Molecular epidemiology and drug resistance patterns of *Mycobacterium tuberculosis* complex isolates from university students and the local community in Eastern Ethiopia

Abiyu Mekonnen[1*] Matthias Merker[2], Jeffrey M Collins[3], Desalegn Addise[4], Abraham Aseffa[5], Beyene Petros[1#], Gobena Ameni [6] and Stefan Niemann[2,7]

[1]Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, Ethiopia

[#]Email: abule2002@yahoo.com

[2]Molecular Mycobacteriology, Research Center Borstel, Borstel, Germany

Email: mmerker@fz-borstel.de

[3]Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine, Atlanta, Georgia

Email: Jeffrey.michael.collins@emory.edu

[4]Ethiopian National Tuberculosis Reference laboratory, Ethiopian Public Health Institute, Addis Ababa, Ethiopia

Email: desalegnaddise@gmail.com

[5]Armauer Hansen Research Institute, Addis Ababa, Ethiopia

Email: aseffaa@gmail.com

[6]Aklilu Lemma Institute of Pathobiology, Addis Ababa University, Ethiopia

Email: gobenachimdi2009@yahoo.co.uk

[7]German Center for Infection Research, Partner Site Borstel, Borstel, Germany

Email: sniemann@fz-borstel.de

Short title: Molecular epidemiology and drug resistance patterns of *M. tuberculosis* in eastern Ethiopia

*Address correspondence to: Abiyu Mekonnen, abiyumg@yahoo.com; Po-box: 11422, Addis Ababa, Ethiopia

Abstract

Background: Previous studies suggest the burden of pulmonary tuberculosis (PTB) in Ethiopia may be greater in university students relative to the overall population. However, little is known about the transmission dynamics of PTB among students and members of the communities surrounding university campuses in Eastern Ethiopia.

Methods: A cross sectional study was conducted in Eastern Ethiopia among culture-confirmed PTB cases from university students (n=36) and community members diagnosed at one of four hospitals (n=152) serving the surrounding area. Drug susceptibility testing (DST) was performed on Mycobacterium Tuberculosis Complex (MTBC) isolates using BD Bactec MGIT 960 and molecular genotyping was performed using spoligotyping and 24-loci MIRU-VNTR. MTBC strains with Identical genotyping patterns were assigned to molecular clusters as surrogate marker for recent transmission and further contact tracing was initiated among clustered patients.

Results: Among all study participants, four MTBC lineages and 11 sub-lineages were identified, with Ethiopia_3 being most common sub-lineage (29.4%) and associated with strain clustering (P= 0.016). We identified 13 (8.1%) strains phylogenetically related to the known Ethiopian sub-lineages with a distinct Spoligotyping patterns and designated as Ethiopia_4. The clustering rate of MTB strains was 52.9% for university students and 66.7% for community members with a Recent Transmission Index (RTI) of 17.6% and 48.4%, respectively. Female gender, urban residence, and new TB cases were significantly associated with strain clustering ($p<0.05$). Forty-eight (30%) of the study participants were resistant to one or more first line anti TB drugs, three patients were classified as multidrug resistant (MDR), defined by isoniazid and rifampicin resistance.

Conclusion: We found evidence of significant PTB cases clustering and recent transmission among Ethiopian university students and the local community in eastern Ethiopia; with Ethiopia_3 being the predominant circulating sub-lineage. A country wide comprehensive molecular surveillance and drug resistance profiling of MTBC strains and Implementation of TB control programs within universities and the surrounding community should be considered to decrease TB transmission.

Key words: 24-loci MIRU-VNTR, Ethiopia_3 sub lineage, University students, MTBC clustering rate, Ethiopia

Background

Tuberculosis (TB) remains a major threat to public health worldwide [1], with an estimated 10.4 million cases in 2016 [2]. Ethiopia is one of the fourteen countries with the highest TB burden with an annual incidence rate of 177/100,000 population [2]. Young adults in Ethiopia have a higher TB incidence than any other age group [3], which has a negative impact on the country's socio-economic development [4]. Halting TB transmission and decreasing TB incidence rates, especially in young adults, will be an essential component of the public health program in Ethiopia. However, in communities of young adults, such as university students, little is known about recent PTB infection rates and the primary drivers of TB transmission. These communities offer an opportunity for interrupt transmission with infection control measures such as TB screening and contact tracing.

Molecular strain typing (genotyping), namely 24-loci Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeat (MIRU-VNTR) technique in combination with Spacer Oligo Nucleotide typing (Spoligotyping), is widely used to investigate local transmission dynamics and evaluate TB control programs [5,6,7]. In addition, certain MTBC lineages (e.g. lineage 2 [Beijing]) have been associated with increased pathogenicity and resistance to specific drugs, which increase the risk of drug resistant MTBC strains in a region [8].

Recent studies on university students in central Ethiopia suggested TB incidence to be much higher on school campuses relative to surrounding communities [9]. Rapid increases in university enrollment in Ethiopia has led to crowded congregate living environments on college campuses with the potential to facilitate TB transmission.

115 However, the relative contribution of recent TB transmission, both on campus and in the
116 surrounding community, to active TB disease among Ethiopian university students is
117 unknown. We studied university students and surrounding community members
118 diagnosed with pulmonary TB at three universities and four hospitals in eastern Ethiopia
119 to determine the genotypic characteristics, transmission dynamics and drug resistance
120 patterns of MTBC strains circulating in the region.

121

122

123

124

125

126

127

128

129

130

131

132

Materials and Methods

Study design and period

A cross-sectional study was conducted among students of three Eastern Ethiopian universities and members of the surrounding community diagnosed with pulmonary tuberculosis (PTB). Inclusion criteria was a positive sputum culture for MTBC. Students with PTB were identified through active case finding between May 2016 and April 2017; community TB cases were enrolled from hospitals serving the geographic areas surrounding the universities from January to April 2017. Participants for whom a valid genotype could not be obtained were excluded from the study. This included evidence of a mixed infection or laboratory cross-contamination as indicated by double alleles at two or more loci during MIRU/VNTR typing and two or more loci with missing data following at least two independent PCR amplifications. Reasons for missing loci included insufficient DNA concentration [10] and nucleotide polymorphisms in the sequence complementary to the PCR primers [11].

