Mycobacterium tuberculosis Lineage 7 Strains Are Associated with Prolonged Patient Delay in Seeking Treatment for Pulmonary Tuberculosis in Amhara Region, Ethiopia

Solomon A. Yimer,^{a,b,c,d} Gunnstein Norheim,^b Amine Namouchi,^a Ephrem D. Zegeye,^a Wibeke Kinander,^b Tone Tønjum,^{a,e} Shiferaw Bekele,^f Turid Mannsåker,^b Gunnar Bjune,^c Abraham Aseffa,^f Carol Holm-Hansen^b

Department of Microbiology, Unit for Genome Dynamics, Oslo University Hospital, Nydalen, Oslo, Norway^a; Department of Bacteriology and Immunology, Norwegian Institute of Public Health, Division of Infectious Disease Control, Nydalen, Oslo, Norway^b; Institute of Health and Society, Section for International Health, Faculty of Medicine, University of Oslo, Blindern, Oslo, Norway^c; Amhara Regional State Health Bureau, Bahir Dar, Bahir Dar, Ethiopia^d; Department of Microbiology, Unit for Genome Dynamics, University of Oslo, Oslo, Norway^e; Armauer Hansen Research Institute, Addis Ababa, Ethiopia^f

Recent genotyping studies of *Mycobacterium tuberculosis* in Ethiopia have reported the identification of a new phylogenetically distinct *M. tuberculosis* lineage, lineage 7. We therefore investigated the genetic diversity and association of specific *M. tuberculosis* lineages with sociodemographic and clinical parameters among pulmonary TB patients in the Amhara Region, Ethiopia. DNA was isolated from *M. tuberculosis*-positive sputum specimens ($n = 240$) and analyzed by PCR and 24-locus mycobacterial interspersed repetitive unit-variable-number tandem-repeat (MIRU-VNTR) analysis and spoligotyping. Bioinformatic analysis assigned the *M. tuberculosis* genotypes to global lineages, and associations between patient characteristics and genotype were evaluated using logistic regression analysis. The study revealed a high diversity of modern and premodern *M. tuberculosis* lineages, among which approximately 25% were not previously reported. Among the *M. tuberculosis* strains ($n = 138$) assigned to seven subgroups, the largest cluster belonged to the lineage Central Asian (CAS) ($n = 60$; 26.0%), the second largest to lineage 7 ($n = 36$; 15.6%), and the third largest to the lineage Haarlem ($n = 35$; 15.2%). Four sublineages were new in the MIRU-VNTRplus database, designated NW-ETH3, NW-ETH1, NW-ETH2, and NW-ETH4, which included 24 (10.4%), 18 (7.8%), 8 (3.5%), and 5 (2.2%) isolates, respectively. Notably, patient delay in seeking treatment was significantly longer among patients infected with lineage 7 strains (Mann-Whitney test, $P < 0.008$) than in patients infected with CAS strains (adjusted odds ratio [AOR], 4.7; 95% confidence interval [CI], 1.6 to 13.5). Lineage 7 strains also grew more slowly than other *M. tuberculosis* strains. Cases of Haarlem (OR, 2.8; 95% CI, 1.2 to 6.6) and NW-ETH3 (OR, 2.8; 95% CI, 1.0 to 7.3) infection appeared in defined clusters. Intensified active case finding and contact tracing activities in the study region are needed to expedite diagnosis and treatment of TB.

Tuberculosis (TB) has been a significant cause of illness and death among humans for centuries. TB is currently responsible for more than 8.6 million new cases and 1.3 million deaths annually (1). The World Health Organization (WHO) stated in 2012 that the incidence, prevalence, and mortality from TB declined significantly in the preceding 10 years. However, progress in contracting TB is being fuelled by the emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) TB and due to a large reservoir of infectivity in TB-positive individuals who remain undetected. An improved understanding of the epidemiology, pathogenesis, and genetic diversity of the causative agent *Mycobacterium tuberculosis* is crucial for better control of the disease.

M. tuberculosis strains are presumed to have coevolved with ancient hominids and continue to evolve today. Based on bioinformatic analysis of mycobacterial genomes, *M. tuberculosis* strains cluster into seven lineages, each associated with specific global geographical locations (2). These lineages are as follows: lineage 1, Indo-Oceanic; lineage 2, East Asian, including “Beijing”; lineage 3, Central Asian (CAS)/Delhi; lineage 4, Euro-American, including the Latin American-Mediterranean (LAM), Haarlem, X type, and T families; lineage 5; lineage 6, West African 1 and West African 2, respectively; and lineage 7 (2). Lineage 2 (Beijing) is highly prevalent in East Asia, lineage 3 is widely distributed in the Indian subcontinent and some countries in East Africa, lineage 4 is commonly found in Europe, the Middle East, America, and some

parts of Africa, lineages 5 and 6 are localized in West Africa, and lineage 7 has recently been reported in Ethiopia and among Ethiopian immigrants in Djibouti (2).

Clinical and epidemiological data reveal that Beijing and Indo-Oceanic strains of lineage 2 are associated with more co-dissemination of TB and meningitis than strains from the Euro-American lineage (3). The Beijing genotype also frequently acquires drug resistance and is associated with high bacillary load in acid-fast bacillus (AFB) smears (4). The East African Indian lineage tends to be transmitted at a lower rate than other lineages (5). *Mycobacte-*

Received 19 December 2014 Accepted 29 January 2015

Accepted manuscript posted online 11 February 2015

Citation Yimer SA, Norheim G, Namouchi A, Zegeye ED, Kinander W, Tønjum T, Bekele S, Mannsåker T, Bjune G, Aseffa A, Holm-Hansen C. 2015. *Mycobacterium tuberculosis* lineage 7 strains are associated with prolonged patient delay in seeking treatment for pulmonary tuberculosis in Amhara Region, Ethiopia. J Clin Microbiol 53:1301–1309. doi:10.1128/JCM.03566-14.

