

RESEARCH

Open Access



Strain structure analysis of *Mycobacterium tuberculosis* circulating among HIV negative, positive and drug resistant TB patients attending chest clinics in Western Kenya

Martin O. Ogwang^{1*}, Lameck Diero², Florence Ng'ong'a³, Gabriel Magoma³, Lucy Mutharia⁴, Mabel Imbuga³ and Caroline Ngugi³

Abstract

Background Despite global tuberculosis (TB) interventions, the disease remains one of the major public health concerns. Kenya is ranked 15th among 22 high burden TB countries globally.

Methods A cross-sectional study was conducted in Western Kenya, which comprises 10 counties. A multistage sampling method was used where a single sub-county was randomly selected followed by sampling two high volume health facility from each sub-county. Identification of spoligotype profiles and their family distribution and lineage level were achieved by comparison with SITVIT database.

Results Lineage distribution pattern revealed that the most predominant lineage was CAS 220 (39.8%) followed by Beijing 128 (23.1%). The other lineages identified were T, LAM, H, X, S and MANU which were quantified as 87 (15.7%), 67 (12.1%), 16 (2.8%), 10 (1.8%), 8 (1.4%) and 5 (0.9%) respectively. CAS and Beijing strains were the most predominant lineage in both HIV negative and positive TB patients. The Beijing lineage was also the most predominant in resistant *M. tuberculosis* strains as compared to wild type. A total of 12 (2.0%) were orphaned *M. tuberculosis* strains which were spread across all the 10 counties of the study site. In multivariate logistic regression adjusting for potential cofounders three potential risk factors were significant. HIV status (OR = 1.52, CI = 0.29–3.68 and *P* value of 0.001), Alcohol use (OR = 0.59, CI = 0.43–3.12 and *P*-value = 0.001) and cross border travel (OR = 0.61, CI = 0.49–3.87 and *P* value = 0.026). Most *M. tuberculosis* clinical isolates showed genetic clustering with multivariate logistic regression indicating three potential risk factors to clustering. HIV status (OR = 1.52, CI = 0.29–3.68 and *P* value of 0.001), Alcohol use (OR = 0.59, CI = 0.43–3.12 and *P*-value = 0.001) and cross border travel (OR = 0.61, CI = 0.49–3.87 and *P* value = 0.026).

Conclusion There exist diverse strains of *M. tuberculosis* across the 10 counties of Western Kenya. Predominant distribution of clustered genotype points to the fact that most TB cases in this region are as a result of recent transmission other than activation of latent TB.

Keywords Spoligotyping, Tuberculosis epidemiology, Strain structure, Drug-resistance

*Correspondence:

Martin O. Ogwang
macogwang2012@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Tuberculosis remains a global threat to public health and is the leading cause of death by a single infectious agent with 1.6 million deaths in 2017. An estimated 10 million people developed TB disease in 2017 but only 6.4 million (61%) were notified [1]. Despite the global aim of a 95% reduction in TB deaths, 90% reduction in incidence and 0% TB affected families facing catastrophic costs due to TB by 2035 [2], the disease still remains a major challenge on livelihood. The emergence of resistant strains to first- and second-line treatment regimen has seen an upsurge of monoresistant, multidrug resistant and even pre-extensive multidrug resistant tuberculosis. Surprisingly, *Mycobacterium tuberculosis*, the causative agent of tuberculosis, is capable of surviving and replicating within macrophages and other human cells.

For this reason, the bacterium has developed a unique mechanism to gain entry and survive in host cells. With the increasing incidence, tuberculosis has become one of the leading causes of death among immunocompromised patients especially those who are HIV-TB coinfectd. Kenya’s TB prevalence survey of 2015 indicated that the disease incidence was 233 per 100,000 (95% CI 188–266)

[3] with the previous survey having been carried out in 1958. The results of the two surveys showed that the drivers of TB have certainly changed over the past 60 years. Another survey conducted in Kenya concluded that TB incidence has been previously underestimated [4].

Western Kenya consists of 10 county governments (Fig. 1) with varying number of sub-counties ranging from 5 to 12 depending on the population, landmass and vastness of the respective county. In that respect, Vihiga is the smallest county with 5 sub-counties while Kakamega is the largest with 12 sub-counties. Western Kenya counties border Tanzania and Uganda with close proximity to Rwanda, South Sudan and Somalia. Increasing interaction of citizens of these countries could be a risk factor to cross border transmission of tuberculosis further complicating the disease management in Kenya. This study established the population structure of *Mycobacterium tuberculosis* circulating in HIV negative and positive, and drug-resistant TB patients Western Kenya. The study also reports the existence of cross border transmission of tuberculosis. Evaluate the enablers, barriers and level of preparedness of western Kenya communities on tuberculosis cross border transmission.