Study sites

The study comprises three regional states and one administration (Oromia, Somali and Harari regional states and Dire Dawa City administration). The towns where the study hospitals were located are: Harar (Harari region), Haramaya (Haramaya district, Oromia region), Dire Dawa (Dire Dawa city administration) and Jigjiga (Somali region). According to the Central Statistical agency of Ethiopian population projection values of 2017, the population of Harari and Somali regional states were 246,000 and 5,748,998, respectively; whereas that of Dire Dawa city administration and Haramaya district were

155 466,000 and 361,787, respectively [12]. The universities studied were Haramaya
156 University, Dire Dawa University and Jigjiga University.

157 Collection of MTBC strains

158 Students: All full-time students attending Haramaya University, Dire Dawa
159 University and Jigjiga University were screened for PTB by active case finding through
160 dormitory-to-dormitory visits using WHO TB screening document [13] between May 2016
161 and April 2017. Two spot sputum samples were collected from students with a positive
162 WHO symptom screen. One sputum sample was processed for Acid Fast Bacilli (AFB)
163 sputum smear microscopy and the other one was transported to Harari Health Research
164 and Regional Laboratory, Harar, Ethiopia for MTB culture. Thirty-six student AFB cultures
165 were positive for MTBC and genotype analysis was performed on all isolates.

166 Local community: Persons diagnosed with PTB from the community
167 surrounding the universities studied were enrolled at four hospitals: Haramaya district
168 hospital (Haramaya), Hiwot Fana specialized university hospital (Harar), Dil Chora
169 hospital (Dire Dawa) and Karamara hospital (Jigjiga). Persons presenting to these
170 facilities with symptoms of PTB and found to have a positive AFB sputum smear between
171 January and April 2017 were approached for study enrollment. Persons giving consent
172 for study participation were administered a standard questionnaire to collect information
173 about relevant clinical and sociodemographic data. An early morning sputum sample was
174 collected from the respective smear positive patient and stored at -20^{0}c until transported
175 to Harari Health Research and Regional Laboratory for TB culture. Initially 171 sputum
176 samples were collected from smear positive PTB cases from the four hospitals, of which
177 19 (11.3%) samples were excluded due to growth of contaminating flora.

Laboratory methods

Sputum samples were collected from students with a positive symptom screen and community members with symptoms and a positive AFB sputum. All specimens were cultivated on LJ (BBL™ Lowenstein-Jensen) media at Harari Health Research and Regional laboratory following the standard operational procedures. Isolates on LJ transported to the National TB reference laboratory, Addis Ababa, were reactivated and phenotypic Drug Susceptibility Testing (DST) was performed using the MGIT SIRE kit at a critical concentration of streptomycin (STM) 1 µg, Isoniazid (INH) 0.1 µg, Rifampicin (RIF) 1 µg and Ethambutol (EMB) 5 µg on liquid Mycobacterium Growth Indicator Tube system (MGIT) 960 as described by the manufacturer [14].

DNA from MTBC isolates were extracted and transported to Borstel Molecular Mycobacteriology laboratory, Germany for genotype analysis. Molecular characterization of all isolates was conducted using Spoligotyping [15] and 24- loci MIRU-VNTR customized kits (Genoscreen, Lilli, France) [16]. Spoligotypes common to more than one strain were designated as shared types (ST) and were assigned a shared international type number (SIT) according to the international spoligotype database SpolDB4 [17]. Basic strain classification and MLVA MTBC 15–9 nomenclature assignment was done using the MIRUVNTRplus database [18]. Samples with complete spoligotyping and MIRU/VNTR-24 results were used for clustering analysis. A cluster was defined as two or more MTBC isolates sharing identical 24-loci MIRU/VNTR and spoligotyping patterns. Samples with no PCR amplicon at only one locus in the 24-loci MIRU/VNTR analysis were included for further analysis by considering missing data at the respective locus [19].

Data analysis

MTBC genotypes were classified in a phylogenetic tree (based on 24-loci MIRU-VNTR profiles) in relation to a MTBC reference collection hosted on the MIRU-VNTRplus website (available at www.miru-vntrplus.org) and considering genotype specific Spoligotyping patterns [18]. Minimum Spanning Trees (MST) were calculated with BioNumerics (Version 7.5; Applied Maths, Sint-Martens-Latem, Belgium) as recommended by the manufacturer (available at http://applied-maths.com). A dendrogram was generated using the Unweighted Pair Group Method with Arithmetic averages (UPGMA) based on the 24-loci MIRU-VNTR profiles. The UPGMA tree was further processed using EvolView, an online visualization and management tool for customized and annotated phylogenetic trees [20] (Fig 1). The Recent Transmission Index (RTI) was calculated as number of clustered patients minus number of clusters divided by total number of patients.

Data were entered and analyzed using IBM SPSS version 23 statistical package software. Logistic regression was used to estimate the strength of association between strain clustering and different variables. A Chi-square test was used for bivariate analysis of categorical variables. P-values <0.05 were considered as statistically significant. Those factors significantly associated with clustering in the univariate analysis were included in the multivariate regression model.

Ethical Considerations

The study was ethically reviewed and approved by Addis Ababa University, College of Natural Sciences Research Ethics Review Board. Written informed consent

223 that included information about the risks and benefits of the study was a prerequisite for
224 all university students and local community study participants.

225

226 Results

227 Study population

228 We enrolled 188 study participants with PTB, all of whom were sputum culture
229 positive for MTBC. Molecular genotyping was performed on all isolates using
230 Spoligotyping and MIRU/VNTR-24 loci techniques. Of these, 28 isolates were excluded
231 for technical reasons; two were found to be a mixed infection as indicated by double
232 alleles at two or more loci during MIRU/VNTR typing and 26 isolates had two or more loci
233 missing following PCR amplification. Therefore, 160 MTBC strains were included in the
234 final genotypic analysis: 34 from university students and 126 from the local community.
235 The majority of the study participants were male (80%) and rural residents (60.6%). Forty-
236 six (28.7%) of the participants had previous treatment for TB, and 10.6% were HIV-
237 positive (Table 1).