Editor: G. A. Land

Address correspondence to Solomon A. Yimer, s.a.yimer@medisin.uio.no.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JCM.03566-14

The authors have paid a fee to allow immediate free access to this article.

rium africanum acquires drug resistance at a lower rate than the Euro-American lineage in Ghana (6).

While the genetic diversity of *M. tuberculosis* lineages in Ethiopia has been investigated (7–11), there is little or no data correlating sociodemographic and clinical factors with *M. tuberculosis* lineages. This study investigates the relationships between the clinical presentation of pulmonary TB (PTB) patients in Amhara Region, Ethiopia, and infection with specific *M. tuberculosis* strains. Clinical specimens and case data were collected between 2008 and 2010, and characterization by spoligotyping identified a high frequency of infection with lineage 7 *M. tuberculosis* international spoligotypes (SIT) 910 and 1729 (12). The analysis of these clinical data is extended to include description of the diverse phylogeny of *M. tuberculosis* strains infecting the patient group and correlation of clinical parameters specific to *M. tuberculosis* lineage 7 strains inducing a longer patient delay in seeking treatment.

MATERIALS AND METHODS

Design and population. Clinical and sociodemographic information and sputum specimens ($n = 240$) were collected from recently diagnosed PTB patients in a cross-sectional study. Data collection sites included selected hospitals and health centers in the study region of Amhara, Ethiopia. Patient delay in seeking treatment is defined as the period from the start of TB symptoms until the first presentation to a medical provider. The study was approved by regional ethical review committees in Norway and Ethiopia. Informed consent was obtained from all subjects.

Statistical analyses. The data were entered and analyzed using the Statistical Package for the Social Sciences (SPSS) version 20 (SPSS, Chicago, IL). Group differences were tested using the Mann-Whitney test. Univariate and multivariate logistic regression analyses were used to identify association between predictors and outcome variables. A two-sided P value of <0.05 was considered statistically significant.

Laboratory methods. (i) **Smear microscopy and sputum culture.** Sputum smear microscopy for acid-fast bacilli (AFB) was done using the Ziehl-Neelsen staining technique. Sputum culture was performed using conventional Löwenstein-Jensen (LJ) slants (13). HIV testing was performed on samples from all participants following the standard procedure in Ethiopia.

(ii) **Cultivation of *M. tuberculosis*.** Clinical *M. tuberculosis* lineage 7 strains were cultured in the semiautomated Bactec mycobacterial growth indicator tube (MGIT) 960 system (BD Biosciences), in modified Middlebrook 7H9 broth supplemented with oleic acid-albumin-dextrose-catalase (OADC) and the antibiotics polymyxin B, amphotericin B, nalidixic acid, trimethoprim, and azlocillin (PANTA). Growth of *M. tuberculosis* lineage 7 and other isolates was studied by cultivating the strains on Middlebrook 7H10 agar plates for 5 weeks, followed by measurement of colony diameter and dry weight.

DST assays. Drug susceptibility testing (DST) was performed by the proportional absolute concentration (14) and Bactec mycobacteria growth indicator tube (MGIT) (Becton Dickinson) methods (15) following the manufacturer's instructions.

DNA isolation and spoligotyping. DNA was isolated from *M. tuberculosis*-positive cultures according to the standard procedure (16). Species identification by PCR targeting the RD9 region was performed as previously described (17). The sputum samples were subjected to spoligotyping according to Kamerbeek et al. (18).

MIRU-VNTR. DNA was prepared from 240 *M. tuberculosis* cultures. The DNA extracts from nine isolates did not provide adequate PCR products and were excluded. The mycobacterial interspersed repetitive-unit-variable-number tandem-repeat (MIRU-VNTR) analysis was performed at the Norwegian Institute of Public Health, Oslo, Norway. DNA samples prepared from 231 *M. tuberculosis* strains were analyzed by the standardized 24-locus MIRU-VNTR assay as previously described (19). In brief, 24 MIRU-VNTR markers were amplified by PCR using the MIRU-VNTR

typing kit (Genoscreen, Lille, France). DNA fragments were separated by capillary electrophoresis using an automated 48-capillary ABI-3730 DNA analyzer (Applied Biosystems, Foster City, CA). PCR fragments were sized, and MIRU-VNTR alleles were assigned using the GeneScan LIZ1200 size standard and customized GeneMapper software package version 3.7 (Applied Biosystems). The results are presented as a set of digits, the MIRU code, that reflect the number of MIRU sequence repeats.

Phylogenetic analyses. The MIRU-VNTRplus online web application was used to classify the *M. tuberculosis* lineages identified in this study (<http://www.miru-vntrplus.org/MIRU/index.faces>). In brief, the MIRU-VNTR 24 loci and spoligotyping patterns were uploaded, and the “similarity” search option was selected to classify lineages based on the best match with the reference strains in the database. A categorical coefficient of 1 and distance cutoff of 0.17 that corresponds to a 4-locus difference were chosen. Thereafter a phylogenetic tree-based classification was applied for strains that were not identified by the best-match approach. A dendrogram was generated using the unweighted-pair group method using average linkages (UPGMA). To confirm the lineage classification, a minimum-spanning tree (MST) was calculated based on the MIRU-VNTR 24-locus data. A cluster was defined as two or more strains with similar genetic patterns. Strains that did not yield identical genetic patterns were referred to as “nonclustered.” The recent transmission index was calculated as the number of clustered patients minus the number of clusters divided by the total number of patients.

Operational definition of lineage 7 isolates. Lineage 7 isolates are defined by several criteria, including (i) a deletion from the 4th to 24th spoligotyping spacers, (ii) a high degree of similarity across the entire genome with other lineage 7 strains (9), and (iii) the presence of 2 or more repeats at the 24th MIRU-VNTR loci.

