presence of transmission clusters.

Results: Mtb culture-positive refugees originated from South Sudan (52/67, 77.6%), Somalia (9/67, 13.4%). Eritrea (4/67, 6%), and Sudan (2/67, 3%). The majority (52, 77.6%) of the isolates belonged to *Mtb* lineage (L) 3, and one L9 was identified from a Somalian refugee. The vast majority (82%) of the isolates were pan-susceptible *Mtb*, and none were multi-drug-resistant (MDR)-TB. Based on the 5-single nucleotide polymorphisms cutoff, we identified eight potential transmission clusters containing 23.9% of the isolates. Contact investigation confirmed epidemiological links with either family or social interaction within the refugee camps or with neighboring refugee camps.

Conclusion: Four lineages (L1, L3, L4, and L9) were identified, with the majority of strains being L3, reflecting the *Mtb* L3 dominance in South Sudan, where the majority of refugees originated from. Recent transmission among refugees was relatively low (24%), likely due to the short study period. The improved understanding of the *Mtb* transmission dynamics using WGS in refugee camps could assist in designing effective TB control programs for refugees.

Abbreviations: AAU, Addis Ababa University; DR, drug resistance; EPHI, Ethiopian Public Health Institute; ETM, ethionamide; INH, isoniazid; LJ, Lowstein Jensen; L, Lineage (s); MAGMA, Maximum Accessible Genome for *Mycobacterium tuberculosis* Analysis; *Mtb*, *Mycobacterium tuberculosis*; MDR, multi-drug resistant; MGIT, Mycobacterium Growth Indicator Tube; NTM, non-tuberculous mycobacteria; QC, Quality Control; RIF, Rifampicin; RR, Resistance to Rifampicin; SNPs, single nucleotide polymorphisms; SM, streptomycin; TB, Tuberculosis; WGS, Whole Genome Sequencing; WHO, World Health Organization; XBS, complex bacterial samples; XDR, extensively drug resistant..

* Corresponding author at: ALIBP-AAU and EPHI, Addis Ababa, Ethiopia.
E-mail address: abimeaza@gmail.com (A. Meaza).

https://doi.org/10.1016/j.meegid.2023.105530
Received 12 July 2023; Received in revised form 15 November 2023; Accepted 19 November 2023
Available online 25 November 2023

1. Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), is one of the world's major causes of death from an infectious agent, with over 1.5 million deaths per year (WHO, 2022). In Ethiopia, TB is a major public health concern, with an estimated TB incidence of 119 cases per 100, 000, annually (WHO, 2022). The prevalence of MDR-TB is 1.03% among new TB patients and 6.52% among previously treated TB patients (FMOH, 2022). Although Ethiopia is no longer included in the World Health Organization (WHO) list of 30 high MDR/Resistance to Rifampicin (*RR*)-TB burden countries, it is still included in the WHO list of 30 high TB burden countries (FMOH, 2022; WHO, 2022).

TB particularly affects poor and vulnerable populations such as refugees, internally displaced persons, migrants, and other populations in humanitarian settings at increased risk of developing TB (WHO-CDC, 2022). Refugees are especially at high risk of developing TB due to the poor living conditions they experience in refugee-hosting countries, leading to high TB morbidity and mortality in refugee populations (Connolly et al., 2007; Figueroa-Munoz and Ramon-Pardo, 2008; WHO-CDC, 2022). Consequently, the incidence and prevalence of TB among refugees and migrant populations are higher than those among non-refugee populations (Meaza et al., 2022). This can significantly affect TB control in refugee-hosting countries by increasing the overall disease burden and the cost of health services (Figueroa-Munoz and Ramon-Pardo, 2008). At the end of June 2021, Ethiopia was the third-largest refugee-hosting country in Africa and hosted 785,322 registered refugees, primarily from neighboring countries such as South Sudan, Sudan, Eritrea, and Somalia (UNHCR, 2022). The number of notified TB cases among refugees in Ethiopia has increased from 138 cases in 2014 to 588 cases in 2017 (Legesse et al., 2021). Similarly, a recent study has shown that the trend of all forms of TB increased among refugees compared to the surrounding communities of Gambella in Ethiopia (Ejeta et al., 2018).

Nearly 235 million people are estimated to be in need of humanitarian assistance and protection globally in 2021. Of these, 26.4 million were refugees (WHO-CDC, 2022). The number of refugees in the Horn of Africa exceeded 3 million at the end of 2022 (UNHCR, 2022). Migrants from the Horn of Africa, particularly from Somalia and Eritrea, have been identified as a potential source of transmission of the *Mtb* strain in Europe (Martínez-Lirola et al., 2021). Efforts toward the prevention and control of TB in refugee settings would not only promote the health of refugees but also the surrounding communities and, thus, the overall public health by reducing *Mtb* transmission.

Using spoligotyping of isolates from refugees in Ethiopia, we found that SIT25, CAS1-Delhi, and L3 were the predominant *Mtb* strain, family, and lineage respectively (Meaza et al., 2023a). Identifying *Mtb* transmission clusters with WGS is a more powerful approach for investigating *Mtb* transmission dynamics (Alaridah et al., 2019; Asare et al., 2020; Perdigao et al., 2017). WGS is becoming the reference method to investigate the phylogenetic background of *Mtb* isolates due to its ability of WGS to generate high-resolution data on transmission clusters and its potential to define the extent and direction of TB transmission. Implementing WGS for transmission studies can be used to distinguish between recent transmission, reinfection, and progression to active disease in previously infected individuals (Gardy et al., 2011; Meehan et al., 2019; Walker et al., 2014). The current study aimed to accurately identify the transmission clusters and circulating *Mtb* lineages among TB cases in selected refugee camps in Ethiopia using WGS.