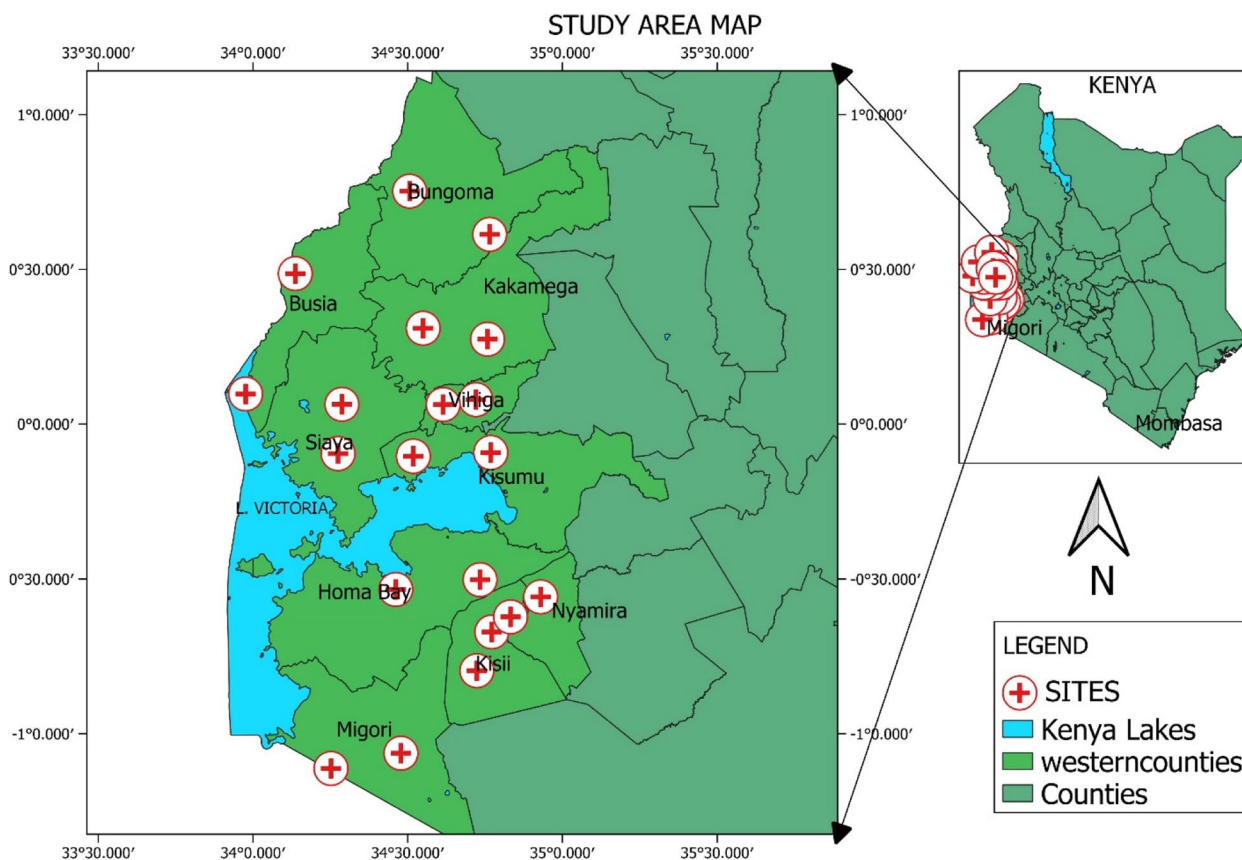


Fig. 1 Map of Kenya showing study sites

Since 1940s when drug resistance was first reported, it has been identified as one of the major global challenge in the control and prevention of TB [5]. Tuberculosis burden in Africa is greatly driven by emergency and spread of Multidrug resistant (MDR) TB strains. While so many studies on molecular mechanisms of multidrug resistant and strain characterization have been carried out, early diagnosis remains a major set back of transmission of the strains due to increasing asymptomatic cases associated with MDR-TB [6]. Currently, drug susceptibility testing has been developed for diagnosis of MDR-TB, culture based phenotypic drug susceptibility has been relied on despite the fact that it takes long for results to be relayed. Delayed detection MDR-TB as a result of slow growth of the bacteria during culture and susceptibility test has been a challenge for early diagnosis and treatment of patients with MDR-TB [7]. Alternative methods for detection of MDR-TB such as the polymerase chain reaction-based methods which include molecular Line Probe Assay [8] has contributed greatly to early diagnosis despite the fact that the technology has not been rolled out in most low- and middle-income countries [9].

Molecular typing has led to great strides in genotypic characterization of *Mycobacterium tuberculosis* hence contributing significantly in the understanding of TB epidemiology. This has contributed immensely in improving TB control by availing critical information on transmission dynamics and determination of reactivation against exogenous reinfection [10, 11]. Genotyping of *M. tuberculosis* has been instrumental in disease TB surveillance with regard to investigating or confirming outbreaks, examining cross contamination that could have occurred in the laboratories and identification of the spread of successful clones like multidrug resistant TB. Genotypically, it's possible to determine different population structures of *M. tuberculosis* thriving in a given geographical area with their possible pathological properties [12].

Spoligotyping technique has been used in *Mycobacterium tuberculosis* typing [13]. It is a PCR reverse-hybridization blotting based technique which has been widely used to study genetic diversity of the Direct Repeat locus of *M. tuberculosis* bacteria which is important in understanding the bacterium evolution, molecular epidemiology and population genetics. This technique is important in studying unique Events Polymorphism and insertion-Sequence-mediated genetic recombination [14]. This technique assesses the genetic diversity of direct repeat (DR) locus [15], which contains multiple 36-base pair (bp) DRs that are separated by 37 to 41 bp unique spacer sequences [16]. The 43 spacers are commonly used for genotyping [15]. Classical spoligotyping is performed by reverse line blot hybridization of biotinylated PCR products to a membrane with 43 covalently bound synthetic

oligonucleotides representing the different spacers selected from *M. tuberculosis* H37Rv (spacers 1–19, 22–32, and 37–43) and *M. bovis* BCG (spacers 20–21 and 33–36) [17]. The presence and absence of each spacer is specific for each individual and used for genotyping.

Strain structure identification will help determining the impact of differing genetic characteristics among Mtb complex strains on vaccine efficacy, treatment outcome, development of MDR/XDR Mtb strains, and epidemiological outbreaks by focusing on the best-adapted human *M. tuberculosis* lineages. Strain structure will help understand factors contributing to *M. tuberculosis* success and survival strategies based on lineages and sub lineages survival characteristics. Cross border transmission of *M. tuberculosis* evaluation will help in form cross border events associated with it hence help in policy formulation of continuous surveillance and intervention at the border entry point. While determining enablers and barriers associated with cross boarder transmission prevention of *M. tuberculosis* is important in informing proper management and strategy for those identified to be carrying cross border strains.

Materials and methods

Study design and population

A cross sectional study design was applied in the study with multistage sampling method to select a single sub-county in each of the study counties. Smear positive tuberculosis patients above 18 years who had resided in the study area for at least 6 months were eligible for enrolment with only those who consented to the study recruited. Tuberculosis-HIV coinfecting and non-coinfecting were proportionately recruited based TB-HIV coinfection prevalence in each county as reported by Ministry of Health.

Case definition

A case of tuberculosis disease was defined as a patient having two acid fast bacilli positive test or a single positive *Mycobacterium tuberculosis* culture. A cluster case was defined as any TB case from a study isolates whose strain type based on standard spoligotyping assay was not distinguishable from at least one other case. On the other hand, unique cases shared a unique strain not found in the study population.

Sampling method

A multistage sampling method was used where a single sub-county was randomly selected followed by sampling two high volume health facility from each sub-county. Consenting study subjects with at least 2 smear positive sputum at the time of enrolment were randomly selected. A total of 553 samples were collected across 10 counties

between 2018 and 2019. Sample size was calculated based on [18].

Collection and transportation of sputum

Two sputum specimens (spot and early morning) were collected from participants with presumptive TB under the supervision of trained and competent medical staff. Upon consenting to the study, the patients were requested to cough so that expectoration came from deep down the chest as possible, and spit into a sterile 50 ml screw cap tubes. The samples were refrigerated at 4°C awaiting transportation in cool boxes to the Mycobacteria Reference Laboratory, Moi University School of Medicine (MRL, MUSOM) weekly for analysis. Samples were processed within 7 days of collection in order to minimize loss of viability of the *Mycobacterium*. Participants also underwent phlebotomy for HIV testing. The blood was delivered into Vacutainer Brand STERILE interior EDTA (K3) tubes and stored at -20°C waiting processing. The samples were transported in cool boxes to MRL, MUSOM, Eldoret, and processed within 2 weeks [19].

Identification of *Mycobacterium* from clinical samples

BACTEC MGIT 960 assay was used in determining *Mycobacterium* strains from clinical samples. Briefly, the specimens were decontaminated and digested with an equal volume of 4% sodium hydroxide for 15 minutes while vortexing after every 5 minutes. Decontaminated sputum was then transferred into a 50 ml corning tube, and then centrifuged at 3000 rpm for 15 minutes and the supernatant discarded. Phosphate-buffered saline (PBS, pH 6.8) was added to the residue raising the volume to 50 ml. The content was then centrifuged at 3000 rpm for 15 minutes and the supernatant discarded. 0.5 ml of PBS, pH 6.8 was added to resuspend the pellets, then 0.5 ml of resuspended deposit was added to the MGIT culture tube which were then incubated into the BACTEC MGIT 960 system until positivity was observed for positive samples or up to 8 weeks for negative samples [20]. Positive tubes were identified, removed and confirmed for acid fastness (AFB), for species identification and drug susceptibility testing and later storage of isolates.

Identification *M. tuberculosis* from mycobacteria other than tuberculosis (MOTT) strains

The Capilla TB Test [21] was performed directly from positive cultures (from broth or from re-suspended colonies from solid media). Briefly, the Capilla TB assay was performed by placing 100 µl from a broth culture onto the specimen placement area of the Capilla TB cartridge, allowing a maximum of 15 min incubation and observing a purple–reddish color change in the test area. This was the indicator that the sample under test was

Mycobacterium tuberculosis complex, and those that did not show the purple–reddish color was considered as MOTTs and therefore eliminated from the study.