RESULTS

Study population. DNA samples ($n = 231$) were prepared from sputum samples from *M. tuberculosis*-positive patients. The patient group included 126 (54.5%) males and 109 (45.5%) females. The mean age was 30 years. Fifty-one (22.1%) of the study subjects were HIV positive.

Distribution of lineages. Using the “identification by similarity” search option of the MIRU-VNTRplus web application, 118 (51.0%) of the strains were assigned to lineages already represented in the database. The majority, 60 (26.0%) strains, were of the Central Asian (CAS) family, followed by 36 (15.6%) lineage 7 strains and 35 (15.2%) Haarlem strains. Approximately half (48.9%) of the isolates ($n = 113$; 48.9%) did not match any of the lineages described in the database. Phylogenetic tree analysis of these strains identified four novel clusters, which were designated NW-ETH3, NW-ETH1, NW-ETH2, and NW-ETH4. These new clusters comprised 24 (10.4%), 18 (7.8%), 8 (3.5%), and 5 (2.2%) strains, respectively. Twenty-three (10%) isolates could not be assigned to any of the available or new lineages (Fig. 1 and 2). To verify lineage classification, a minimum-spanning tree (MST) was derived (Fig. 1). All lineages detected by the UPGMA analysis were also identified as clonal complexes in the MST (Fig. 2), and thus the deduced phylogeny of these *M. tuberculosis* strains was validated.

Patient delay in seeking treatment and *M. tuberculosis* lineage association. For the complete data set, the median self-reported patient delay was 21 days (interquartile range [IQR], 7 to 60 days). However, the median duration of first seeking health care for patients infected with lineage 7 strains was significantly higher than median duration of seeking health care for all patients (Mann-Whitney test, $P < 0.008$) (Fig. 3). Patients infected with lineage 7 strains were highly likely to delay seeking care compared

FIG 1 UPGMA tree of *Mycobacterium tuberculosis* strains based on MIRU-VNTR 24-locus copy numbers. Abbreviations: CAS, Central Asian strain; LAM, Latin American-Mediterranean.

to patients infected with CAS family strains which is the predominant lineage in our data (adjusted odds ratio [AOR], 4.7; 95% confidence interval [CI], 1.6 to 13.5) (Table 1).

MIRU-VNTR patterns and clustering. A high discriminatory power was obtained by a combination of MIRU-VNTR and spoligotyping. A total of 157 distinct MIRU-VNTR genotype patterns were observed. Seventy-four (32%) isolates shared a genotype pattern with at least one other isolate and were grouped in 22 clusters comprising 2 to 11 isolates. Spoligotyping identified in 65 distinct typing patterns grouped into 26 clusters, comprising 2 to 34 isolates (12). The recent transmission index and clustering rates were 22.5% and 32%, respectively. Cluster 12 was the largest cluster, with 11 strains of NW-ETH3; the multilocus variable-number tandem-repeat analysis (MLVA) *M. tuberculosis* complex (MTC) 15-9 type for this strain was 594-15. The second largest cluster consisted of the CAS/Delhi family ($n = 7$; cluster 7, MLVA MTC

15-9 type, 1557-32). The third largest cluster was Haarlem ($n = 5$; cluster 9, MLVA MTC 15-9 type, 291-31).

***M. tuberculosis* strain clustering and drug resistance.** The odds of clustering for streptomycin-resistant strains was 2.9 times higher than that for streptomycin-sensitive strains (OR, 2.9; 95% CI, 1.2 to 7.1). Strains with any drug resistance to first-line anti-TB drugs (OR, 2.1; 95% CI, 1.1 to 4.3) exhibited a higher degree of clustering than sensitive strains. Haarlem (OR, 2.8; 95% CI, 1.2 to 6.6) and NW-ETH3 (OR, 2.8; 95% CI, 1.0 to 7.3) strains were more likely to cluster than CAS family strains (Table 2).

The overall frequency of resistance to any drug tested among the isolates was 16.9% ($n = 39$). The proportions of mono-drug-resistant, poly-drug-resistant, MDR, and MDR+ TB isolates (where poly drug resistance is resistance to more than one first-line anti-TB drug, MDR indicates resistance to rifampicin and isoniazid, and MDR+ indicates resistance to rifampicin and iso-

FIG 2 Minimum-spanning tree (MST) of *Mycobacterium tuberculosis* strains based on MIRU-VNTR 24-locus copy numbers. The *M. tuberculosis* clonal complexes are represented by different colors. Circle size is proportional to the number of MIRU-VNTR types belonging to each complex. Abbreviations: CAS, Central Asian strain; LAM, Latin American-Mediterranean.

niazid plus resistance to any other first-line anti-TB drugs) were 12.6%, 3.9%, 0.9%, and 0.9%, respectively (Table 3). Using the CAS family as a reference, NW-ETH3 (OR, 3.4; 95% CI, 1.1 to 10.2) isolates were more likely to be resistant to any first-line anti-TB drug. Haarlem strains (OR, 5.1; 95% CI, 1.3 to 20.9) were more likely to be resistant to streptomycin (Table 4). In contrast, Haarlem isolates were more resistant to any of the first-line anti-TB drugs and streptomycin than any other lineage group

(OR, 2.5, and 95% CI, 1.1 to 5.7, and OR, 3.5, and 95% CI, 1.4 to 9.3, respectively) (Table 5).

Relative growth of *M. tuberculosis* lineage 7 strains. Cultivation of the *M. tuberculosis* lineage 7 isolates on Middlebrook 7H10 agar plates showed significantly smaller colony diameter and weight among lineage 7 strains compared to those in lineages 3 and 4 combined (Fig. 4).