238 Table 1. Socio-demographic and other characteristics by study participants, eastern
239 Ethiopia, May 2016 to April 2017

Variable		Students n (%)	Local Community n (%)	Total n (%)	P-value
Gender	M	33 (97.1)	95 (75.4)	128 (80.0)	0.005
	F	1 (2.9)	31 (24.6)	32 (20.0)	

Age group	<18	0 (0)	16 (12.7)	16 (10.0)	NA
	18-24	33 (97.1)	27 (21.4)	60 (37.5)	
	25-34	1 (2.9)	40 (31.7)	41 (25.6)	
	35-44	0 (0)	23 (18.3)	23 (14.4)	
	≥ 45	0(0)	20 (15.9)	20 (12.5)	
Location/Region	Haramaya/Harar	18 (53.0)	46 (36.5)	64 (40.0)	0.177
	Dire Dawa	6 (17.6)	38 (30.2)	44 (27.5)	
	Jigjiga	10 (29.4)	42 (33.3)	52 (32.5)	
Residence	Urban	3 (8.8)	60 (47.6)	63 (39.4)	NA
	Rural	31 (91.2)	66 (52.4)	97 (60.6)	
Previous Rx for TB	No	28 (82.4)	86 (68.3)	114 (71.3)	0.107
	Yes	6 (17.6)	40 (31.7)	46 (28.7)	
HIV status	Negative	32 (94.1)	111 (88.1)	143 (89.4)	0.614
	Positive	2 (5.9)	15 (11.9)	17 (10.6)	
Clustering	No	16 (47.1)	42 (33.3)	58 (36.3)	0.140
	Yes	18 (52.9)	84 (66.7)	102 (63.8)	

Phylogenetic analysis of *M. tuberculosis* strains

The two predominant MTBC lineages in this study were lineage 4 (Euro-American, 71.3%) and lineage 3 (Delhi-CAS, 22.5%). MTBC strains classified as lineage 1 (East African Indian), and lineage 2 (Beijing) were only identified in six and four patients, respectively (Table 2).

246 Table 2. Phylogenetic Sub-lineages and Drug Sensitivity patterns of MTBC strains
247 isolated by study participants, eastern Ethiopia, May 2016 to April 2017
248

Variable		Students n (%)*	Local Community n (%)*	P-Value
Phylogenetic Lineage	Ethiopia_3	10 (29.4)	38 (30.2)	0.074
	Ethiopia_4	1 (2.9)	12 (9.5)	
	Ethiopia_H37 Rv like	7 (20.6)	15 (11.9)	
	Delhi/CAS	7 (20.6)	29 (23.0)	
	Beijing	0 (0)	4 (3.2)	
	LAM	1 (2.9)	6 (4.8)	
	Haarlem	1 (2.9)	7 (5.6)	
	X-type	0 (0)	4 (3.2)	
	EAI	0 (0)	6 (4.8)	
	UgandaII	1 (2.9)	0 (0)	
	URAL	1 (2.9)	1 (0.8)	
	Not defined	5 (14.7)	4 (3.2)	
Streptomycin	S	27 (79.4)	109 (86.5)	0.304
	R	7 (20.6)	17 (13.5)	
Isoniazid	S	28 (82.4)	111 (88.1)	0.379
	R	6 (17.6)	15 (11.9)	
Rifampin	S	34 (100)	117 (92.9)	0.109
	R	0 (0)	9 (7.1)	

Ethambutol	S	34 (100)	117 (92.9)	0.109
	R	0 (0)	9(7.1)	
Resistance to ≥1 Anti-TB drugs	S	23 (67.6)	89 (70.6)	0.734
	R	11 (32.4)	37 (29.4)	
MDR	S	34 (100)	123 (97.6)	0.364
	R	0 (0)	3 (2.4)	

*Percentage is calculated from column total

MDR: Multidrug resistant, EAI: East-African Indian; LAM: Latin American Mediterranean; S: Susceptible; R: Resistance

The most prevalent MTBC genotypes within the Euro-American super-lineage were Ethiopia_3 (29.4%) and Ethiopia-H37Rv-like (12.8%), both previously described among Ethiopian patients with PTB [10], and TB lymphadenitis [21]. Based on the phylogenetic structure/topology (UPGMA- and MST-based), we further termed the third largest monophyletic lineage 4 group “Ethiopia_4” (8.1%) accordingly (Fig 1). Ethiopia_4 strains are closely related to Ethiopia_3 strains but with a distinct Spoligotyping pattern. Both, Ethiopia_3 and Ethiopia_4 strains, have a shared common ancestor with TUR-genotype strains, but with unique Spoligotyping patterns, justifying their own nomenclature in the context of the molecular epidemiology in Ethiopia.

Fig 1. Minimum spanning tree based on MIRU/VNTR profiles of 24- loci of 160 MTBC isolates in eastern Ethiopia. A circle representing a specific genotype is divided in to the number of strains clustering in it.

MTBC strains classified as Beijing, East Africa India (EAI), or X-type were isolated exclusively from community members and not identified among university students. Particularly Beijing strains were only isolated from members of the surrounding Dire Dawa community. Nine MTBC isolates (5.6%) could not be classified within any of the known genotypes but these strains were part of lineage 4 with an unknown sub-group and are labeled “Not defined”.

Molecular MTBC clusters and associated risk factors

The overall cluster rate of MTBC strains derived from all patients was 63.8%, including 21 clusters with 2 to 22 patients. Twelve clusters contained at least one university student. The overall Recent Transmission Index (RTI) was 50.6% (Table 3). In a univariate regression model, we found the following demographic and treatment related factors associated to clustered cases: female gender (P=0.004), urban residence (P=0.012) and those without prior TB diagnosis of active TB disease (P=0.001). Additionally, MTBC strains classified as Ethiopia_3 were significantly more abundant among clustered cases compared to other MTBC genotypes/lineages (Table 4, Fig 2).

Table 3. Clustering rate and Recent Transmission Index (RTI) analysis using Spoligotyping and MIRU-VNTR 24-loci methods for local community, university students and all study participants, eastern Ethiopia, May 2016 to April 2017

Methods	Source of MTBC isolate	No. of different patterns	Unique Patterns	Number of clusters	Number of isolates in cluster	Clustering Rate	RTI

Spoligotyping	Students	18	6	12	28	82.4%	26.5%
	Local community	38	12	26	114	90.5%	69.8%
	Both students and local community	40	18	22	142	88.7%	75%
MIRU-VNTR 24-loci	Students	28	16	12	18	52.9%	17.6%
	Local community	63	38	25	88	69.8%	50%
	Both students and local community	77	55	22	105	65.6%	51.9%
Spoligotyping + MIRU-VNTR 24-loci	Students	28	16	12	18	52.9%	17.6%
	Local community	65	44	21	84	66.7%	48.4%
	Both students and local community	79	58	21	102	63.8%	50.6%

Fig 2. Radial UPGMA tree based on MIRU/VNTR 24-loci copy number.