Culture

The ABBL MGIT tube (from Becton Dickinson) containing 7 mL modified middle brook 7H9 broth was used, to which an enrichment supplement as well as a mixture of antibiotics consisting of polymyxin B, amphotericin B, nalidixic acid, trimethoprim, and azlocillin were added [22]. After inoculation, the tubes were incubated at 37°C. Readings were taken daily for the first 3 weeks and once a week thereafter for culture positivity until the end of 6 weeks using the BBL Micro MGIT system [23]. All the positive tubes were further confirmed by ZN staining and a sub culturing on blood agar plate and a LJ slant. The time to detection (TTD) of Mycobacteria were based on the date of the earliest instrumental indication of positivity.

Extraction of *M. tuberculosis* DNA

Each sample pellet was dispersed in 10 mM Tris-1 mM EDTA buffer containing 0.1% Tween 80 and 2 mg/ml lysozyme. The tube was incubated for 2 hrs at 37°C with intermittent shaking and centrifuged. The pellet was lysed by redissolving in TE buffer containing Proteinase K (100 µg/ml) and 1% (w/v) sodium dodecyl sulphate and incubated for 1 hr. at 37°C. Then, an equal volume of TE saturated phenol: chloroform: iso-amyl alcohol (25:24:1 v/v/v) was added. The aqueous phase was transferred to another sterile tube and 0.1 volume of cold 3 M sodium acetate (pH 5.2) was added. The sample was then mixed by inversion and placed on ice for 10 min before centrifugation for 1 min at a speed of $11,000\times g$ at room temperature. The supernatant was transferred to another sterile tube and DNA precipitated by the adding acrylamide (20 µg/ml), 0.05 volumes of 3 M sodium acetate and 2.5 volume of ethanol. The DNA was then washed, dried, and dissolved in 25 µl sterile triple distilled water [24].

Quantification of extracted *M. tuberculosis* DNA

Quantification of DNA was done spectrophotometrically by measuring absorbance at 260 nm for comparative quantitative analysis of DNA [25]. Serial dilutions of DNA extracted were prepared ranging from 0.5 µg to 1.0 µg. All of these dilutions were amplified to determine the minimum amplifiable quantity of DNA.

Spoligotyping

In this study, the classical 43-spacer format of spoligotyping was performed as previously described [16]. DNA samples of the *M. tuberculosis* H37Rv and *Mycobacterium bovis* BCG strains were included as positive

controls. Molecular biology-grade water was used as negative control. The spoligotypes (presence and absence of spacers) were then recorded in 43-digit binary format and compared with those recorded in the SpolDB4 database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) to identify the Spoligotype International Type (SIT) and family [14]. For the spoligotypes that matched the SITs, but could not be related to any family (i.e., unknown), and for the spoligotypes that were not present in the SpolDB4 database (e.g., orphan), the SPOTCLUST program, which were built from the spolDB3 database (http://tbinsight.cs.rpi.edu/run_spotclust.html) [26], was used to search for *M. tuberculosis* family similarity. In the SPOTCLUST analyses, the family assignment was retained when the probability is $\geq 90\%$. Nevertheless, the final designation of families and subfamilies was also based on the MIRU-VNTR data.

Spoligotype cluster analysis

Association between patient characteristics and genotype clustering were evaluated using logistic regression and frequency analysis. The potential risk factors for clustering included age, HIV status, sex, level of education, occupation, monthly income, size of the house, cigarette smoking, alcohol use and cross boarder travels. Bivariate analysis was performed to determine the association of potential risk factors with genotypic clustering followed by multivariate logistic regression adjusting for potential cofounders. Since TB cases were from adjacent and approximate counties and that all cases were notified within the same calendar year, the specific geographical and temporal timing of cases were not used in the definition of a cluster. Lineage clustering rate was computed from the difference between the total strains clustered cases in each lineage and the number of clusters divided by the sum of cases in the MTB lineage.

Results

Social demographic characteristics of the study population (N = 553)

Participants in this study were 18 years and above with the most affected being between 25 and 30 years representing 33.1% of the total sample size. 48.3% of the participants were HIV positive while 51.7% were HIV negative. Majority of the participants were males (62.4%) while the rest were females. Most of the participants (44%) were primary school leavers, 26% with informal education, 23% secondary school leavers while participants with tertiary education were 7%. Self-employed participants 39%, those with informal employment were 23.6%, unemployed were 16.6% while 15% were employed with 1.8% still in school. Participants with monthly income in the bracket of 10,000–19,999 were 31.5% with the least being

those with monthly income of Ksh. 50,000 and above (Table 1).

In terms of housing, 37.4% of the participants were living in single bed roomed houses, 34.5% in two bed roomed houses, 2% in bedsitters, 22.1% in single rooms, the rest were 2%. Also, among the participants, 76.1%

Table 1 Social demographic Characteristics of the study population (N=553)

Variable	Status	Frequency	Percentage
Age	18–24	145	26.1
	25–30	183	33.1
	31–36	107	19.3
	37–42	47	8.5
	43–48	35	6.3
	49–54	18	3.3
	55–60	12	2.2
	61+	6	1.1
HIV status	Positive	267	48.3
	Negative	286	51.7
Sex	Female	208	37.6
	Male	345	62.4
Education	No formal	144	26
	Primary	243	44
	Secondary	127	23
	Tertiary	39	7
Occupation	Employed	83	15
	Self employed	218	39
	Informal employment	128	23
	Unemployed	103	16.6
	Student	10	1.8
	Other	11	2
Monthly income	Less than Ksh. 1000	105	19
	Ksh. 1000–9,999	160	28.9
	Ksh. 10,000–19,999	174	31.5
	Ksh. 20,000–29,999	48	8.7
	Ksh. 30,000–39,999	34	6.1
	Ksh. 40,000–49,999	18	3.3
	Ksh. 50,000 +	14	2.5
Size of the house	Single room	122	22.1
	Bedsitter	11	2
	One bedroom	207	37.4
	Two bedroom	191	34.5
	Other	22	4
Cigarette smoking	Yes	132	23.9
	No	421	76.1
Alcohol use	Yes	151	27.3
	No	402	72.7
Cross boarder travels	Yes	211	38.2
	No	342	61.8

were non cigarette smokers while 23.9% were smokers. On the same note, 72.7% did not take alcohol while 27.3% were alcohol users.

In terms of local and international travels, 38.2% were involved in cross boarder travel while 61.8% did not (Table 1).

Spoligotyping

The distribution of *M. tuberculosis* strains in Western Kenya is highly diversified with Euro American strains (H, LAM, T, S and X), Indo Oceanic (MANU), East Asia strains and Beijing strains being identified across the 10 counties of Western Kenya. Lineage distribution pattern revealed that the most predominant lineage was CAS 220 (39.8%) followed by Beijing 128 (23.1%). The other lineages identified were T, LAM, H, X, S and MANU which were quantified as 87(15.7%), 67(12.1%), 16(2.8%), 10(1.8%), 8(1.4%) and 5(0.9%) respectively (Table 2, Fig. 1).

Clade distribution among HIV negative and positive, CAS and Beijing strains were the most equally distributed among HIV positive and negative participants. T and unknown strains were predominantly detected mostly in HIV positive respondents. While CAS, H, LAM, MANU, S, T and X were predominant in HIV negative (Table 2, Fig. 1).

Clade distribution among wild type and resistant *M. tuberculosis* TB patients, revealed that Beijing lineage was most identified in resistant strains than wild type strains with orphaned strains also being detected (Table 2, Fig. 1).

Across all the 10 counties of Western Kenya, CAS strain was the most predominant followed by Beijing, T and finally LAM (Fig. 2). In Bungoma county, Beijing, CAS, H, LAM and T *M. tuberculosis* lineages were detected representing 6 (1.1%), 17 (3.1%), 2 (0.4%), 5 (0.9%) and 7 (1.2%) respectively (Table 3).