DISCUSSION

This study documents a very high genetic diversity of modern and premodern *M. tuberculosis* lineages among PTB patients in Amhara Region, Ethiopia. Our previous study provided preliminary evidence that a high proportion of PTB patients in this geographic region carry strains belonging to lineage 7 (12). *M. tuberculosis* lineages, including lineage 7 strains, are along with other lineages thought to have emerged approximately 70,000 years ago, near the time modern humans migrated out of Africa (2).

Criteria for identifying lineage 7 strains were previously established (9, 20–22) and were adhered to in the present study. Accordingly, all lineage 7 strains identified here ($n = 36$) have deletions from the 4th to 24th spacers and have a high degree of sequence similarity to the genomes of previously reported lineage 7 strains (9). The presence of two or more repeats at the 24th MIRU-VNTR locus is also considered highly discriminatory for modern and ancient *M. tuberculosis* lineages, as reported in genotyping studies from Bagdad (20), Singapore (21), and Bangladesh (22). All 36 strains in the present study have two alleles at the 24th MIRU-VNTR locus. Moreover, phylogenetic tree analyses support this result; both UPGMA and MST analyses clearly show that

FIG 3 Patient delay associations with infection by non-lineage 7 and lineage 7 *Mycobacterium tuberculosis* strains. $P < 0.008$.

TABLE 1 Association between sociodemographic/clinical factors and patient delay in seeking treatment^a

Variable	No. (%) of patients with treatment delay of:		OR (95% CI)		P value
	≤30 days	>30 days	Crude	Adjusted	
Sex					
Male	88 (53.7)	38 (56.7)	1	1	
Female	76 (46.3)	29 (43.3)	0.9 (0.5–1.6)	0.8 (0.4–1.8)	0.61
Age					
≤30 yr	113 (68.9)	41 (61.2)	1	1	
>30 yr	51 (31.1)	26 (38.8)	1.4 (0.7–2.5)	0.8 (0.4–1.8)	0.62
Distance to health care facility					
≤2 h	147 (89.6)	53 (79.1)	1	1	
>2 h	17 (10.4)	14 (20.9)	2.3 (1.1–4.9)*	1.8 (0.6–4.9)	0.27
Education					
Illiterate	67 (40.9)	36 (53.7)	1	1	
1st through 6th grades	48 (29.3)	19 (28.4)	0.7 (0.4–1.4)	1.2 (0.5–3.0)	0.70
7th through 12th grades	36 (21.6)	10 (13.4)	0.5 (0.2–1.1)	1.2 (0.3–4.8)	0.82
College	13 (7.9)	3 (4.5)	0.4 (0.1–1.6)	1.0 (0.2–6.4)	0.99
Employment					
Student	35 (21.5)	6 (9.0)	1	1	
Day laborer	26 (16.0)	9 (13.4)	2.0 (0.6–6.4)	1.9 (0.4–8.6)	0.39
Government employee	8 (4.9)	3 (4.5)	2.2 (0.4–10.7)	2.9 (0.4–19.9)	0.28
Private business	15 (9.2)	3 (4.5)	1.2 (0.3–5.3)	0.4 (0.0–5.3)	0.40
Housewife	13 (8.4)	4 (6.0)	1.8 (0.4–7.4)	2.0 (0.3–12.4)	0.44
Farmer	66 (40.5)	42 (62.7)	3.7 (1.4–9.6)*	4.7 (1.2–19.2)	0.03**
Symptom severity					
No work limitation	91 (56.2)	33 (49.3)	1	1	
Work limitation	49 (30.2)	26 (38.8)	1.4 (0.8–2.7)	0.8 (0.4–1.8)	0.56
Bedridden	22 (13.6)	8 (11.9)	1.0 (0.4–2.5)	0.7 (0.2–2.3)	0.52
Any drug resistance/sensitivity					
Sensitive	124 (81.0)	46 (82.1)	1	1	
Resistant	29 (19.0)	10 (17.9)	0.9 (0.4–2.1)	1.4 (0.5–3.7)	0.49
Hemoptysis					
Yes	47 (28.7)	25 (37.3)	1	1	
No	117 (71.3)	42 (62.7)	1.3 (0.5–3.4)	0.9 (0.4–1.9)	0.70
HIV serostatus					
Positive	38 (23.2)	13 (19.4)	1	1	
Negative	126 (76.8)	54 (80.6)	1.3 (0.6–2.5)	0.7 (0.2–1.9)	0.50
Lineages					
CAS	46 (28.0)	14 (20.9)	1	1	
Lineage 7	14 (8.5)	22 (32.8)	5.2 (2.1–12.7)*	4.7 (1.6–13.5)	0.00**
Haarlem	29 (17.7)	6 (9.0)	0.7 (0.2–1.9)	0.7 (0.2–2.2)	0.67
NW-ETH3	19 (11.6)	5 (7.5)	0.9 (0.3–2.7)	0.4 (0.1–1.8)	0.24
NW-ETH1	14 (8.5)	3 (4.5)	0.7 (0.2–2.8)	0.5 (0.1–2.9)	0.46
NW-ETH2	7 (4.3)	1 (1.5)	0.5 (0.1–4.1)	0.4 (0.0–3.6)	0.39
TUR	11 (6.7)	5 (7.5)	1.5 (0.4–5.0)	1.5 (0.4–5.7)	0.58
URAL	0 (0.0)	3 (4.5)			
Undefined	19 (11.6)	4 (6.0)	0.7 (0.2–2.5)	0.6 (0.1–2.2)	0.40
Beijing	1 (0.6)	0 (9.0)			
LAM	0 (0.0)	1 (0.6)			
NEW-1	1 (0.6)	0			
S	1 (0.6)	0			
NW-ETH4	2 (1.2)	3 (4.5)	4.9 (0.7–32.5)	5.9 (0.7–51.4)	0.11

^a *, crude odds ratio statistically significant at $P < 0.05$; **, AOR statistically significant at $P < 0.05$.