Table 4. Factors associated with MTBC strain clustering, eastern Ethiopia, May 2016 to April 2017

Variable		Genotype		AOR (95% C.I)	P-value
		Clustered n (%)*	Unique n (%)*		
Gender	Male	75 (58.6)	53 (41.4)	1	

	Female	27 (84.4)	5 (15.6)	7.14 (1.85, 27.52)	0.004
Residence location	Urban	48 (76.2)	15 (23.8)	4.19 (1.36, 12.89)	0.012
	Rural	54 (55.7)	43 (44.3)	1	
Previous DX for TB	No	83 (72.8)	31 (27.2)	6.50 (2.14, 19.79)	0.001
	Yes	19 (41.3)	27 (58.7)	1	
Isoniazid	Susceptible	84 (60.4)	55 (39.6)	1	
	Resistance	18 (85.7)	3 (14.3)	7.81 (0.75, 81.69)	0.086
Any resistance to FLDs	No	62 (55.4)	50 (44.6)	1	
	Yes	40 (83.3)	8 (16.6)	3.23 (0.80, 13.01)	0.100
Genotype	Ethiopia_3	41 (87.2)	7 (12.8)	13.37 (1.63,109.6)	0.016
	Ethiopia_4	7	6	1.11 (0.12, 10.65)	0.931
	Ethiopia_H37 Rv like	14	8	1.39 (0.17, 11.49)	0.760
	Delhi/CAS	25	11	3.49 (0.48, 25.66)	0.219
	Beijing	4	0	NA	NA
	LAM	2	5	0.10 (0.01, 1.48)	0.094
	Haarlem	0	8	NA	NA
	X-type	0	4	NA	NA
	EAI	4	2	3.26 (0.19, 55.52)	0.414
	UgandaII	0	1	NA	NA
	URAL	0	2	NA	NA
	Not defined	5	4	1	

*Percentage is calculated from row total

There was no a statistically significant difference in the proportion of clustered strains between university students and local community (p=0.142) (Table 1). With regard to the university cohort, 18/34 (52.9%) patients were part of a molecular cluster. 17 (94%) lived in an area at least 400 km away prior to attending university. There were two clusters with multiple student cases. One cluster contained three students, all of whom were attending Haramaya University. Two of the three students shared a common area of study, but none were living in the same dormitory or building. The other cluster contained five student cases, with three of the students attending Haramaya University and the other two attending Jigjiga and Dire Dawa Universities, respectively. None of the students in this cluster shared a clear epidemiologic link. Of the eight students clustered with other student cases, six (75%) reported recent exposure to someone with a cough, but none had a known exposure to an active TB case on campus. The RTI among students was 17.6% (Table 3).

Patterns of *M. tuberculosis* drug resistance

Forty-eight (30%) of the study participants were resistant to one or more first line anti TB drugs. Among the first line anti-TB drugs examined, resistance to streptomycin and isoniazid was most common at 15%, and 13.1%, respectively. Rates of resistance to at least one first line anti-TB drug were similar in students and the community (32.4% vs 29.4%, p=0.734) (Table 2). Few participants had multidrug resistant (MDR) TB, with three cases occurring in community members (infected with Ethiopia_3, Ethiopia H37 Rv like, and LAM strain, respectively) and none occurring among university students.

315 Ethiopia_3 strains were observed with a higher proportion (9.4%) of resistance to at least
316 one first-line drug compared to other MTBC genotypes (Table 5).
317
318 Table 5. Patterns of drug resistance to first line anti-TB drugs by MTBC lineage eastern
319 Ethiopia, May 2016 to April 2017

Lineage (N=160)	STM n (%)*	INH n (%)*	RIF n (%)*	EMB n (%)*	Any resistance n (%)*	MDR n (%)*
Ethiopia_3 (48)	8 (5.0)	8 (5.0)	3 (1.9)	1 (0.6)	15 (9.4)	1 (0.6)
Ethiopia_4 (13)	3 (1.9)	1 (0.6)	1 (0.6)	4 (2.5)	6 (3.8)	0
Ethiopia_H37 Rv like (22)	2 (1.3)	5 (3.1)	1 (0.6)	2 (1.3)	8 (5.0)	1 (0.6)
Delhi/CAS (36)	5 (3.1)	2 (1.3)	1 (0.6)	1 (0.6)	9 (5.6)	0
Beijing (4)	2	1	1	0	3	0
LAM (7)	1 (0.6)	2 (1.3)	1 (0.6)	1 (0.6)	3 (1.9)	1 (0.6)
Haarlem (8)	0	0	0	0	0	0
X-type (4)	1	1	1	0	2	0
EAI (6)	0	0	0	0	0	0
UgandaII (1)	0	0	0	0	0	0
URAL (2)	0	0	0	0	0	0
Not defined (9)	2 (1.3)	1 (0.6)	0	0	2 (1.3)	0
Total	24 (15.0)	21 (13.1)	9 (5.6)	9 (5.6)	48 (30.0)	3 (1.9)

320 *Percentage calculated from the total isolates (N) and for those lineages with more than
321 5 strains; MDR: Multi Drug Resistance; n= number of resistance strains in the specific
322 group.
323
324
325
326
327
328
329
330
331
332
333
334
335
336

Discussion

By using molecular MTBC strain typing (24-loci MIRU-VNTR typing and Spoligotyping) in combination with epidemiological investigations we point out the importance of MTBC Ethiopia_3 strains in the context of recent transmission and drug resistance in Eastern Ethiopia. Cluster rates among students and other community members were similar (52.9% and 66.7%, respectively [P =0.142]) and all clusters with university cases had at least one case from the community cohort associated. Being aware of the limitations mediated by the small sample size and the cross-sectional sampling scheme, our data suggest that MTBC transmission is still a major public health concern in Eastern Ethiopia and is not limited to congregate living facilities, e.g. universities or prisons [22].

Most university students with active TB originated from areas at least 400 km from the university campuses which might well be the source of the MTCB infection of non-clustered cases. Given the high cluster rate and the links to TB cases in the local community also suggests a significant proportion of active TB disease among students is due to transmission either on campus or in surrounding areas. Students within a molecular cluster were not observed with a direct epidemiologic link, indicating rather random short-term exposures as source of infection. These findings indicate that TB-transmission is not within university campuses but is rather a community associated phenomenon. Larger scale molecular surveillance of MTBC strains in the region may be required to comprehensively characterize community transmission dynamics and design the most impactful interventions to interrupt transmission.