In Busia, four major lineages were detected with CAS being the dominant strain with a proportion of 20 (3.4%), followed by Beijing (2.7%), T (2.0%) and LAM (1.6%). Subsequently, two unknown strains of *M. tuberculosis* were detected. Similarly, the most prevalent lineage detected in Homa Bay was CAS (8.1%), followed by Beijing (4.9%) with 0.5% of the total samples typed were unknown (Table 3).

In Kakamega, Beijing, CAS, H, LAM, T and X lineages were detected at the proportion of 2.4, 4.5, 0.2, 0.5, 1.6, 0.5 and 0.2% respectively. Subsequently, in Kisumu county just like in Kakamega County CAS lineage was the most prevalent at a proportion of 5.2%, followed by Beijing and LAM both with the proportion of 3.4 and 1.1%, H had a proportion of 0.2%, T and X having a proportion

of 1.1 and 0.4% respectively. The proportion of unknown lineage in Kisumu County was 0.5% (Table 3).

In Migori 0.7% of the samples typed were unknown and did not belong to any lineage were detected, however the majority of strains identified belonged to CAS lineage (4.1%) followed by T lineage (1.3%). Beijing lineage identified in Migori were 3.6% in proportion to the total sample typed while LAM, H and X were represented by 0.5, 0.4 and 0.4% respectively (Table 3).

In Kisii, 0.3% were unknown strains not belonging to any lineage as compared to SITVIT database. In this county, majority of samples typed belonged to Beijing strain followed by CAS, LAM, T, H and MANU at the proportion of 3.4, 0.2, 1.1, 0.2 and 0.2% respectively (Table 3).

In Vihiga, the predominant clades were CAS, Beijing and T just as many other Western Counties represented by a proportion of 3.3, 2.3 and 1.6% respectively. On the other hand, in Siaya County, four lineages were found to dominate among TB patients; CAS, Beijing, LAM and T at the proportion of 4.1, 2.3, 2.0% and then 2.0%. Then Finally, Nyamira recorded CAS at a proportion 2.9%, LAM had a proportion 1.3% and finally Beijing 1.1% (Table 3).

Bivariate analysis of potential risk factors and genotypic clustering of *M. tuberculosis* showed that six factors showed significant association. These factors included age of 25–30 with odd ratio of 0.95, confidence interval of 0.35–2.27 and *P*-value of 0.003. HIV status was significantly associated with genotypic clustering with OR (Odd Ratio) of 1.73, CI (Confidence Interval) of 0.42–5.21 and *P* value of 0.04. Those with monthly income of between Ksh. 1000–Ksh. 9999 were significantly associated with genotypic clustering having OR of 1.28, CI of 0.62–4.00 and *P*-value of 0.046. There was significant association between those living in single rooms with genotypic clustering with OR of 1.43, CI of 0.46–4.31 and *P*-value of 0.049. Alcohol and cross boarder travels were also significantly associated with genetic clustering with OR of 0.87, 0.93; CI of 0.34–5.34, 0.46–3.92 having *P*-values of 0.003, 0.002 respectively. In multivariate logistic regression adjusting for potential cofounders three potential risk factors were significant. HIV status (OR=1.52, CI=0.29–3.68 and *P* value of 0.001), Alcohol use (OR=0.59, CI=0.43–3.12 and *P*-value =0.001) and cross border travel (OR=0.61, CI=0.49–3.87 and *P* value =0.026) (Table 4).

Discussion

Strain structure identification is key in understanding the circulating *M. tuberculosis* clades within TB patients hence help in informing targeted policy formulation with regards to enhancing preventive and control strategies.

Table 2 Showing *M. tuberculosis* strain distribution pattern as compared to SITVIT database. Strain distribution in HIV positive and negative TB patients and also their distribution patterns in TB drug resistant and drug sensitive patients

Clade	Lineage	SIT	Spoligotype	Number	Resistance Pattern		HIV STATUS		
					Any resistance	Wild type	HIV-VE	HIV+VE	
BEIJING	BEIJING	1		125(22.6%)	33 (6.0%)	92(17.0%)	58(10.5%)	69(12.5%)	
	BEIJING	265		3(0.5%)	1(0.2%)	2(0.4%)	3(0.5%)	0(0.0%)	
CAS	CAS	2100		3(0.5%)	1(0.2%)	2(0.4%)	0(0.0%)	1(0.2%)	
	CAS	203		3(0.5%)	1(0.2%)	2(0.4%)	2(0.4%)	1(0.2%)	
	CAS1_DELHI	26		54(9.8%)	12(2.2%)	42(7.6%)	36(6.5%)	18(3.3%)	
	CAS1_DELHI	357		3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)	
	CAS1_DELHI	25		3(0.5%)	0(0.0%)	3(0.5%)	2(0.4%)	1(0.2%)	
	CAS1_KILI	21		129(23.3%)	22(4.0%)	107(19.3%)	67(12.1%)	64(11.6%)	
	CAS2	288		10(1.8%)	3(0.5%)	7(1.3%)	7(1.3%)	3(0.5%)	
	Not in SITVIT	Not in SITVIT			3(0.5%)	0(0.0%)	3(0.5%)	0(0.0%)	3(0.5%)
	Not in SITVIT	Not in SITVIT			3(0.5%)	0(0.0%)	3(0.5%)	2(0.4%)	1(0.2%)
	Not in SITVIT	Not in SITVIT			3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)
H	H3	50		10(1.8%)	2(0.4%)	7(1.3%)	4(0.7%)	6(1.1%)	
	H1	45		3(0.5%)	1(0.2%)	2(0.4%)	0(0.0%)	3(0.5%)	
	Not in SITVIT	Not in SITVIT			3(0.5%)	1(0.2%)	2(0.4%)	2(0.4%)	1(0.2%)
	LAM	LAM	4		33(6.0%)	3(0.5%)	30(5.4%)	19(3.4%)	14(2.5%)
		LAM3	2042		3(0.5%)	1(0.2%)	2(0.4%)	3(0.5%)	0(0.0%)
		LAM	2345		3(0.5%)	1(0.2%)	2(0.4%)	0(0.0%)	3(0.5%)
		LAM9	42		3(0.5%)	1(0.2%)	2(0.4%)	2(0.4%)	1(0.2%)
		LAM9	1470		3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)
		LAM11_ZWE	59		13(2.4%)	3(0.5%)	10(1.8%)	9(1.6%)	4(0.7%)
		Not in SITVIT	Not in SITVIT			3(0.5%)	1(0.2%)	2(0.4%)	2(0.4%)
Not in SITVIT		Not in SITVIT			3(0.5%)	2(0.4%)	1(0.2%)	1(0.2%)	2(0.4%)
Not in SITVIT	Not in SITVIT			3(0.5%)	3(0.5%)	0(0.0%)	2(0.4%)	1(0.2%)	
MANU	MANU2	54		5(0.9%)	1(0.2%)	4(0.7%)	3(0.5%)	2(0.4%)	
	S	S	34		5(0.9%)	2(0.4%)	3(0.5%)	2(0.4%)	1(0.2%)
S		156		3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)	
T	T1	53		10(1.8%)	3(0.5%)	7(1.3%)	4(0.7%)	6(1.1%)	
	T1	522		10(1.8%)	3(0.5%)	8(1.4%)	3(0.5%)	7(1.3%)	
	T1	649		10(1.8%)	1(0.2%)	9(1.6%)	6(1.1%)	4(0.7%)	
	Not in SITVIT	Not in SITVIT			5(0.9%)	2(0.4%)	3(0.5%)	1(0.2%)	4(0.7%)
	T1	196		5(0.9%)	3(0.5%)	2(0.4%)	3(0.5%)	2(0.4%)	
	T1	122		3(0.5%)	3(0.5%)	0(0.0%)	0(0.0%)	3(0.5%)	
	T2	52		10(1.8%)	1(0.2%)	9(1.6%)	3(0.5%)	7(1.3%)	
	T2	ORPHAN			3(0.5%)	0(0.0%)	3(0.5%)	2(0.4%)	1(0.2%)
	T2		848		3(0.5%)	0(0.0%)	3(0.5%)	3(0.5%)	0(0.0%)
	T2		2867		3(0.5%)	0(0.0%)	3(0.5%)	3(0.5%)	0(0.0%)
	T2	128		3(0.5%)	1(0.2%)	2(0.4%)	3(0.5%)	0(0.0%)	
	Not in SITVIT	Not in SITVIT			3(0.5%)	2(0.4%)	1(0.2%)	2(0.4%)	1(0.2%)
	T2	135		3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)	
	T3	37		5(0.9%)	3(0.5%)	2(0.4%)	4(0.7%)	1(0.2%)	
	T3_ETH	149		5(0.9%)	2(0.4%)	3(0.5%)	3(0.5%)	2(0.4%)	
Not in SITVIT	Not in SITVIT			3(0.5%)	1(0.2%)	2(0.4%)	0(0.0%)	3(0.5%)	
T	78		3(0.5%)	1(0.2%)	2(0.4%)	2(0.4%)	1(0.2%)		
X	X2	137		10(1.8%)	7(1.3%)	3(0.5%)	7(1.3%)	3(0.5%)	
Orphaned	Not in SITVIT	Not in SITVIT		3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)	
	unknown	1778		3(0.5%)	2(0.4%)	1(0.2%)	0(0.0%)	3(0.5%)	
	Not in SITVIT	Not in SITVIT		3(0.5%)	1(0.2%)	2(0.4%)	1(0.2%)	2(0.4%)	
	Not in SITVIT	Not in SITVIT		3(0.5%)	1(0.2%)	2(0.4%)	2(0.4%)	1(0.2%)	
	no spoligo	(repeated)		0					
				553(100%)	140(25.3%)	413(74.7%)	286(51.7%)	267(48.3%)	