2. Methods

2.1. Study design and study inclusion

This study was a secondary analysis of a cross-sectional study of 610 consecutive refugees presenting for symptoms of TB between February and August 2021 in 12 refugee camps in Ethiopia (Mai Ani, Asaita, Pugnido-Nuer, Pugnido-Agnwak, Kule, Terkeidi, Sherkole, Bambasi, Tongo, Kebribeyah, Sheder, and Melkadida) (Meaza et al., 2022; Meaza et al., 2023b). Of the 610 participants, 71 (11.6%) were culture-positive, 37 (6.1%) were NTM, 96 (15.7%) were contaminated, and 406 (66.6%) were culture-negative.

Of the 71 Mycobacterium Growth Indicator Tube (MGIT) *Mtb*-positive isolates, 3 were excluded due to poor growth in the Lowenstein Jensen (LJ) subculture. DNA was extracted from 68 *Mtb* isolates with typical growth characteristics (appearance of brown and granular colonies on LJ media (Fig. 1). One sample was excluded due to insufficient DNA quality, resulting in 67 isolates with WGS data for analysis.

2.2. DNA extraction and WGS

DNA was extracted from the LJ sub-cultures of 68 *Mtb* isolates at the Mycobacteriology Research Center of Jimma University using the Qiagen DNeasy UltraClean Microbial Kit (lot number: 172046262) (Qiagen, 2017). The extracted *Mtb* DNA was shipped to CD Genomics (NY, USA) for WGS. Quality control (QC) was performed using both Nanodrop and Qubit to determine DNA concentration and quantity. Of the 68 samples, 67 passed QC (Supporting Material 1). From each sample, 0.3 μg of extracted DNA was used as input for library preparation. Genomic DNA libraries were constructed using the NEBNextR Ultra™ DNA Library Prep Kit (NEB, USA) according to the manufacturer's instructions (Illumina, 2022). Libraries were analyzed for size distribution by an Agilent 2100 bioanalyzer and quantified using a real-time polymerase chain reaction. DNA libraries were then sequenced on an Illumina NovaSeq 6000 instrument according to standard protocols (Illumina, 2020) (Fig. 1).

2.3. WGS data analysis

The generated reads were analyzed using the Maximum Accessible Genome for *Mtb* Analysis (MAGMA) pipeline (Heupink et al., 2023). MAGMA is a bioinformatics pipeline that implements the XBS variant calling core (Heupink et al., 2021). The MAGMA pipeline is implemented in Nextflow and is available open source on GitHub (https://github.com/TORCH-Consortium/MAGMA). Briefly, Illumina FASTQ sequence data was quality-controlled and mapped to the H37RV reference genome. All samples were checked for selection criteria of median coverage (>90×), breadth of coverage (>0.90), variant frequency (>0.80), and NTM frequency (<0.20) to filter out samples that negatively impact the downstream analysis. As part of the MAGMA pipeline, lineage and drug resistance-conferring variants for each sample, SNP, and INDEL variants, are analyzed by TBprofiler v. 4.1.1 (Verboven et al., 2022). The filtered SNP and INDEL variant call format (VCF) files produced were further processed using the downstream analysis modules to construct annotated phylogenetic trees, perform drug-resistant variant detection, and transmission cluster analysis (Supporting Material 2). Given the high TB incidence of the study setting, a pairwise distance cutoff of 5-SNPs was selected to define transmission clusters (Luo et al., 2014; Oostvogels et al., 2022). A sensitivity analysis using a 12 SNP cutoff was also performed.

2.3.1. Analysis of L9 data

We compared our L9 sample's (R-192) spoligotype data (Meaza et al., 2023a) with the spoligotype data of six L9 strains reported to date (Coscolla et al., 2021) using the methods described by Kamerbeek et al., 1997 (Kamerbeek et al., 1997) with a commercially available membrane following the manufacturer's instructions (Mapmygenome, India). To assess whether the phylogenetic position of the R-192 sample is situated within the clade of recently discovered L9 panel strains, a maximum likelihood phylogeny was inferred using FigTree v. 1.4.4 for the L9 sample (R-192) from this study and the six previously identified L9 strains. The bootstraps were generated using the standard IQtree with a bootstrap of 1000 runs.