Molecular clusters defined by the applied genotyping methods are a surrogate marker for recent transmission and epidemiological linked cases [23]. The high clustering rate of 63.8% in this study indeed indicates that the majority of active TB cases in Eastern Ethiopia are due to recent transmission and not reactivation. Studies from other parts of the world, including South Tawara, Kiribat [24], have shown an even higher clustering rate of 75.3%. Although there were no previous studies conducted in eastern Ethiopia using the standard combined application of 24 loci MIRU/VNTR and Spoligotyping methods, a study from Northwestern Ethiopia demonstrated a clustering rate of 45.1% [10], differences might be explained by different living environments in the areas studied (e.g. rural vs. metropolitan setting) or other demographic differences (e.g. age, gender) [22].

Here, multivariable analysis demonstrated several demographic and clinical factors were associated with clustering. Females were about 7 times more likely than males to be a part of a cluster. This finding is similar to previous studies from Ethiopia [25] and Botswana [26] and could be linked to an increased tendency of females spending more time in close contact with their relatives in social sittings like market places compared to males. Urban residents were more than four times more likely part of a cluster, compared to their rural counterparts, which could be a result of the dense living conditions in cities. However, the burden of tuberculosis was higher among rural residents, which may reflect higher rates of poverty and less access to health care among the rural population.

Univariate analysis showed that strains with resistance to at least one FLD were more likely to be part of a cluster and were strongly associated with the Ethiopia_3

382 genotype. This is consistent with previous findings that even mono-resistant isolates may
383 demonstrate selective pressure for MTBC transmission in Ethiopia [10, 27]. Similar to
384 previous studies in Eastern Ethiopia [29], the prevalence of MDR TB was low. As recent
385 data from South Africa suggests person-to-person transmission of drug resistant MTBC
386 strains is the primary mechanism for the propagation of drug resistance, it will be
387 important to improve public health programs to minimize transmission of these isolates
388 [28].

389 The predominant MTBC lineages in this study were lineage 4 followed by lineage-
390 3; findings that are similar to other studies conducted in Ethiopia [10, 19, 22]. We also
391 demonstrate that local sub-linages such as the newly described “Ethiopia_4” and the
392 closely related to Ethiopia_3 (both part of lineage 4 and related to TUR genotype strains)
393 dominate among Ethiopian TB cases but do not play a major role in the global TB
394 epidemic. This might be another example of a specialized, locally adapted MTBC strain
395 type as recently suggested by Stucki and colleagues for the lineage 4 strains [30].
396 However, it is also important to note that strains with the Ethiopia H37Rv-like genotype
397 can be found in many other world regions [24, 31, 32], and shares a common ancestor
398 with the H37Rv laboratory reference strains (MTBC lineage 4.7, 4.8, and 4.9 according
399 to Coll and his colleagues [33].

400 Strains from the Ethiopia_3 sub-lineage were more likely to be part of a cluster,
401 indicating active transmission in the study area; an observation that was also found in
402 Northwestern Ethiopia [10] and another study from Eastern Ethiopia [21]. The association
403 between the Ethiopia_3 genotype and resistance to at least one first line drug in our study
404 might be one factor that contribute to the expansion of this strain type. This finding

405 highlights the need to conduct a large scale and more detailed characterization of the
406 Ethiopia_3 sub-lineage in Ethiopia. Studies conducted in Northern, Northwestern and
407 Southwestern Ethiopia have found the Dehli/CAS to be the predominant sub-lineage [10,
408 21, 27, 34], which may be attributable to these regions bordering Sudan, where Dehli/CAS
409 was the most prevalent MTBC genotype [35]. All active TB cases caused by the Beijing
410 strain, which is associated with high virulence, multidrug resistance and increased
411 mortality [36], originated from the Dire Dawa community, which also supports the need to
412 monitor the disease in the region. This study did not reveal any lineage-7 isolates (also
413 referred to as "Ethiopia_1" in previous studies [10. 21]), which is in agreement with
414 previous work demonstrating the predominance of lineage-7 in the northern part of the
415 country [37].

416 This study is subject to several limitations. Our study in the local community
417 included only smear positive PTB cases who have access to and visited the study hospital
418 laboratories, which may not be representative of all active TB cases in the region. While
419 our study is the first to provide insight about how TB is transmitted among persons
420 attending university in eastern Ethiopia, the number of isolates from university students
421 available for genotyping was small, limiting the ability to draw definitive conclusions about
422 TB transmission.

423

424

425

426

Conclusion

This study suggests TB-transmission in Ethiopia still is a major public health concern, and that transmission among college students is not limited to the university settings but also occurs in the general community. While the burden of MDR-TB in eastern Ethiopia remains low, we present evidence that the sub-lineage Ethiopia_3 is linked to recent TB transmission and also associated with resistance to at least one first-line drug. Country-wide comprehensive molecular surveillance and DST profiling of MTBC strains may be useful to guide ongoing and future TB control programs.

Supporting Information

S1 Fig. Neighbor joining (NJ) phylogenetic tree based on 24 loci MIRU-VNTR profiles of 160 MTBC isolates from eastern Ethiopia (PDF) in relation to the MTBC reference collection hosted on miru-vntrplus.org.

Acknowledgements

We would like to express our appreciation to the administrations of the respective universities and hospitals where the study was conducted. We thank all study participants and data collectors of the study. We are grateful to Harari Health Research and Regional Laboratory and Ethiopian Public Health Institute for their lab facility support. We acknowledge Anja Lüdemann, Tanja Struve-Sonnenschein, and Tanja Ubben for their support in MIRU-VNTR typing and Spoligotyping. We appreciate all-rounded support from department of Microbial, Cellular and Molecular Biology, Addis Ababa University.

Funding: This work was supported in part by the NIH/Fogarty International Center Global Infectious Diseases Grant D43TW009127 and NIH grants T32 AI074492

449 and UL1 TR002378 (Georgia Clinical and Translational Science Alliance). Also, the
450 United States Agency for International Development, through USAID Challenge TB,
451 financially supports this study under the terms of Agreement No. AID-OAA-A-14-00029,
452 by the 3rd round grant of TB Research Advisory Committe (TRAC) and KNCV
453 Tuberculosis Foundation, Ethiopia. The funders had no role in study design, data
454 collection and analysis, decision to publish, or preparation of the manuscript.