Strain distribution pattern in ten Counties of Western Kenya

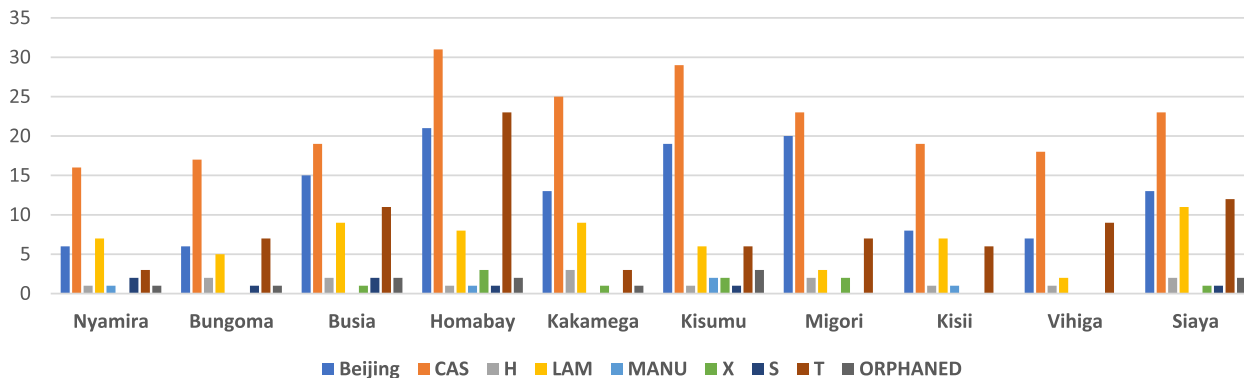


Fig. 2 Clade distribution in 10 counties of Western Kenya

Table 3 Strain distribution pattern in Western Kenya counties

	Beijing	CAS	H	LAM	MANU	X	S	T	ORPHANED
Nyamira	6(1.1%)	16(2.9%)	1(0.2%)	7(1.3%)	1(0.2%)	0(0.0%)	2(0.4%)	3(0.5%)	1(0.2%)
Bungoma	6(1.1%)	17(3.1%)	2(0.4%)	5(0.9%)	0(0.0%)	0(0.0%)	1(0.2%)	7(1.2%)	1(0.2%)
Busia	15(2.7%)	19(3.4%)	2(0.4%)	9(1.6%)	0(0.0%)	1(0.2%)	2(0.4%)	11(2.0%)	2(0.4%)
Homabay	21(3.8%)	31(5.6%)	1(0.2%)	8(1.4%)	1(0.2%)	3(0.5%)	1(0.2%)	23(4.2%)	2(0.4%)
Kakamega	13(2.4%)	25(4.5%)	3(0.5%)	9(1.6%)	0(0.0%)	1(0.2%)	0(0.0%)	3(0.5%)	1(0.2%)
Kisumu	19(3.4%)	29(5.2%)	1(0.2%)	6(1.1%)	2(0.4%)	2(0.4%)	1(0.2%)	6(1.1%)	3(0.5%)
Migori	20(3.6%)	23(4.1%)	2(0.4%)	3(0.5%)	0(0.0%)	2(0.4%)	0(0.0%)	7(1.3%)	0(0.0%)
Kisii	8(1.4%)	19(3.4%)	1(0.2%)	7(1.2%)	1(0.2%)	0(0.0%)	0(0.0%)	6(1.1%)	0(0.0%)
Vihiga	7(1.3%)	18(3.3%)	1(0.2%)	2(0.4%)	0(0.0%)	0(0.0%)	0(0.0%)	9(1.6%)	0(0.0%)
Siaya	13(2.3%)	23(4.1%)	2(0.4%)	11(2.0%)	0(0.0%)	1(0.2%)	1(0.2%)	12(2.2%)	2(0.4%)
TOTAL	128 (23.1%)	220(39.8%)	16(2.8%)	67(12.1%)	5(0.9%)	10(1.8%)	8(1.4%)	87(15.7%)	12(2.2%)

This was the first study to determine strain diversity and distribution in Western Kenya circulating among HIV negative and positive patients in the region. HIV/AIDS is the major risk factor to acquiring TB in Western Kenya due to its effecting compromising human immune system [27]. high prevalence rates of the HIV epidemic in this region especially among the five counties; Kisumu, Homa Bay, Siaya, Busia and Migori [28]. The High prevalence rate has been contributed by socio culture behavior of the natives of this area [29]. It is therefore important to understand the *M. tuberculosis* strain structure circulating in clinical isolates of this region. Additionally, there has been increased circulation of drug resistant *M. tuberculosis* of varying degrees ranging from mono resistant, multi resistant and poly resistant strains. Drug resistance structure of *M. tuberculosis* under study was reported by Ogwang et al., [30] thus this study intended to establish *M. tuberculosis* drug resistance strain structure through spacer oligonucleotide typing (Spoligotyping). Our findings indicate a circulation of diverse strains

of *M. tuberculosis* in the study area. The study established that despite the existing of predominant strains like CAS, Beijing and LAM, there was no clear association with either HIV status of the patient or drug resistant *M. tuberculosis*.

The distribution of *M. tuberculosis* strains in Western Kenya is highly diversified with Euro American strains (H, LAM, T, S AND X), Indo oceanic (MANU) and East African strains being identified in addition to Beijing strain. Studies by Merker et al., [31] showed that Africa is the only continent with all the seven human adapted lineages with further evidence indicating that MTBC originated from Africa and spread globally through human migration events and slave trade [32]. This could explain the findings of this study with regards to genetic diversity of *M. tuberculosis* lineages in Western Kenya.