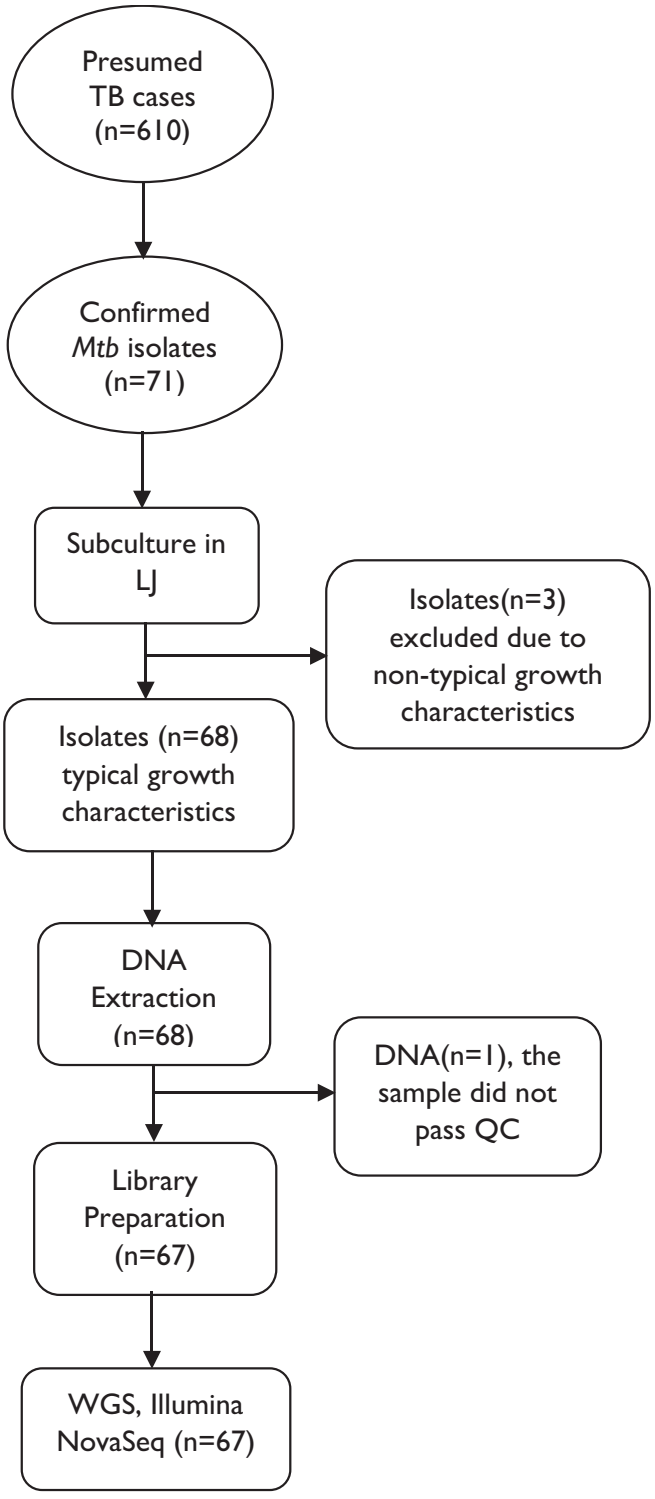


Fig. 1. Laboratory investigation process diagram. Typical growth characteristics: *Mtb* on LJ media appears as brown, and granular colonies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Demographic and clinical characteristics

Of the 67 refugees included in the analysis, 52 (78%) originated from South Sudan, 9 (13%) from Somalia, 4 (6%) from Eritrea, and 2 (3%) from Sudan. Most were male (67%) and age from 19 to 38 years old (63%). Fifteen (22.4%) reported a history of TB contact.

3.2. WGS quality control

All QC results of the DNA extraction and WGS are provided in supporting materials 2 and 3, respectively. The median coverage of the samples ranged from 158 to 330×, and on average, 99% (0.99×) of the genomes were covered. Mixed strains were not identified, possibly due to the use of subcultures for WGS (Supporting Material 4).

3.3. Lineage identification

Most isolates were L3, both overall (52/67, 77.6%) and in samples collected from South Sudanese (47/52, 90.4%) and Sudanese (2/2, 100%) refugees. Eleven (11/67, 16.4%) isolates were L4, and three isolates (3/67, 4.5%) were L1. One isolate (R192) from Somali refugee, was identified as L9 by TBprofiler (Fig. 2). In total, 14 different sublineages belonging to the major lineages (L1, L3, and L4) were identified (Supporting Material 5).

A comparison of spoligotyping data demonstrated that sample R-192 has a different spoligotyping pattern (Meaza et al., 2023a) compared to the spoligotyping patterns of the six previously reported L9 strains (Coscolla et al., 2021) (Fig. 3a). Based on the maximum likelihood phylogeny tree, the clade of previously identified L9 strains and sample R-192 were phylogenetically closely related and descended from a common ancestor (Fig. 3b). The SNP distances of the R-192 sample to the previously identified L9 isolates ranged from 86 to 512 SNPs (supporting Material 6).

3.4. Genotypic DR-causing variant detection

Most isolates (55/67, 82%) were genotypically pan-susceptible; ten (10/67, 15.0%) were resistant to a single anti-TB drug, mainly (8/10) streptomycin. All eight samples predicted to be resistant to streptomycin (SM) were identified as L3. Resistance to more than one drug was found in only two (2/67, 3.0%) samples. Sample R-044 was predicted to be resistant to SM and ethionamide (ETM), while sample R-171 was predicted to be resistant to isoniazid (INH) and ETM. None of the isolates collected in this study were identified as MDR-TB, pre-extensively drug-resistant (XDR)-TB, or XDR-TB (Table 1).

3.5. Transmission clusters

Of the 67 samples analyzed in this study, 16 (16/67, 23.9%) belonged to one of eight putative transmission clusters with a $\leq$ 5 SNP distance cut-off (Fig. 2). The pairwise maximum genetic distance between samples in putative transmission clusters ranged from zero to two SNPs (Supporting Material 7). Five of the clusters were pairs of South Sudanese refugees; two clusters were pairs of Somali refugees; and one cluster was a pair of Sudanese refugees. Contact investigation demonstrated that our clusters had epidemiological links, either family history or social interaction within the refugee camps and with neighboring refugee camps (Supporting Material 8). In a sensitivity analysis using a 12 SNP cutoff, the same 8 clusters were identified as compared to the 5 SNP cutoff.