455 **Authors contribution:** AM conceived and designed the study. AM, BP,
456 JC, GA and AA reviewed the proposal, contributed to the experiment, analysis and
457 interpretation of the data, as well as the composition of this manuscript. AM, MM and SN
458 participated in cluster analysis, strain classification, interpretation of the results and
459 reviewed the initial and final manuscript. AM and DA participated in Drug Sensitivity
460 Testing, interpretation of results and reviewed the manuscript. All authors read and
461 approved the final manuscript.

462

463

464

465

466

467

468

References

1. Palomino J. Molecular detection, identification and drug resistance detection in Mycobacterium tuberculosis. FEMS Immunol Med Microbiol. 2009;56(2):103-11.

2. Global tuberculosis report 2017. Geneva: World Health Organization; 2017.

3. World Health Organization. Ethiopia: on the road to ending tuberculosis. http:// www.who.int/features/2016/ethiopia-tuberculosis/en/

4. Reves R and Angelo S. As Ethiopia moves toward tuberculosis elimination, success requires higher investment. Washington: Center for Strategic & International Studies; 2016. Available at https://csis-prod.s3.amazonaws.com/ s3fspublic/legacy_files/files/publication/160229_Reves_ EthiopiaMovesTowardTB_Web.pdf

5. Firdessa R, Berg S, Hailu E, Schelling E, Gumi B, Erenso G, et al. Mycobacterial lineages causing pulmonary and extrapulmonary tuberculosis, Ethiopia. Emerg Infect Dis. 2013; 19: 460±3. https://doi.org/ 10.3201/eid1903.120256 PMID: 23622814

6. Niemann S and Supply P. Diversity and evolution of Mycobacterium tuberculosis: moving to whole-genome-based approaches. Cold Spring Harb Perspect Med. 2014;4:a021188. doi:10.1101/cshperspect.a021188.

7. Fauville-Dufaux M and Supply P. Three-year population-based evaluation of standardize mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of Mycobacterium tuberculosis. J Clin Microbiol. 2008; 46(4):1398–406. doi: 10.1128/JCM.02089-07 PMID: 18234864

491 8. Thwaites G, Caws M, Chau T, D'Sa A, Lan N, Huyen M, et al. Relationship between
492 Mycobacterium tuberculosis genotype and the clinical phenotype of pulmonary
493 and meningeal tuberculosis. J Clin Microbiol. 2008; 46: 1363±8.
494 https://doi.org/10.1128/JCM.02180-07 PMID: 18287322
495 9. Mekonnen A and Petros B. Burden of tuberculosis among students in two Ethiopian
496 universities. EMJ. 2016;54 (4):189–196.
497 10. Tessema B, Beer J, Merker M, Emmrich F, Sack U, Rodloff AC, et al. Molecular
498 epidemiology and transmission dynamics of Mycobacterium tuberculosis in
499 Northwest Ethiopia: new phylogenetic lineages found in Northwest Ethiopia. BMC
500 Infect Dis. 2013;13:131
501 11. Affolabi D, Anyo G, Faïhun F, Sanoussi N, Shamputa I, et al. First molecular
502 epidemiological study of Mycobacterium tuberculosis in Benin. Int J Tuberc Lung
503 Dis. 2009; 13:317–322.
504 12. Central Statistical Agency (Ethiopia). Population Projection of Ethiopia for All
505 Regions at Wereda Level from 2014–2017. Federal Democratic Republic of
506 Ethiopia Central Statistical Agency, Addis Ababa, Ethiopia. 2013.
507 http://www.csa.gov.et/images/general/news/pop_pro_wer_2014–2017
508 13. Maher D, Grzemska M, Coninx R, and Reyes H. Guidelines for the control of
509 tuberculosis in prisons. Geneva: World Health Organization.1998.
510 http://whqlibdoc.who.int/hq/1998/WHO_TB_98.250.pdf
511 14. Salm H Saddiqqi and Sabine Rüsch-Gerdes Mycobacteria Growth Indicator Tube
512 (MGIT) Culture and Drug Susceptibility Demonstration Projects, 2006 MGIT™

513 procedure manual for Bactech™ and MGIT 960™ TB system available at
514 http://www.finddx.org/wp- content/uploads/2016/02/mgit_manual_nov2006.pdf.
515 15. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al.
516 Simultaneous detection and strain differentiation of Mycobacterium tuberculosis
517 for diagnosis and epidemiology. J Clin Microbiol. 1997;35:907–14.
518 16. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, et al. Proposal
519 for standardization of optimized mycobacterial interspersed repetitive unit-
520 variable-number tandem repeat typing of Mycobacterium tuberculosis. J Clin
521 Microbiol. 2006;44:498–510.
522 17. Dou H, Tseng F, Lin C, Chang J, Sun J, Tsai W, et al. Molecular epidemiology and
523 evolutionary genetics of Mycobacterium tuberculosis in Taipei. BMC Infect Dis.
524 2008;8:170. doi:10.1186/1471-2334-8-170.
525 19. Weniger T, Krawczyk J, Supply P, Niemann S and Harmsen D. MIRU-VNTRplus: a
526 web tool for polyphasic genotyping of Mycobacterium tuberculosis complex
527 bacteria. Nucleic Acids Res. 2010;38:W326–31. doi:10.1093/nar/gkq351.
528 19. Barletta F, Otero L, Jong B, Iwamoto T, Arikawa K, Van der Stuyft P, et al.
529 Predominant Mycobacterium tuberculosis families and high rates of recent
530 transmission among new cases are not associated with primary multidrug
531 resistance in Lima. J Clin Microbiol. 2015;53:1854–63.
532 20. He Z, Zhang H, Gao S, Lercher M, Chen W and Hu S. Evolview v2: an online
533 visualization and management tool for customized and annotated phylogenetic
534 trees. Nucleic Acids Research. 2016; 44 doi: 10.1093/nar/gkw370

21. Biadglegne F, Merker M, Sack U, Rodloff A, and Niemann S. 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. 2015;10(9).

22. Ali S, Beckert P, Haileamlak A, Wieser A, Pritsch M, Heinrich N, et al. Drug resistance and population structure of M. tuberculosis isolates from prisons and communities in Ethiopia. BMC Infect Dis. 2016;16(1):687.

23. Chevrel-dellagi D, Abderrahman A, Haltiti R, Koubaji H, Gicquel B and Dellagi K. Large-scale DNA fingerprinting of mycobacterium tuberculosis strains as a tool for epidemiological studies of tuberculosis. J Clin Microbiol. 1993;31:2446–50.