Spoligotyping patterns of some clinical isolates could not be matched to other reported in international database hence referred to us orphaned. This unique strain needs further analysis using other molecular

Table 4 Bivariate and Multivariate risk factor association with genotypic patterns of *M. tuberculosis* strain isolates

Variable	Status	TOTAL	Genotypic pattern		Bivariate Analysis			Multivariate Analysis			
			Unique	Clustered	Unadjusted			Adjusted			
					OR	CI(95%)	P-Value	OR	CI (95%)	P-Value	
Age	18–24	145 (26.22)	14(2.53)	131(23.69)	Ref						
	25–30	183(33.09)	23(4.15)	160(28.93)	0.95	0.35–2.27	0.003**	0.63	0.26–2.32	0.056	
	31–36	107(19.35)	17(3.07)	90(16.27)	2.13	0.96–6.12	0.81	1.67	0.33–3.23	0.72	
	37–42	47(8.50)	3(0.54)	44(7.96)	1.93	0.41–4.61	0.94	0.72	0.21–1.89	0.65	
	43–48	35(6.32)	3(0.54)	32(5.79)	2.16	0.98–6.78	0.31	1.53	0.43–2.87	1.45	
	49–54	18(3.25)	4(0.72)	14(2.53)	1.34	0.35–3.72	0.13	1.21	0.27–2.67	1.20	
	55–60	1(0.18)	1(0.18)	11(1.99)	0.85	0.23–3.10	0.12	0.45	0.12–2.98	0.35	
	61+	6(1.08)	0(0.00)	6(1.08)	1.23	0.16–3.81	0.48	0.94	0.43–2.78	0.87	
HIV status	Positive	267(48.28)	26(4.70)	241(43.58)	1.73	0.42–5.21	0.04**	1.52	0.29–3.68	0.001**	
	Negative	286(51.71)	17(3.07)	269(48.64)	Ref						
Sex	Female	208(37.61)	13(2.35)	195(35.26)	1.67	0.23–3.02	0.27	1.32	0.23–3.78	0.97	
	Male	345(62.38)	17(3.07)	326(58.95)	Ref						
Education	No formal	144(26.04)	18(3.25)	126(22.78)	2.61	0.87–5.92	0.92	2.21	0.49–3.98	1.34	
	Primary	243(43.94)	14(2.53)	229(41.41)	Ref						
	Secondary	127(22.97)	19(3.43)	108(19.52)	2.32	0.61–5.24	0.71	2.13	0.49–3.09	1.98	
	Tertiary	39(7.05)	4(0.72)	35(6.32)	3.61	1.02–7.31	0.21	3.32	0.67–5.76	2.34	
Occupation	Employed	83(15.01)	15(2.71)	68(2.30)	1.67	0.24–3.06	0.36	1.43	0.45–2.97	1.23	
	Self employed	218(39.42)	15(2.71)	203(36.71)	Ref	Ref					
	Informal employment	128(23.15)	21(3.80)	107(19.35)	1.52	0.21–2.98	0.54	1.22	0.22–3.21	0.98	
	Unemployed	103(18.62)	13(2.35)	90(16.27)	1.43	0.19–2.261	0.25	1.31	0.57–2.61	1.13	
	Student	10(1.81)	2(0.36)	8(1.45)	0.98	0.37–3.12	0.46	0.69	0.22–2.22	0.34	
	Other	11(1.99)	1(0.18)	10(1.80)	1.73	0.31–2.98	0.41	1.57	0.23–2.01	0.87	
Monthly income	Less than Ksh. 1000	105(18.98)	10(1.80)	95(17.18)	Ref						
	Ksh. 1000–9, 999	160(28.93)	11(1.99)	149(26.94)	1.28	0.63–4.00	0.046**	1.19	0.27–3.02	0.52	
	Ksh. 10, 000–19, 999	174(31.46)	17(3.07)	157(28.75)	1.73	0.58–4.32	0.82	1.53	0.21–2.56	0.98	
	Ksh. 20,000–29, 999	48(8.68)	4(0.72)	44(7.96)	1.82	0.69–3.92	0.76	1.67	0.36–2.63	1.98	
	Ksh. 30,000–39,999	34(6.15)	2(0.36)	32(5.79)	1.23	0.42–3.62	0.34	0.98	0.32–3.21	0.71	
	Ksh. 40,000–49,999	18(3.25)	1(0.18)	17(3.07)	1.68	0.41–4.31	0.36	1.42	0.51–2.21	1.34	
	Ksh. 50,000 +	14(2.53)	2(0.36)	12(2.17)	1.53	0.29–3.93	0.54	1.37	0.31–2.87	1.43	
Size of the house	Single room	122(22.06)	17(3.07)	105(18.99)	1.43	0.46–4.31	0.049**	1.27	0.29–3.02	0.056	
	Bedsitter	11(1.99)	3(0.54)	8(1.45)	1.83	0.32–3.61	0.71	1.59	0.36–3.42	0.78	
	One bedroom	207(37.43)	21(3.80)	186(33.63)	Ref						
	Two bedroom	191(34.54)	18(3.25)	173(31.28)	1.92	0.67–5.42	0.83	1.63	0.41–3.67	0.73	
	Other	22(3.98)	2(0.36)	20(3.62)	1.86	0.20–5.31	0.47	1.53	0.37–3.45	0.56	
	Cigarette smocking	Yes	132(23.87)	7(1.26)	125(22.6)	1.81	0.43–4.82	0.41	1.51	0.46–3.32	1.21
	No	421(76.13)	14(2.53)	407(73.60)	Ref						
Alcohol use	Yes	151(27.31)	16(2.89)	135(24.41)	0.87	0.34–5.34	0.003**	0.59	0.43–3.12	0.001**	
	No	402(72.69)	11(1.99)	391(70.70)	Ref						
Cross boarder travels	Yes	211(38.15)	16(2.89)	195(35.26)	0.93	0.46–3.92	0.002**	0.61	0.49–3.87	0.026**	
	No	342(61.84)	17(3.07)	325(58.77)	Ref						

** P-value of <0.05 was considered significant. CI Confidence interval

epidemiology tools including Mycobacterium Interspersed Repetitive Units (MIRU) or Variable Number of Tandem Repeats (VNTR) to shed more insights on their epidemiology. These orphaned strains could

indicate either recent or sporadic TB infections in Western Kenya.

The detection of different lineages of *M. tuberculosis* in Western Kenya maybe an indicator of global

diversity of *M. tuberculosis* strains with different origin given that there are several counties under study that are bordering neighboring countries like Uganda and Tanzania where most of these strains has been reported [33–35]. The findings of this study is in agreement with previous studies which indicated that CAS is wide spread and predominant in Kenya, Uganda, Tanzania and Ethiopia [36] which are East African Community countries which allow for free movement of its citizen across the border.

The discovery of wide spread Beijing strain which has been reported as the most lethal driver of drug resistance and reinfection [37] could be the contributing factor to persistent of the disease in Western Kenya. The Beijing strain has been reported as the most virulent strain globally [38]. This has been because it's the most favored strain with mutations hence easily develops drug resistant [39].

On the other hand, despite Western Kenya Counties bearing the largest burden of HIV pandemic, there was no evidence that specific strains have an association of HIV even though the status is a risk factor to acquiring tuberculosis due to suppressed immune system. Despite other studies associating TB drug resistance to Beijing lineage, this study did not find any correlation between the strain and resistance. Therefore, resistance in this case could be associated to poor drug adherence, failure in prompt diagnosis, migration, poverty and increased TB/HIV cross infection [30].