TABLE 2 Association between sociodemographic/clinical factors and *M. tuberculosis* strain clustering

Variable ^a	No. (%) of cases		Crude OR (95% CI)	P value ^b
	Nonclustered	Clustered		
Resistance/sensitivity				
Any drug				
Sensitive	121 (85.2)	49 (73.1)	1	
Resistant	21 (14.8)	18 (26.9)	2.1 (1.0–4.3)	0.03**
Streptomycin				
Sensitive	132 (93.0)	55 (82.1)	1	
Resistant	10 (7.0)	12 (17.9)	2.9 (1.2–7.1)	0.02**
Lineages				
CAS	42 (27.8)	18 (24.3)	1	
Lineage 7	30 (19.1)	6 (8.1)	0.5 (0.2–1.3)	0.15
Haarlem	16 (10.2)	19 (25.7)	2.8 (1.2–6.6)	0.02**
NW-ETH3	11 (7.0)	13 (17.6)	2.8 (1.0–7.3)	0.04**
NW-ETH1	13 (8.9)	2 (2.7)	0.7 (0.2–2.5)	0.60
NW-ETH2	6 (3.8)	2 (2.7)	0.8 (0.1–4.2)	0.77
TUR	14 (8.9)	2 (2.7)	0.3 (0.0–1.6)	0.17
URAL	2 (1.3)	1 (1.4)	1.2 (0.1–13.7)	0.90
Undefined	18 (11.5)	5 (6.8)	0.6 (0.2–2.0)	0.45
Beijing	1 (0.6)			
LAM	1			
NEW-1	1			
S	1			
NW-ETH4	1 (0.6)	4 (5.4)	9.3 (0.9–89)	0.05

^a The variables sex, age, distance to health care facility, education, employment status, hemoptysis, HIV serostatus, and symptom severity did not show significant differences and have not been included in the table.

^b **, significant at $P < 0.05$.

the 36 strains reside in a single branch separate from other *M. tuberculosis* lineages. These data confirm the spoligotyping results of our previous study, in which we concluded that the 36 SIT 910 and 1729 strains belong to lineage 7 (12).

Notably, we observed longer patient delay in seeking treatment associated with infection with lineage 7 strains. A possible explanation is that infection with lineage 7 strains may more often cause mild rather than severe symptoms, and that these

patients wait longer before seeking medical attention. This may result in silent TB cases, increasing the risk of transmission to uninfected individuals. Thus, more intense efforts to identify active cases are recommended in order to control the spread of TB disease. The duration of TB symptoms, time of health care seeking, and associated factors have been studied in many countries (23–29). Various sociodemographic and clinical factors have been identified as predictors of delay in seeking care.

TABLE 3 Drug resistance stratified by *Mycobacterium tuberculosis* lineage

Lineage	No. of strains with resistance ^a :									Total no. (%)
	S	H	R	MDR	MDR+	H + E	S + H	S + H + E	S + E	
CAS	3	4	1	1	0	0	0	0	0	9 (23.0)
Lineage 7	2	0	0	0	0	0	0	0	0	2 (8.9)
Haarlem	8	1	1	0	1	0	0	0	0	11 (28.2)
NW-ETH3	4	1	1	1	1	0	0	1	0	9 (23.0)
NW-ETH1	1	1	0	0	0	0	0	0	0	2 (5.1)
NW-ETH2	1	0	0	0	0	1	0	0	0	2 (8.9)
TUR	2	0	0	0	0	0	0	0	0	2 (8.9)
URAL	1	0	0	0	0	0	0	0	0	1 (2.6)
Undefined	0	0	0	0	0	0	0	0	0	0 (0.0)
Beijing	0	0	0	0	0	0	0	0	0	0 (0.0)
NEW-1	0	0	0	0	0	0	0	0	0	0 (0.0)
S	0	0	0	0	0	0	0	0	0	0 (0.0)
NW-ETH4	0	1	0	0	0	0	1	0	0	2 (5.1)
Total										39

^a S, streptomycin; H, isoniazid; R, rifampin; E, ethambutol; MDR, multidrug resistance.

TABLE 4 Association between sociodemographic/clinical factors and *M. tuberculosis* drug resistance

Variable ^a	Any drug resistance		OR (95% CI) ^b	Streptomycin resistance		OR (95% CI) ^b
	No	Yes		No	Yes	
Lineages						
CAS	44	9	1	50	3	1
Lineage 7	28	2	0.3 (0.0–1.7)	28	2	1.1 (0.2–7.6)
Haarlem	23	11	2.3 (0.8–6.4)	26	8	5.1 (1.3–20.9)**
NW-ETH3	13	9	3.4 (1.1–10.3)**	18	4	3.7 (0.8–18.1)
NW-ETH1	13	2	0.7 (0.1–3.6)	15	1	1.1 (0.1–12.3)
NW-ETH2	6	2	1.6 (0.3–9.4)	7	2	2.4 (0.2–26.1)
TUR	14	2	0.6 (0.0–5.5)	14	2	2.3 (0.4–15.7)
URAL	1	1	4.9 (0.3–85.6)	1	1	16.7 (0.8–337)
Undefined	21	0		21	0	
Beijing	1	0		1	0	
LAM	1	0		1	0	
NEW-1	1	0		1	0	
S	1	0		1	0	
NW-ETH4	6	2	1.2 (0.1–12.3)	5	0	
Clustering						
Yes	121	21	1	132	10	1
No	49	18	2.9 (1.2–7.1)**	55	12	2.9 (1.2–7.1)**

^a The variables sex, age, distance to health care facility, education, employment, hemoptysis, HIV serostatus, and symptom severity did not show significant differences and have not been included in the table.

^b **, significant at $P < 0.05$.