24. Aleksic E, Merker M, Cox H, Reiher B, Sekawi Z, Hearps A, et al. First Molecular Epidemiology Study of Mycobacterium tuberculosis in Kiribati. PLoSONE. 2013;8:e55423.

25. Bruchfeld J, Aderaye G, Palme IB, Bjorvatn B, and Ghebremichael S. Molecular epidemiology and drug resistance of Mycobacterium tuberculosis isolates from Ethiopian pulmonary tuberculosis patients with and without human immunodeficiency virus infection. J Clin Microbiol. 2002;40:1636–1640.

26. Lockman S, Sheppard J, Braden C, Mwasekaga M, Woodley C, Kenyon T, et al. Molecular and conventional epidemiology of *Mycobacterium tuberculosis* in Botswana: a population-based prospective study of 301 pulmonary tuberculosis patients. J. Clin. Microbiol. 2001;39:1042–1047.

557 27. Yimer S, Norheim G, Namouchi A, Zegeye E, Kinander W, Tønjum T, et al.
558 Mycobacterium tuberculosis lineage 7 strains are associated with prolonged
559 patient delay in seeking treatment for pulmonary tuberculosis in Amhara Region,
560 Ethiopia. J Clin Microbiol. 2015;53:1301–9.
561 28. Shah N, Auld S, Brust J, Mathema B, Ismail N, Moodley P, et al. Transmission of
562 Extensively Drug-Resistant Tuberculosis in South Africa. N ENGL J MED. 2017;
563 376;3 DOI: 10.1056/NEJMoa1604544
564 29. Seyoum B, Demissie M, Worku A, Bekele S, and Aseffa A. Prevalence and Drug
565 Resistance Patterns of *Mycobacterium tuberculosis* among New Smear
566 Positive Pulmonary Tuberculosis Patients in Eastern Ethiopia. Hindawi
567 Publishing Corporation. Tuberculosis Research and Treatment 2014; Volume
568 2014, Article ID 753492, 7 pages http://dx.doi.org/10.1155/2014/753492
569 30. Stucki D, Brites D, Jeljeli L, Coscolla M, Liu Q, Trauner A et al. *Mycobacterium*
570 *tuberculosis* Lineage 4 comprises globally distributed and geographically
571 restricted sublineages. Nat Genet. 2016; 48(12): 1535–1543.
572 31. Malm S, Linguissi LS, Tekwu EM, Vouvoungui JC, Kohl TA, Beckert P et al. New
573 Mycobacterium tuberculosis Complex Sublineage, Brazzaville, Congo. Emerg
574 Infect Dis. 2017;23(3): 423-429.
575 32. Niemann S, Diel R, Khechinashvili G, Gegia M, Mdivani N and Tang YW.
576 *Mycobacterium tuberculosis* Beijing Lineage Favors the Spread of Multidrug-
577 Resistant Tuberculosis in the Republic of Georgia. J. Clin. Microbiol.
578 2010;48(10): 3544– 3550.

579 33. Coll F, McNerney R, Guerra-Assunção JA, Glynn JR, Perdigão J, Viveiros M et al. A
580 robust SNP barcode for typing Mycobacterium tuberculosis complex strains.Nat.
581 Commun. 2014; 5:4812. doi: 10.1038/ncomms5812 (2014).
582 34. Tadesse M, Abebe G, Bekele A, Bezabih M, Rijk P and Meehan C. The predominance
583 of Ethiopian specific Mycobacterium tuberculosis families and minimal contribution
584 of Mycobacterium bovis in tuberculous lymphadenitis patients in Southwest
585 Ethiopia, Infection, Genetics and Evolution. 2017;
586 doi:10.1016/j.meegid.2017.09.016
587 35. Sharaf-Eldin G, Fadl-Elmula I, Ali M, Ali A and Salih A. Tuberculosis in Sudan: a
588 study of *Mycobacterium tuberculosis* strain genotype and susceptibility to anti-
589 tuberculosis drugs. BMC Infect Dis. 2011;11:219.
590 36. Parwati I, Crevel R and Soolingen D. Possible underlying mechanisms for
591 successful emergence of the Mycobacterium tuberculosis Beijing genotype
592 strains. Lancet Infect Dis. 2010; 10: 103–11.
593 37. Comas I, Elina H, Teklu K, Shiferaw B, Wondale M, Balako G, et al. Population
594 genomics of Mycobacterium tuberculosis in Ethiopia contradicts the virgin soil
595 hypothesis for human tuberculosis in sub-Saharan Africa. Curr.Biol. *2015;25,*
596 *3260– 3266.*
597
598
599