4. Discussion

This study aimed to use WGS to accurately investigate the TB transmission dynamics and circulating lineages among refugees residing in refugee camps in Ethiopia.

Four human-adapted lineages (L1, L3, L4, and L9) of *Mtb* were identified. Most *Mtb* isolates in this study were identified as L3, and one sample from a Somalian refugee was identified as L9, making this the first additional report of L9 sample since the discovery of the lineage in

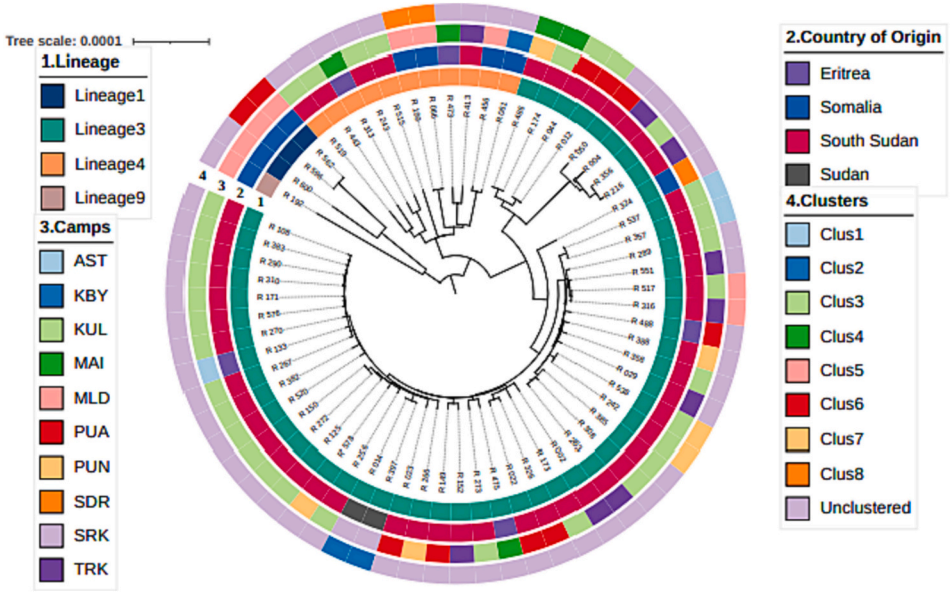
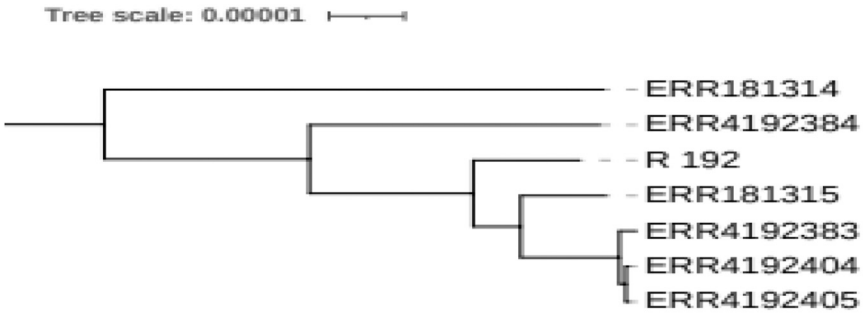


Fig. 2. Phylogeny tree showing lineages of *Mtb* isolates [circle 1], country of origin [circle 2], refugee camps [circle 3] and clusters (Clus) [circle 4] identified in our samples. Camps and country of origin: AST-Asaita, and MAI-MaiAni (Eritrea), KBY-Kebribeya, MLD-Melkadida and SDR-Sheder (Somalia), KUL-Kule, PUA-PugnidoAgnwak, PUN-PugnidoNuer, and TRK-Terkedi (South Sudan), SRK-Sherkole (Sudan).

Sample number	Country of Origin	Spoligo 43	Spoligo Octal	Source
R-192	Somalia		772701002060771	This study
ERR181315	Somalia		700000007177771	Coscolla et al, 2021
ERR4192383	Somalia		772700000003671	Coscolla et al, 2021
ERR4192404	Somalia		772700000003671	Coscolla et al, 2021
ERR4192405	Somalia		772700000003671	Coscolla et al, 2021
ERR181314	Djibouti		772000007775671	Coscolla et al, 2021
ERR4192384	Europe		772600000003631	Coscolla et al, 2021

a,



b,

Fig. 3. a. Spoligotyping patterns and their country of origin of our L9 sample (R-192) and six samples from previously described by Cosccola et al. b. a maximum likelihood phylogeny tree showing the phylogenetic position of R-192 in the context of a previously described panel of six L9 strains by Coscolla et al. The phylogenetic rectangular tree was midpoint rooted using FigTree v.1.4.4 and further visualized automatically on the nodes using iTol, version 6. The tree has two major branches with seven leaves and the bootstraps was generated using the standard IQtree with a bootstrap of 1000 runs.

2021 (Coscolla et al., 2021). Eight *Mtb* clusters, representing 23.9% of the study population, were identified and confirmed by epidemiological links.