Most of the *M. tuberculosis* isolates from the study population were clustered which may inform that a large proportion of were recent infection. Further, this could indicate that the clustered strain might be having a higher transmissibility of higher chances off progressing into disease stage immediately after transmission. There was also a sizable number of unique strains, indicating the main drivers of persistent tuberculosis in the population that undergoes latent stage before reactivation occurs. This could be the main public health change in terms of control and prevention of tuberculosis in the population. Latency affects early diagnosis and therefore management of tuberculosis.

Significant association in bivariate analysis especially age, monthly income and size of houses used by patients could be as a result of potential cofounders, this could further be explained in multivariate logistic regression adjusting for potential cofounders which showed that the above factors were not significant. Multivariate logistic regression analysis showed HIV status, alcohol use and cross border travels as having significant association with genetic clustering of tuberculosis. HIV virus is known to fight the immune system hence exposing populations to opportunistic diseases like tuberculosis. Compromised

immunity could accelerate disease progression to active tuberculosis immediately after infection. Alcohol is a social drink taken bay revelers who share cups, syphons and other drinking utensils. Western Kenya is one of the leading traditional, crude and other forms of alcohol users [40] further alcohol use have been reported to increase the burden of tuberculosis [41]. Local and international travels have also been reported to increase the risk of tuberculosis transmission [42]. Emigrants globally have been reported to contribute to global distribution of tuberculosis. This has resulted to genetic diversity of *M. tuberculosis* strains globally [43].

Conclusion

There exist diverse strains of *M. tuberculosis* across the 10 counties of Western Kenya. Predominant distribution of clustered genotype as opposed to unique points due to the fact that most TB cases in this region are as a result of resent transmission other than activation of latent TB.

Recommendation

The identification of diverse *Mycobacterium tuberculosis* in 10 counties of Western Kenya calls for aggressive control and prevention strategies to curb mortalities associated with pulmonary tuberculosis. Spoligotyping patterns of some clinical isolates could not be matched to others reported in international database hence referred to us orphaned need to be further characterized using other molecular epidemiology tools including Mycobacterium Interspersed Repetitive Units (MIRU) or Variable Number of Tandem Repeats (VNTR) to shed more insights on their epidemiology. These orphaned strains could indicate either recent or sporadic TB infections hence the need for aggressive intervention.

Strength and limitation of the study

The study has demonstrated the existence of genetic diversity of *M. tuberculosis* circulating in Western Kenya population. Further we have demonstrated the existence of fast progressing pathogens to full blown disease stage and those that undergo latency through M tuberculosis cluster analysis. Additionally, the study has identified potential risk factors that are likely to be the drivers of tuberculosis in Western Kenya. However, we were unable to further characterize the orphaned strains through other available genotypic characterization tools to strengthen our findings.

Acknowledgments

We thank George Kamau and Betty Khamala both from the national Tuberculosis reference laboratory (Nairobi Kenya), Moi Teaching and referral Hospital TB laboratory for field and laboratory assistance, Kenya Medical Research Institute Kisumu, Tuberculosis division and Stellenbosch University, Division of Molecular Biology and Human Genetics. This study was supported by

Africa-ai-Japan (JKU/ADM/10B). The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of Jomo Kenyatta University of Agriculture and Technology or that of the Government of Kenya.

Authors' contributions

Martin Ogwang, Lamec Diero and Lucy Mutharia conceived the study. Gabriel Magoma, Mabel Imbuga and Lucy Mutharia supervised the implementation of the project. Caroline Ngugi and Martin Ogwang participated in field sample processing. All authors contributed equally in writing the manuscript. Martin Ogwang wrote the first draft of the manuscript. Lucy Mutharia, Florence Ng'ong'a, Gabriel Magoma, Mabel Imbuga, Lamec Diero and Caroline Ngugi participated in reviewing the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by Africa-ai-Japan (JKU/ADM/10B). The funding body had no role in the design of the study, the collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The data that support the findings of this study are available from Division of Tuberculosis, Leprosy and Lung Diseases, Ministry of Health and Research Division Jomo Kenyatta University of Agriculture and Technology, Kenya but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Head of Division Tuberculosis, Leprosy and Lung Diseases Kenya P.O. Box 20781–00202 Nairobi, Kenya. Tel: +(254) 773 977 440, Jomo Kenyatta University of Agriculture and Technology P.O. Box 62000–00200, Nairobi, Telephone +254–67–5870000/1/2/3/4/5 and Corresponding author through email: macogwang2012@gmail.com.

Declarations

Ethics approval and consent to participate

The approval for the study was provided by Moi Teaching and Referral Hospital/Moi University College of Health Sciences-Institutional Research and Ethics committee (MTRH/MU-IREC) under the reference number 0001519. All participants in the study were taken through the study objectives allowed to seek clarification. It was also made clear to them that they were free to opt out at any point of the study. The informed consent was obtained from all individuals through appending their signatures or use of thumb print for those unable to write as was approved by MTRH/MU-IREC. All procedures in the manuscript were performed in accordance with relevant guidelines. Administrative permissions were acquired by our team to access the data used in this paper from the Division of Tuberculosis, Leprosy and Lung Diseases, a Division of the Ministry of Health. P.O. Box 20781–00202 Nairobi, Kenya. Tel: +(254) 773 977 440.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Public Health Nairobi Kenya, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya. ²School of Medicine, Moi University, Eldoret, Kenya. ³School of Biomedical Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya. ⁴Department of Cellular and Molecular Biology, University of Guelph, Guelph, ON, Canada.