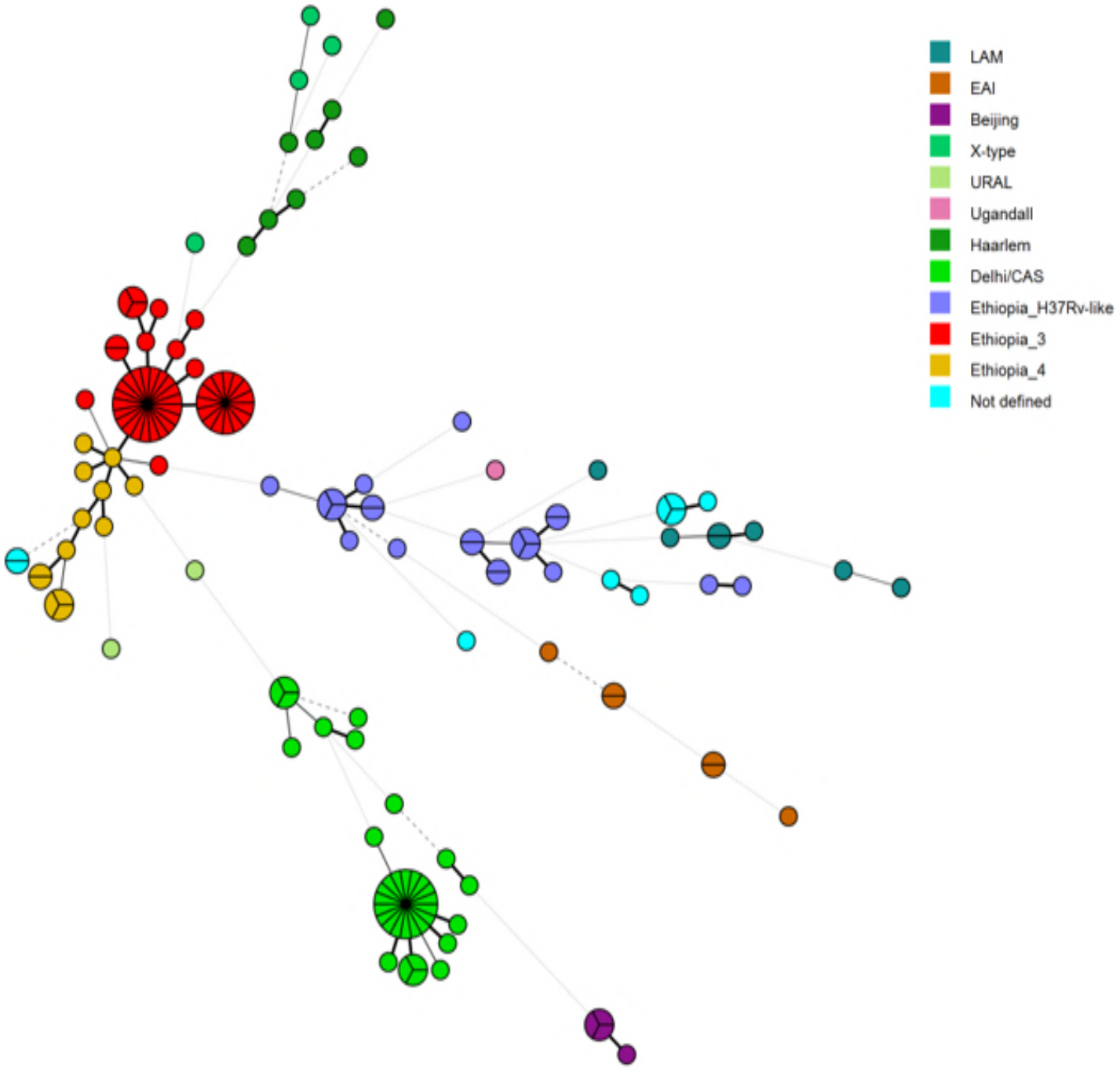

Fig 1. Minimum spanning tree based on MIRU/VNTR profiles of 24- loci of 160 MTBC isolates in eastern Ethiopia. A circle representing a specific genotype is divided in to the number of strains clustering in it. EAI: East Africa India, LAM=M. tuberculosis Latin American Mediterranean.

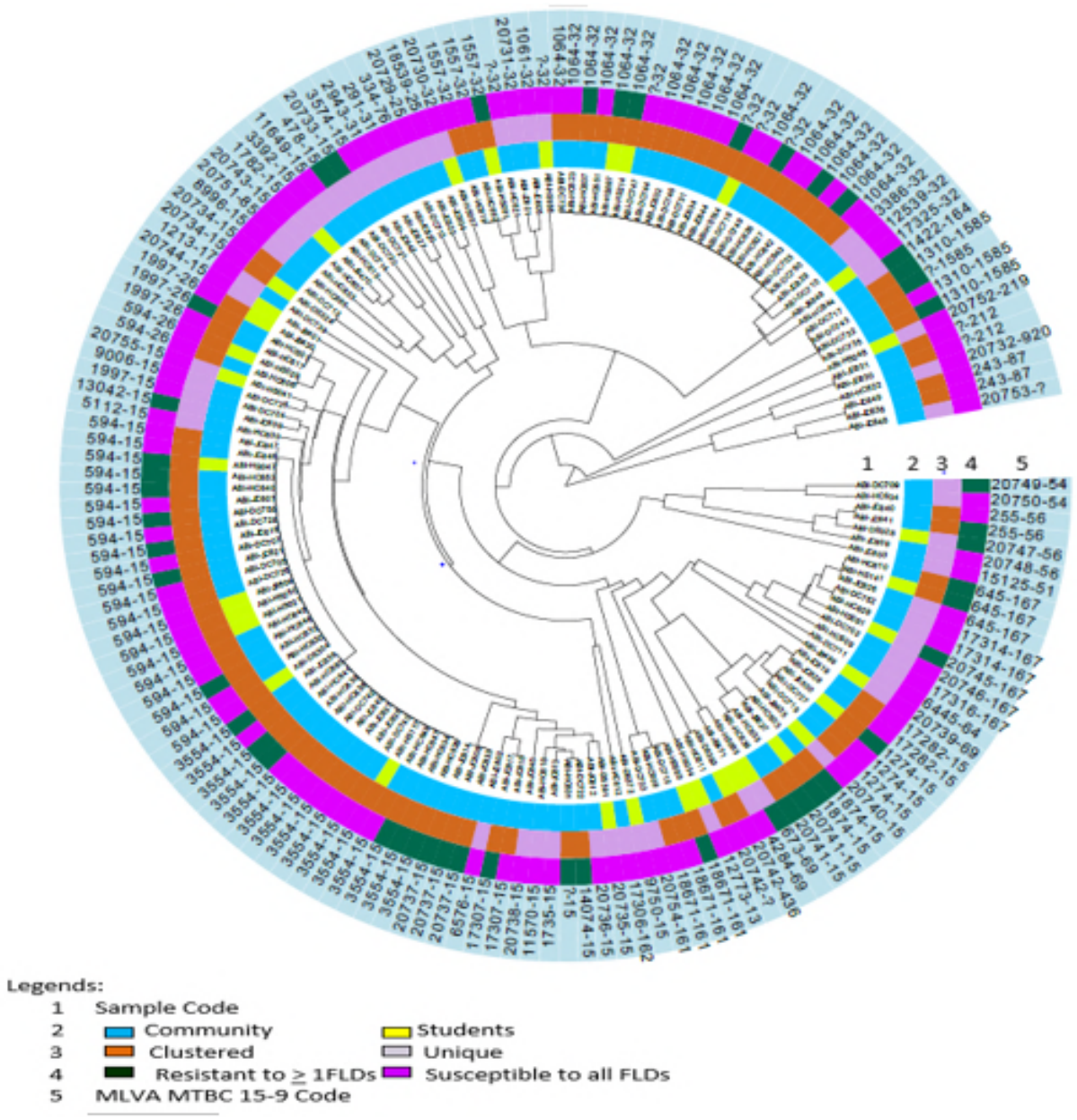

Fig 2. Radial UPGMA tree based on MIRU/VNTR 24-loci copy number. patient origin (Community or student), Clustering, Resistance to one or more FLDs (First line Anti TB drugs), and MLVA MTBC 15-9 Code.