The lineage distribution of the *Mtb* isolates in the 12 refugee camps in Ethiopia was different from the distribution of *Mtb* lineages in the general population of Ethiopia. In contrast to the dominance (77.6%) of L3

Table 1
Drug resistance variants among DR-TB isolates ($n = 12$) and lineages (L3 & L4) by WGS.

		anti-TB drug															
ID	**L**	RMP	INH	EMB	PZA	SM	FQ	AGS	KAN	AMK	CAP	ETM	*P*-AGS	CLO	LZD	BQ	DM
R-002	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S
R-012	3	S	S	S	S	S	S	S	S	S	S	R	S	S	S	S	S
R-044	3	S	S	S	S	R	S	S	S	S	S	R	S	S	S	S	S
R-061	4	S	S	S	S	S	S	S	S	S	R	S	S	S	S	S	S
R-171	3	S	R	S	S	S	S	S	S	S	S	R	S	S	S	S	S
R-256	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S
R-273	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S
R-316	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S
R-382	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S
R-456	4	S	S	S	S	S	S	S	S	S	R	S	S	S	S	S	S
R-517	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S
R-551	3	S	S	S	S	R	S	S	S	S	S	S	S	S	S	S	S

AGS-Aminoglycosides AMK-Amikacin, BQ-Bedaquiline, CAP-Capreomycin, CLO-Clofazimine, DM-Delamanid EMB-Ethambutol, ETM-Ethionamide, FLQ-Fluoroquinolones, INH-Isoniazid, KAN-Kanamycin, LEV-Levofloxacin, LZD-Linezolid OFL-Ofloxacin, *P*-AGA-Para-aminosalicylic acid, PZA-Pyrazinamide, RIF-Rifampicin, SM-Streptomycin, L-lineage, R-Resistant, S-Susceptible.

in our study of refugees, L4 is the dominant *Mtb* lineage in the general population of Ethiopia, as shown by two recent systematic reviews (L4 in 62.3% and 64.8%) (Mekonnen et al., 2019; Tulu and Ameni, 2018). Similarly, 74.7% of 91 *Mtb* isolates belonged to L4 in an Ethiopian national TB survey (Getahun et al., 2015). In contrast, studies in Sudan and Eritrea showed that the dominant lineage is L3 (Eldirdery et al., 2015; Elegail et al., 2018; Mesfin et al., 2021). Our results show that the distribution of *Mtb* lineages among the refugees reflects the distribution of *Mtb* lineages in their countries of origin, which suggests limited transmission between refugees and the Ethiopian population. These data highlight the challenges of TB control in unstable regions and the importance of cross-border TB surveillance, especially in areas of intense migration.

One sample originating from a Somalian refugee was identified as L9. This L9 *Mtb* strain was phylogenetically closely related to the six previously reported L9 strains (Coscolla et al., 2021), indicating that these strains descended from a common ancestor. The identification of a L9 strain from a Somalian refugee confirms previous findings that L9 is currently geographically primarily restricted to Somalia (Coscolla et al., 2021; Silva et al., 2022). The identification of the L9 sample (R-192) in this study adds to the limited set of sequenced L9 *Mtb* strains and can support studies aimed at identifying the phylogenetic markers and phenotypic data of L9 *Mtb* strains.

The vast majority (82%) of the 67 isolates were pan-susceptible by WGS, and only 12 (18%) contained one or more DR-causing variants, mostly to streptomycin. No MDR-TB or pre-XDR-TB isolates were identified. Recently, the WHO removed Ethiopia from the list of high MDR/RR-TB countries, highlighting the successful implementation of an effective TB control program in the country (FMOH, 2022; WHO, 2022). Compared to the findings of this study, the WHO global TB report (WHO, 2022) shows a higher proportion [1.6 to 4.4%] of MDR/RR-TB among new cases in the four refugee countries of origin (Eritrea, Somalia, South Sudan, and Sudan), with Somalia having the highest (4.4%) estimated MDR/RR-TB prevalence. Interestingly, WGS failed to detect rifampicin resistance in a sample (R-171) that was rifampicin resistant on phenotypic drug susceptibility testing (Meaza et al., 2023a). This could be due to culture bias introduced by the sub-culture on LJ performed prior to DNA extraction.

In this study, 23.9% (16/67) of the isolates belonged to one of eight phylogenetic clusters. Other studies among refugees, albeit not in refugee camps, have reported varying proportions of recent transmission. A study among immigrants in Switzerland reported a lower proportion (6.5%, 26/401) of immigrants belonging to transmission clusters (Stucki et al., 2016). In contrast, a cross-border transmission study from Canada showed that a very high proportion (87.5%, 28/32 in 6 clusters) of *Mtb* isolates were grouped in a phylogenetic cluster (Guthrie et al., 2019). Similarly, in a study from Sweden, 56% (52/93 in 18 clusters) of isolates originating from immigrants and homeless individuals were part of phylogenetic clusters (Alaridah et al., 2019). The lower clustering proportion in our study could be due to differences in sampling methods, SNP cut-off used to define a cluster difference, or the short sampling period of 6 months compared to the studies in Canada and Sweden, where samples were collected over a period of ten years. A long sample collection period has been shown to be required to accurately determine the number of secondary TB cases produced by a primary infectious TB case (Basu et al., 2009; Liao et al., 2012).