However, to our knowledge, an association between *M. tuberculosis* lineage and patient delay has not been reported previously. Our results suggest for the first time a possible relationship between *M. tuberculosis* lineage-specific disease and health-care-seeking behavior. The longer patent delay in seeking treatment observed in the *M. tuberculosis* lineage 7 strain might be explained by the longer time required for cultivation of these particular *M. tuberculosis* strains.

Another implication from this study is that *M. tuberculosis*

strains belonging to lineage 7 may be less virulent than strains of other *M. tuberculosis* lineages. Our study demonstrated that the average diameter of colonies and colony weight were significantly smaller for lineage 7 strains than those for strains of other lineages studied. These observations indicate that lineage 7 strains grow more slowly than non-lineage 7 *M. tuberculosis* strains *in vitro*. However, it is not known whether *M. tuberculosis* lineage 7 strains also grow relatively slowly *in vivo*, nor is it known whether slow growth of lineage 7 strains is associated with low virulence.

TABLE 5 Association between Haarlem or non-Haarlem *M. tuberculosis* strains and sociodemographic or clinical factors

Variable ^a	No. (%) of strains		Crude OR (95% CI)	P value ^b
	Non-Haarlem	Haarlem		
Resistance/sensitivity				
Any drug				
Sensitive	147 (86.5)	23 (67.6)	1	
Resistant	28 (16.0)	11 (32.4)	2.5 (1.1–5.7)	0.03**
Streptomycin resistance				
Yes	161 (92.1)	26 (76.5)	1	
No	14 (8.0)	8 (23.5)	3.5 (1.4–9.3)	0.01**
Clustering				
Yes	141 (71.9)	16 (45.7)	1	
No	55 (28.1)	19 (54.3)	3.0 (1.5–6.5)	0.00**
HIV serostatus				
Negative	158 (80.6)	22 (62.9)	1	
Positive	38 (19.4)	13 (37.1)	2.5 (1.1–5.3)	0.02**

^a The variables sex, age, distance to health care facility, education, employment, hemoptysis, and symptom severity did not show significant differences and have not been included in the table.

^b **, significant at $P < 0.05$.

FIG 4 Growth of *Mycobacterium tuberculosis* lineage 7 versus other lineages on agar plates. *M. tuberculosis* isolates belonging to lineages 3, 4, and 7 were streaked onto Middlebrook 7H10 plates and incubated at 37°C for 5 weeks, and colony morphology parameters were studied. (A) Colony diameter. The diameters of 4 to 8 representative colonies were measured, and average values were calculated for each isolate. The data presented are mean diameters $\pm$ standard deviations (SD) among the isolates. (B) Colony weight. Three representative colonies were picked up, and dry weight was measured for each isolate. The data presented are mean colony weight $\pm$ SD among the isolates (*, $P < 0.000$).

A previous study showed that Beijing strains of *M. tuberculosis* are more likely to cause severe disease and to develop multidrug resistance than other strains. It has been proposed that the virulence of these strains is due to their ability to produce a phenolic glycolipid (PGL) that reduces the efficiency of the host immune system (4). Although many *M. tuberculosis* strains do not produce PGLs, it is not yet known whether lineage 7 strains do or do not produce a similar compound. Further proteomic and lipidomic studies are warranted to address this question.

Clustering is a marker of recent transmission (30, 31). We observed a clustering rate of 32%, which is lower than that in a similar study that showed 45.1% clustering in Ethiopia (8). A recent TB transmission index in a study population can be used to assess the efficacy of the TB control program in a certain geographical area (30). The relatively lower clustering rate observed in our study may be linked to the general trend of decreasing TB incidence and prevalence in Ethiopia. The health extension pro-

gram is functioning well at the community level, and suspected TB patients are generally detected and referred for diagnosis and treatment within a few weeks of onset of symptoms. However, the clustering rate observed in our data indicates an unacceptable rate of transmission. It is desirable to reduce the rate of transmission, and this will require increased effort to actively trace patients' contacts within the suspected transmission area. Haarlem strains were more likely to cluster, and these strains were more likely to demonstrate resistance to streptomycin and any of the first-line anti-TB drugs compared to other lineages. A similar finding was previously observed in Ethiopia (8). This suggests high rates of transmission of drug-resistant *M. tuberculosis* Haarlem strains. The Amhara Region has nearly 20 million inhabitants, and there is only one TB culture facility in the region. It is crucial to expand the capabilities for TB culture and drug susceptibility testing to combat the spread of drug-resistant TB in the region. Patients carrying NW-ETH3 strains were also more likely to cluster, and these strains were more likely to be resistant to any of the first-line anti-TB drugs, as observed in another study from Ethiopia (8).

Among the modern *M. tuberculosis* lineages identified in our study, the majority, 60 (26.4%), belonged to CAS/Delhi. Another study from Ethiopia reported 38.9% CAS strains (8), and CAS/Delhi was the predominant lineage in a study from Sudan (32). CAS/Delhi strains have been circulating in Central Asia, the Middle East, Saudi Arabia (33) and India (34) for a long time. The presence of a high proportion of CAS/Delhi strains in this study may be due to the fact that CAS/Delhi predominates in this region, or the presence of these strains may be related to the increase in migration, trade, and tourism between Ethiopia and the Middle East.

The novel *M. tuberculosis* lineages identified in this study were designated NW-ETH1, NW-ETH2, NW-ETH3, and NW-ETH4. According to the SIVIT/SPOLGD4 database, all strains in these groups belong to the ill-defined T spoligotype. Strains with the name NW-ETH3 are similar to spoligotype T3-ETH (<http://www.pasteur-guadeloupe.fr/tb/spolddb4/spolddb4.pdf>). NW-ETH3 strains were more likely to cluster and be resistant to any of the first line anti-TB drugs. A previous study demonstrated a high frequency of clustered cases of infection with novel multidrug-resistant T3-ETH strains of *M. tuberculosis* in Ethiopia (35).