Epidemiological contact tracing coupled with WGS data is important to understand TB transmission (Alaridah et al., 2019). In our study, all genomic clusters identified were supported by the epidemiological links from contact investigations. Similar findings were reported in other studies (Abascal et al., 2019; Alaridah et al., 2019; Stucki et al., 2016). In contrast, no epidemiological links were found in the clustered cases in the study in Canada (Guthrie et al., 2019). The epidemiological links we found for all clusters in the refugee camps highlight the importance of contact tracing for early diagnosis and treatment of missed cases to slow the spread of TB infection.

We observed discrepancies in six samples in lineage classification by WGS and spoligotyping. WGS classified three samples as L3, which spoligotyping identified as L4 in two samples and L1 in one sample. WGS identified L4 and L1 for each sample that had been identified as L3 and L2, respectively, by spoligotyping. The sample classified as L9 by WGS was classified as L4 by spoligotyping (Meaza et al., 2023a). This is likely due to the higher discriminatory power of WGS to identify *Mtb* lineages compared to spoligotyping (Napier et al., 2023; Oudghiri et al., 2018).

While this is the first study applying WGS to study the molecular epidemiology of TB in refugee camps, our results should be interpreted in light of its limitations. First, the short study duration (6 months) likely resulted in an underestimation of *Mtb* transmission in refugee camps. Future studies should extend the duration of data collection to more comprehensively capture transmission. Second, we only studied the transmission dynamics within refugee camp populations and not the neighboring communities. To optimally design interventions, future studies should expand to investigate TB transmission within a broader context that also includes the communities surrounding refugee camps. Third, performing WGS on an LJ subculture could have resulted in culture bias. Future studies should perform WGS directly on the primary MGIT culture or, when this becomes feasible, directly on a sputum sample.

5. Conclusion

In conclusion, our results highlight the importance of WGS for

improved understanding and surveillance of TB in refugee camps. We identified the distribution of lineages that represent the country of origin of the refugees, the occurrence of an L9 *Mtb* strain from a Somalian refugee, and a low level of DR. We also observed a relatively low proportion of transmission clusters, which were confirmed by epidemiological links that generated important evidence. This evidence can be used to design effective interventions to reduce the burden of TB in refugee camps. The findings highlight the value of cross-border surveillance and the need for increased efforts to prevent transmission in the vulnerable refugee population.

Contributors

AM designed the study; BG and AVR oversaw the study. AM, ER, ZB, G. Abebe and GG contributed to the DNA extraction and the shipment process of kits and DNA. Bioinformatics Analysis was done by AM, VR, and MDF. The statistical analysis was undertaken by AM and reviewed by GM. The manuscript draft was developed by AM and reviewed by VR, ER, CM, AVR, GM, G. Abebe, GA, and BG. All authors read and approved the final version of the manuscript.

Funding source

AM received an award for sample preparation, DNA extraction, DNA shipment, and WGS from the VLIRUOS project (project number: ET2018JOI008A102), University of Antwerp. The funder has no role in the study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Ethical approval and Consent to participate

The study was approved by the institutional review board of EPHI (protocol number: EPHI-IRB-200-2019) and AAU (reference number: ALIPB IRB/33/2013/20). Written informed consent or assent was obtained from the study participants prior to data and sample collection. The extracted DNA was shipped to CD Genomics for WGS based on a signed service agreement including terms and conditions for material transfer between AAU, Ethiopian Biodiversity Institute, and CD Genomics.

Raw sequence accession number

The raw sequence data for all samples (67) including one sample containing L9 (isolate number: R-192) from this study are deposited to the GenBank Data Libraries (https://www.ncbi.nlm.nih.gov) under accession No. PRJNA1010675.

Declaration of Competing Interest

We declare that there is no conflict of interest that might have influenced the performance and presentation of the work described in this manuscript.

Data availability

All relevant data are within the manuscript and additional data also available as supporting materials. Furthermore, the raw sequence data for 67 samples (one L9 sample and 66 samples containing L1, L3 and L4) from this study are deposited to the GenBank Data Libraries (https://www.ncbi.nlm.nih.gov/sra) under accession No. PRJNA1010675.

Acknowledgments

We would like to thank AAU, EPHI, Jimma University, and CD Genomics for supporting the study. We also would like to thank VLIR project (project number: ET2018JOI008A102) and the TORCH research team at the University of Antwerp for funding the cost for sample preparation and WGS.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2023.105530.

References

Abascal, E., Pérez-Lago, L., Martínez-Lirola, M., Chiner-Oms, Á., Herranz, M., Chaoui, I., Comas, I., El Messaoudi, M.D., Cárdenas, J.A.G., Santantón, S., 2019. Whole genome sequencing–based analysis of tuberculosis (TB) in migrants: rapid tools for cross-border surveillance and to distinguish between recent transmission in the host country and new importations. Eurosurveillance 24 (4), 1800005.

Alaridah, N., Hallbäck, E.T., Tångrot, J., Winqvist, N., Sturegård, E., Florén-Johansson, K., Jönsson, B., Tenland, E., Welinder-Olsson, C., Medstrand, P., 2019. Transmission dynamics study of tuberculosis isolates with whole genome sequencing in southern Sweden. Sci. Rep. 9 (1), 4931.

Asare, P., Otchere, I.D., Bedeley, E., Brites, D., Loiseau, C., Baddoo, N.A., Asante-Poku, A., Osei-Wusu, S., Prah, D.A., Borrell, S., 2020. Whole genome sequencing and spatial analysis identifies recent tuberculosis transmission hotspots in Ghana. Front. Med. 7, 161.