In summary, our study demonstrates that the TB epidemic in the study area is caused by a highly heterogeneous group of *M. tuberculosis* strains belonging to modern and premodern *M. tuberculosis* lineages. Patients with lineage 7 strains had increased duration of TB symptoms and were more likely to delay seeking medical attention. Haarlem strains frequently demonstrated resistance to one or more first-line anti-TB drugs. The considerable clustering of *M. tuberculosis* cases indicates ongoing transmission. Increased efforts to identify active cases and trace patient contacts will be essential to contain the spread of disease. Additional laboratory facilities in which TB culture and drug susceptibility testing can be performed are needed in the study region.

ACKNOWLEDGMENTS

We are grateful to Norwegian Institute of Public Health, Oslo University Hospital, and the Amhara Regional State Health Bureau for facilitating this study. We thank all of the study participants for providing valuable samples and the health care workers for facilitating sample and data collection.

This work was supported by the Research Council of Norway, GLOB-VAC Program (grants 192468/S50 and 234506/H10 to G.N. and T.T., respectively).

REFERENCES

- World Health Organization. Global tuberculosis control: surveillance, planning. World Health Organization, Geneva, Switzerland. http://www.who.int/tb/publications/global_report/en/.
- Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, Parkhill J, Malla B, Berg S, Thwaites G, Yeboah-Manu D, Bothamley G, Mei J, Wei L, Bentley S, Harris SR, Niemann S, Diel R, Aseffa A, Gao Q, Young D, Gagneux S. 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet* 45:1176–1182. <http://dx.doi.org/10.1038/ng.2744>.
- Krishnan N, Malaga V, Constant P, Caws M, Tran TH, Salmons J, Nguyen TN, Nguyen DB, Daffé M, Young DB, Robertson BD, Guilhot C, Thwaites GE. 2011. *Mycobacterium tuberculosis* lineage influences innate immune response and virulence and is associated with distinct cell envelope lipid profiles. *PLoS One* 6(9):e23870. <http://dx.doi.org/10.1371/journal.pone.0023870>.
- Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD, Warren RM. 2011. *Mycobacterium tuberculosis* Beijing genotype: a template for success. *Tuberculosis (Edinb)* 91:510–523. <http://dx.doi.org/10.1016/j.tube.2011.07.005>.
- Albanna AS, Reed MB, Kotar KV, Fallow A, McIntosh FA, Behr MA, Menzies D. 2011. Reduced transmissibility of East African Indian strains of *Mycobacterium tuberculosis*. *PLoS One* 6(9):e25075. <http://dx.doi.org/10.1371/journal.pone.0025075>.
- Yeboah-Manu D, Asante-Poku A, Bodmer T, Stucki D, Koram K, Bonsu F, Pluschke G, Gagneux S. 2011. Genotypic diversity and drug susceptibility patterns among *M. tuberculosis* complex isolates from South-Western Ghana. *PLoS One* 6(7):e21906. <http://dx.doi.org/10.1371/journal.pone.0021906>.
- Garedew L, Mihret A, Mamo G, Abebe T, Firdessa R, Bekele Y, Ameni G. 2013. Strain diversity of mycobacteria isolated from pulmonary tuberculosis patients at Debre Birhan Hospital, Ethiopia. *Int J Tuberc Lung Dis* 17:1076–1081. <http://dx.doi.org/10.5588/ijtld.12.0854>.
- Tessema B, Beer J, Merker M, Emmrich F, Sack U, Rodloff AC, Niemann S. 2013. Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia. *BMC Infect Dis* 13:131. <http://dx.doi.org/10.1186/1471-2334-13-131>.
- Firdessa R, Berg S, Hailu E, Schelling E, Gumi B, Erenso G, Gadisa E, Kirots T, Habtamu M, Hussein J, Zinsstag J, Robertson BD, Ameni G, Lohan AJ, Loftus B, Comas I, Gagneux S, Tschopp R, Yamuah L, Hewinson G, Gordon SV, Young DB, Aseffa A. 2013. Mycobacterial lineages causing pulmonary and extrapulmonary tuberculosis, Ethiopia. *Emerg Infect Dis* 19:460–463. <http://dx.doi.org/10.3201/eid1903.120256>.
- Belay M, Ameni G, Bjune G, Couvin D, Rastogi N, Abebe F. 2014. Strain diversity of *Mycobacterium tuberculosis* isolates from pulmonary tuberculosis patients in Afar pastoral region of Ethiopia. *Biomed Res Int* 2014:238532. <http://dx.doi.org/10.1155/2014/238532>.
- Mihret A, Bekele Y, Loxton AG, Jordan AM, Yamuah L, Aseffa A, Howe R, Walzl G. 2012. Diversity of *Mycobacterium tuberculosis* isolates from new pulmonary tuberculosis cases in Addis Ababa, Ethiopia. *Tuberc Res Treat* 2012:892079. <http://dx.doi.org/10.1155/2012/892079>.
- Yimer SA, Hailu E, Derese Y, Bjune GA, Holm-Hansen C. 2013. Spoligotyping of *Mycobacterium tuberculosis* isolates among pulmonary tuberculosis patients in Amhara Region, Ethiopia. *APMIS* 121(9):878–885. <http://dx.doi.org/10.1111/apm.12046>.
- Christie JD, Callinit DR. 1995. The laboratory diagnosis of mycobacterial diseases. *Clin Lab Med* 15:297–307.