Basu, S., Friedland, G.H., Medlock, J., Andrews, J.R., Shah, N.S., Gandhi, N.R., Moll, A., Moodley, P., Sturm, A.W., Galvani, A.P., 2009. Averting epidemics of extensively drug-resistant tuberculosis. Proc. Natl. Acad. Sci. 106 (18), 7672–7677.

Connolly, M.I.A., Gayer, M., Ottmani, S.-E., Organization, W. H, 2007. Tuberculosis care and control in refugee and displaced populations: an interagency field manual. World Health Organization.

Coscolla, M., Gagneux, S., Menardo, F., Loiseau, C., Ruiz-Rodriguez, P., Borrell, S., Otchere, I.D., Asante-Poku, A., Asare, P., Sánchez-Busó, L., 2021. Phylogenomics of Mycobacterium africanum reveals a new lineage and a complex evolutionary history. Microb. Genom. 7 (2).

Ejeta, E., Beyene, G., Balay, G., Bonsa, Z., Abebe, G.J.P., o., 2018. Factors associated with unsuccessful treatment outcome in tuberculosis patients among refugees and their surrounding communities in Gambella Regional State, Ethiopia, 13 (10), e0205468.

Eldirdery, M.M., Alrayah, I.E., ElkareIm, M.O.A., Khalid, F.A., Ma, A., Elegail, S., Ibrahim, N.Y., Nour, E.O.M., Ali, R.H., Hailu, E., 2015. Genotyping of pulmonary Mycobacterium tuberculosis isolates from Sudan using spoligotyping. Am. J. Microbiol. Res. 3 (4), 125.

Elegail, A., Mohamed, N.Y.I., Nour, E.O.M., Hoffner, S., Haile, M., 2018. Molecular characterization of Mycobacterium tuberculosis isolates from pulmonary tuberculosis patients in Khartoum, Sudan. Int. J. Mycobacteriol. 7 (3), 236.

Figueroa-Munoz, J.I., Ramon-Pardo, P., 2008. Tuberculosis control in vulnerable groups. Bull. World Health Organ. 86, 733–735.

FMOH, 2022. Annual Performance Report, 2014 EFY/2021–2022. Ministry of Health, Ethiopia.

Gardy, J.L., Johnston, J.C., Sui, S.J.H., Cook, V.J., Shah, L., Brodkin, E., Rempel, S., Moore, R., Zhao, Y., Holt, R., 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N. Engl. J. Med. 364 (8), 730–739.

Getahun, M., Ameni, G., Kebede, A., Yaregal, Z., Hailu, E., Medihn, G., Demssie, D., Girmachew, F., Fiseha, Y., Meaza, A., 2015. Molecular typing and drug sensitivity testing of Mycobacterium tuberculosis isolated by a community-based survey in Ethiopia. BMC Public Health 15 (1), 1–7.

Guthrie, J., Strudwick, L., Roberts, B., Allen, M., McFadzen, J., Roth, D., Jorgensen, D., Rodrigues, M., Tang, P., Hanley, B., 2019. Whole genome sequencing for improved understanding of Mycobacterium tuberculosis transmission in a remote circumpolar region. Epidemiol. Infect. 147.

Heupink, T.H., Verboven, L., Warren, R.M., Van Rie, A., 2021. Comprehensive and accurate genetic variant identification from contaminated and low-coverage Mycobacterium tuberculosis whole genome sequencing data. Microb. Genom. 7 (11).

Heupink, Tim H., Verboven, Lennert, Sharma, Abhinav, Rennie, Vincent, de Diego Fuertes, Miguel, Warren, Robin M., Van Rie, Annelies, 2023. The MAGMA pipeline for comprehensive genomic analyses of clinical Mycobacterium tuberculosis samples. BMJ. https://doi.org/10.1101/2023.10.04.23296533. Preprint.

Illumina, 2020. NovaSeq 6000 Sequencing System Guide.

Illumina, 2022. NEBNext® Ultra™ DNA Library Prep Kit for Illumina. In: Instruction Manual.

Kamerbeek, J., Schouls, L., Kolk, A., Van Agterveld, M., Van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., 1997. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. J. Clin. Microbiol. 35 (4), 907–914.

Legesse, T., Admenur, G., Gebregzabher, S., Woldegebriel, E., Fantahun, B., Tsegay, Y., Bayssa, A., Darge, B., Denbu, Y., Michalel, H.J.B., i. d., 2021. Tuberculosis (TB) in the refugee camps in Ethiopia: trends of case notification, profile, and treatment outcomes, 2014 to 2017. BMC Infect. Dis. 2021 (21), 139.

Liao, C.M., Cheng, Y.H., Lin, Y.J., Hsieh, N.H., Huang, T.L., Chio, C.P., Chen, S.C., Ling, M.P., 2012. A probabilistic transmission and population dynamic model to assess tuberculosis infection risk. Risk Analysis Int. J. 32 (8), 1420–1432.

Luo, T., Yang, C., Peng, Y., Lu, L., Sun, G., Wu, J., Jin, X., Hong, J., Li, F., Mei, J., 2014. Whole-genome sequencing to detect recent transmission of Mycobacterium

tuberculosis in settings with a high burden of tuberculosis. Tuberculosis 94 (4), 434–440.
Martínez-Lirola, M., Jajou, R., Mathys, V., Martin, A., Cabibbe, A.M., Valera, A., Sola-Campoy, P.J., Abascal, E., Rodríguez-Maus, S., Garrido-Cárdenas, J.A., 2021. Integrative transnational analysis to dissect tuberculosis transmission events along the migratory route from Africa to Europe. J. Travel Med. 28 (4), taab054.
Meaza, A., Tola, H.H., Eshetu, K., Mindaye, T., Medhin, G., Gumi, B., 2022. Tuberculosis among refugees and migrant populations: systematic review. PLoS One 17 (6), e0268696.
Meaza, A., Diriba, G., Girma, M., Wondimu, A., Worku, G., Medhin, G., Ameni, G., Gumi, B., 2023a. Molecular typing and drug sensitivity profiles of M. Tuberculosis isolated from refugees residing in Ethiopia. J. Clinical Tuberculosis Other Mycobacterial Dis. 100371.
Meaza, A., Yenew, B., Amare, M., Alemu, A., Hailu, M., Gamtesa, D.F., Kaba, M., Medhin, G., Ameni, G., Gumi, B., 2023b. Prevalence of tuberculosis and associated factors among presumptive TB refugees residing in refugee camps in Ethiopia. BMC Infect. Dis. 23 (1), 1–9.
Meehan, C.J., Goig, G.A., Kohl, T.A., Verboven, L., Dippenaar, A., Ezewudo, M., Farhat, M.R., Guthrie, J.L., Laukens, K., Miotto, P., 2019. Whole genome sequencing of Mycobacterium tuberculosis: current standards and open issues. Nat. Rev. Microbiol. 17 (9), 533–545.
Mekonnen, D., Derbie, A., Chanie, A., Shumet, A., Biadglegne, F., Kassahun, Y., Bobosha, K., Mihret, A., Wassie, L., Munshea, A., 2019. Molecular epidemiology of M. Tuberculosis in Ethiopia: a systematic review and meta-analysis. Tuberculosis 118, 101858.
Mesfin, A., Araia, Z., Beyene, H., Mebrahtu, A., Suud, N., Berhane, Y., Hailu, D., Kassahun, A., Auguet, O., Dean, A., 2021. First molecular-based anti-TB drug resistance survey in Eritrea. Int. J. Tuberc. Lung Dis. 25 (1), 43–51.
Napier, G., Couvin, D., Refrégier, G., Guyeux, C., Meehan, C.J., Sola, C., Campino, S., Phelan, J., Clark, T.G., 2023. Comparison of in silico predicted Mycobacterium tuberculosis spoligotypes and lineages from whole genome sequencing data. Sci. Rep. 13 (1), 11368.
Oostvogels, S., Ley, S.D., Heupink, T.H., Dippenaar, A., Streicher, E.M., De Vos, E., Meehan, C.J., Dheda, K., Warren, R., Van Rie, A., 2022. Transmission, distribution and drug resistance-conferring mutations of extensively drug-resistant tuberculosis in the Western Cape Province, South Africa. Microb. Genom. 8 (4).
Perdigao, J., Clemente, S., Ramos, J., Masakidi, P., Machado, D., Silva, C., Couto, I., Viveiros, M., Taveira, N., Portugal, I., 2017. Genetic diversity, transmission dynamics and drug resistance of Mycobacterium tuberculosis in Angola. Sci. Rep. 7 (1), 1–10.
Oudghiri, A., Chaoui, I., Elmzibri, M., 2018. Molecular Epidemiology of Tuberculosis: A Review of Tools and Applications. J. Infect Dis. Ther. 6 (6) https://doi.org/10.4172/2332-0877.1000386.
Qiagen, June 2017. DNeasy® UltraClean® 96 Microbial Kit Handbook. In: Qiagen.
Silva, M.L., Cá, B., Osório, N.S., Rodrigues, P.N., Maceiras, A.R., Saraiva, M., 2022. Tuberculosis caused by Mycobacterium africanum: knowns and unknowns. PLoS Pathog. 18 (5), e1010490.
Stucki, D., Ballif, M., Egger, M., Furrer, H., Altpeter, E., Battegay, M., Droz, S., Bruderer, T., Coscolla, M., Borrell, S., 2016. Standard genotyping overestimates transmission of Mycobacterium tuberculosis among immigrants in a low-incidence country. J. Clin. Microbiol. 54 (7), 1862–1870.
Tulu, B., Ameni, G., 2018. Spoligotyping based genetic diversity of *Mycobacterium tuberculosis* in Ethiopia: a systematic review, Vol. 18.
UNHCR, 2022. United Nation High Commissioner for Refugees. East and Horn of Africa 2022 end year report. 2022.
Verboven, L., Phelan, J., Heupink, T.H., Van Rie, A., 2022. TBProfiler for automated calling of the association with drug resistance of variants in Mycobacterium tuberculosis. PLoS One 17 (12), e0279644.
Walker, T.M., Lalor, M.K., Broda, A., Ortega, L.S., Morgan, M., Parker, L., Churchill, S., Bennett, K., Golubchik, T., Giess, A.P., 2014. Assessment of Mycobacterium tuberculosis transmission in Oxfordshire, UK, 2007–12, with whole pathogen genome sequences: an observational study. Lancet Respir. Med. 2 (4), 285–292.
WHO, 2022. Global Tuberculosis Report-2022.
WHO-CDC, 2022. Tuberculosis Prevention and Care among Refugees and Other Populations in Humanitarian Settings: An Interagency Field Guide.