- van Klingeren B, Dessens-Kroon M, van der Laan T, Kremer K, van Soolingen D. 2007. Drug susceptibility testing of *Mycobacterium tuberculosis* complex by use of a high-throughput, reproducible, absolute concentration method. *J Clin Microbiol* 45:2662–2668. <http://dx.doi.org/10.1128/JCM.00244-07>.
- Siddiqi SH, Rüsch-Gerdes S. 2006. MGIT procedure manual. Foundation for Innovative New Diagnostics, Geneva, Switzerland.
- Zwadyk P, Down JA, Myers N, Dey MS. 1994. Rendering of mycobacteria safe for molecular diagnostic studies and development of a lysis method for strand displacement amplification and PCR. *J Clin Microbiol* 32:2140–2146.
- Parsons LM, Brosch R, Cole ST, Somoskövi A, Loder A, Bretzel G, Van Soolingen D, Hale YM, Salfinger M. 2002. Rapid and simple approach for identification of *Mycobacterium tuberculosis* complex isolates by PCR-based genomic deletion analysis. *J Clin Microbiol* 40:2339–2345. <http://dx.doi.org/10.1128/JCM.40.7.2339-2345.2002>.
- Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 35:907–914.
- Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Loch C, van Soolingen D. 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* 44:4498–4510. <http://dx.doi.org/10.1128/JCM.01392-06>.
- Mustafa Ali R, Trovato A, Couvin D, Al-Thwani AN, Borroni E, Dhaer FH, Rastogi N, Cirillo DM. 2014. Molecular epidemiology and genotyping of *Mycobacterium tuberculosis* isolated in Baghdad. *Biomed Res Int* 2014:580981. <http://dx.doi.org/10.1155/2014/580981>.
- Sun YJ, Bellamy R, Lee AS, Ng ST, Ravindran S, Wong SY, Loch C, Supply P, Paton NI. 2004. Use of mycobacterial interspersed repetitive unit-variable-number tandem repeat typing to examine genetic diversity of *Mycobacterium tuberculosis* in Singapore. *J Clin Microbiol* 42:1986–1993. <http://dx.doi.org/10.1128/JCM.42.5.1986-1993.2004>.
- Banu S, Gordon SV, Palmer S, Islam R, Ahmed S, Alam KM, Cole ST, Brosch R. 2004. Genotypic analysis of *Mycobacterium tuberculosis* in Bangladesh and prevalence of the Beijing strain. *J Clin Microbiol* 42:674–682. <http://dx.doi.org/10.1128/JCM.42.2.674-682.2004>.
- Storla DG, Yimer S, Bjune GA. 2008. A systematic review of delay in the diagnosis and treatment of tuberculosis. *BMC Public Health* 8:15. <http://dx.doi.org/10.1186/1471-2458-8-15>.
- Yimer S, Holm-Hansen C, Yimaldu T, Bjune G. 2009. Health care seeking among pulmonary tuberculosis suspects and patients in rural Ethiopia: a community-based study. *BMC Public Health* 9:454. <http://dx.doi.org/10.1186/1471-2458-9-454>.
- Yimer S, Bjune G, Alene G. 2005. Diagnostic and treatment delay among pulmonary tuberculosis patients in Ethiopia: a cross sectional study. *BMC Infect Dis* 5:112. <http://dx.doi.org/10.1186/1471-2334-5-112>.
- Gele AA, Bjune G, Abebe F. 2009. Pastoralism and delay in diagnosis of TB in Ethiopia. *BMC Public Health* 9:5. <http://dx.doi.org/10.1186/1471-2458-9-5>.
- Belay M, Bjune G, Ameni G, Abebe F. 2012. Diagnostic and treatment delay among Tuberculosis patients in Afar Region, Ethiopia: a cross-sectional study. *BMC Public Health* 12:369. <http://dx.doi.org/10.1186/1471-2458-12-369>.
- Yimer SA, Bjune GA, Holm-Hansen C. 2014. Time to first consultation, diagnosis and treatment of TB among patients attending a referral hospital in Northwest, Ethiopia. *BMC Infect Dis* 14:19. <http://dx.doi.org/10.1186/1471-2334-14-19>.
- Mesfin MM, Newell JN, Walley JD, Gessesew A, Madeley RJ. 2009. Delayed consultation among pulmonary tuberculosis patients: a cross sectional study of 10 DOTS districts of Ethiopia. *BMC Public Health* 9:53. <http://dx.doi.org/10.1186/1471-2458-9-53>.
- Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J. 1994. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 330:1703–1709.
- Barnes PF, Yang Z, Preston-Martin S, Pogoda JM, Jones BE. 1997. Patterns of tuberculosis transmission in Central Los Angeles. *JAMA* 278:1159–1163.
- Sharaf Eldin GS, Fadl-Elmula I, Ali MS, Ali AB, Salih AL. 2011. Tuberculosis in Sudan: a study of *Mycobacterium tuberculosis* strain genotype and susceptibility to anti-tuberculosis drugs. *BMC Infect Dis* 11:219. <http://dx.doi.org/10.1186/1471-2334-11-219>.
- Varghese BI, Supply P, Allix-Béguec C, Shoukri M, Al-Omari R, Herbawi M, Al-Hajj S. 2013. Admixed phylogenetic distribution of drug resistant *Mycobacterium tuberculosis* in Saudi Arabia. *PLoS One* 8(2):e55598. <http://dx.doi.org/10.1371/journal.pone.0055598>.
- Vijaya-Bhanu N, van Soolingen D, van Embden JDA, Dar L, Pandey RM. 2005. Predominance of a novel *Mycobacterium tuberculosis* genotype in the Delhi region of India. *Tuberculosis* 82:105–112. <http://dx.doi.org/10.1054/tube.2002.0332>.
- Agonafir M, Lemma E, Wolde-Meskel D, Goshu S, Santhanam A, Girmachew F, Demissie D, Getahun M, Gebeyehu M, van Soolingen D. 2010. Phenotypic and genotypic analysis of multidrug-resistant tuberculosis in Ethiopia. *Int J Tuberc Lung Dis* 14:1259–1265.