Received: 13 May 2023 Accepted: 30 November 2023

Published online: 09 December 2023

References

- World Health Organization. WHO report on surveillance of antibiotic consumption: 2016–2018 early implementation.
- Uplekar M, Weil D, Lonroth K, Jaramillo E, Lienhardt C, Dias HM, et al. WHO's new end TB strategy. *Lancet*. 2015;385(9979):1799–801.
- Kenya National Bureau of Statistics. Economic survey. Kenya National Bureau of Statistics. 2018:2018.
- Enos M, Sitienei J, Ong'ang'o J, Mungai B, Kamene M, Wambugu J, et al. Kenya tuberculosis prevalence survey 2016: challenges and opportunities of ending TB in Kenya. *PLoS One*. 2018;13(12):1–19.
- Lead A. The global prevention and elimination of tuberculosis among health workers. 2018. Available from: <https://apha.org/policies-and-advocacy/public-health-policy-statements/policy-database/2019/01/28/global-prevention-and-elimination-of-tuberculosis>.
- Allué-Guardia A, García JI, Torrelles JB. Evolution of drug-resistant Mycobacterium tuberculosis strains and their adaptation to the human lung environment. *Front Microbiol*. 2021;12(February):1–21.
- Xi Y, Zhang W, Qiao RJ, Tang J. Risk factors for multidrug-resistant tuberculosis: a worldwide systematic review and meta-analysis. *PLoS One*. 2022;17(6):1–15. <https://doi.org/10.1371/journal.pone.0270003>.
- Huang Y, Ai L, Wang X, Sun Z, Wang F. Review and updates on the diagnosis of tuberculosis. *J Clin Med*. 2022;11(19)
- Timire C, Ngwenya M, Chirenda J, Metcalfe JZ, Kranzer K, Pedrazzoli D, et al. Catastrophic costs among tuberculosis-affected households in Zimbabwe: a national health facility-based survey. *Trop Med Int Heal*. 2021;26(10):1248–55.
- Parvaresh L, Crighton T, Martinez E, Bustamante A, Chen S, Sintchenko V. Recurrence of tuberculosis in a low-incidence setting: a retrospective cross-sectional study augmented by whole genome sequencing. *BMC Infect Dis*. 2018;18(1):4–9.
- Wollenberg K, Harris M, Gabrielian A, Ciobanu N, Chesov D, Long A, et al. A retrospective genomic analysis of drug-resistant strains of M. Tuberculosis in a high-burden setting, with an emphasis on comparative diagnostics and reactivation and reinfection status. *BMC Infect Dis*. 2020;20(1):1–12.
- Addo KK, Addo SO, Mensah GI, Mosi L, Bonsu FA. Genotyping and drug susceptibility testing of mycobacterial isolates from population-based tuberculosis prevalence survey in Ghana. *BMC Infect Dis*. 2017;17(1):1–7.
- Yin C, Mijiti X, Liu H, Wang Q, Cao B, Anwaijiang A, et al. Molecular epidemiology of clinical Mycobacterium tuberculosis isolates from southern Xinjiang, China using Spoligotyping and 15-locus MIRU-VNTR typing. *Infect Drug Resist*. 2023;16(February):1313–26.
- Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajj SA, et al. Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol*. 2006;6:1–17.
- Zeng X, Li H, Zheng R, Kurepina N, Kreiswirth BN, Zhao X, et al. Spoligotyping of mycobacterium tuberculosis complex isolates by use of ligation-based amplification and melting curve analysis. *J Clin Microbiol*. 2016;54(9):2384–7.
- Belakehal F, Barth SA, Menge C, Mossadak HT, Malek N, Moser I. Evaluation of the discriminatory power of spoligotyping and 19-locus mycobacterial interspersed repetitive unit-variable number of tandem repeat analysis (MIRU-VNTR) of Mycobacterium bovis strains isolated from cattle in Algeria. *PLoS One*. 2022;17(1):1–13.
- Van der Zanden AGM, Kremer K, Schouls LM, Caimi K, Cataldi A, Hulleman A, et al. Improvement of differentiation and interpretability of spoligotyping for Mycobacterium tuberculosis complex isolates by introduction of new spacer oligonucleotides. *J Clin Microbiol*. 2002;40(12):4628–39.
- Organization WH. WHO consolidated guidelines on tuberculosis: tuberculosis preventive treatment. World Health Organization; 2020.
- Murphy ME, Phillips PPJ, Mendel CM, Bongard E, Bateson ALC, Hunt R, et al. Spot sputum samples are at least as good as early morning samples for identifying Mycobacterium tuberculosis. *BMC Med*. 2017;15:1–10.
- Tortoli E, Cichero P, Piersimoni C, Simonetti MT, Gesu G, Nista D. Use of BACTEC MGIT 960 for recovery of mycobacteria from clinical specimens: multicenter study. *J Clin Microbiol*. 1999;37(11):3578–82.
- Ramos A, Carvalho T, Ribeiro M, Guimaraes JT. Capilia™ TB-neo assay: a new tool for rapid distinction between tuberculous and non-tuberculous mycobacteria. *Int J Tuberc Lung Dis*. 2016;20(6):753–6.

22. Brittle W, Marais BJ, Hesselring AC, Schaaf HS, Kidd M, Wasserman E, et al. Improvement in mycobacterial yield and reduced time to detection in pediatric samples by use of a nutrient broth growth supplement. *J Clin Microbiol.* 2009;47(5):1287–9.
23. Chang CL, Park TS, Oh SH, Kim HH, Lee EY, Son HC, et al. Reduction of contamination of mycobacterial growth indicator tubes with a modified antimicrobial combination. *J Clin Microbiol.* 2002;40(10):3845–7.
24. Votintseva AA, Pankhurst LJ, Anson LW, Morgan MR, Gascoyne-Binzi D, Walker TM, et al. Mycobacterial DNA extraction for whole-genome sequencing from early positive liquid (MGIT) cultures. *J Clin Microbiol.* 2015;53(4):1137–43.
25. García-Alegría AM, Anduro-Corona I, Pérez-Martínez CJ, Corella-Madueño MAG, Rascón-Durán ML, Astiazaran-García H. Quantification of DNA through the nanodrop spectrophotometer: methodological validation using standard reference material and Sprague dawley rat and human DNA. *Int J Anal Chem.* 2020;2020.
26. Vitol I, Driscoll J, Kreiswirth B, Kurepina N, Bennett KP. Identifying Mycobacterium tuberculosis complex strain families using spoligotypes. *Infect Genet Evol.* 2006;6(6):491–504.
27. Lungu P, Njelesani E, Sukwa T, Ngamika O, Munsaka S, Kilembe W, et al. Immune correlates of Mycobacterium tuberculosis patients in Zambia stratified by HIV serostatus and level of immunity—a cross-sectional analytical laboratory based study. *PLoS One.* 2022;17(1):e0262454.
28. Bershteyn A, Mutai KK, Akullian AN, Klein DJ, Jewell BL, Mwalili SM. The influence of mobility among high-risk populations on HIV transmission in Western Kenya. *Infect Dis Model.* 2018;3:97–106.
29. Otieno EJ, Okuku MT. Socio-Cultural Factors Contributing to the Spread of HIV and AIDs in Homa Bay County, Kenya. In *Catholic University of Eastern Africa and Center for Democracy Research and~...*; 2017.
30. Ogwang MO, Imbuga M, Ngugi C, Mutharia L, Magoma G, Diero L. Distribution patterns of drug resistance Mycobacterium tuberculosis among HIV negative and positive tuberculosis patients in Western Kenya. *BMC Infect Dis.* 2021;21(1):1–15. <https://doi.org/10.1186/s12879-021-06887-x>.
31. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, et al. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. *Nat Genet.* 2015;47(3):242–9.
32. Brites D, Gagneux S. The nature and evolution of genomic diversity in the Mycobacterium tuberculosis complex. *Strain Var Mycobacterium Tuberc complex its role Biol Epidemiol Control.* 2017:1–26.
33. Lukoye D, Katabazi FA, Musisi K, Kateete DP, Asiimwe BB, Okee M, et al. The T2 Mycobacterium tuberculosis genotype, predominant in Kampala, Uganda, shows negative correlation with antituberculosis drug resistance. *Antimicrob Agents Chemother.* 2014;58(7):3853–9.
34. Mfinanga SG, Warren RM, Kazwala R, Ngadaya E, Kahwa A, Kazimoto T, et al. Genetic profile of Mycobacterium tuberculosis and treatment outcomes in human pulmonary tuberculosis in Tanzania. *Tanzan J Health Res.* 2014;16(2)
35. Chihota VN, Niehaus A, Streicher EM, Wang X, Sampson SL, Mason P, et al. Geospatial distribution of Mycobacterium tuberculosis genotypes in Africa. *PLoS One.* 2018;13(8):e0200632.
36. Tulu B, Ameni G. Spoligotyping based genetic diversity of Mycobacterium tuberculosis in Ethiopia: a systematic review. *BMC Infect Dis.* 2018;18(1):1–10.
37. Cowley D, Govender D, February B, Wolfe M, Steyn L, Evans J, et al. Recent and rapid emergence of W-Beijing strains of Mycobacterium tuberculosis in Cape Town. *South Africa Clin Infect Dis.* 2008;47(10):1252–9.
38. Ates LS, Dippenaar A, Ummels R, Piersma SR, van der Woude AD, van Der Kuij K, et al. Mutations in ppe38 block PE₃_PGRS secretion and increase virulence of *Mycobacterium tuberculosis*. *Nat Microbiol.* 2018;3(2):181–8.
39. Shuaib YA, Khalil EAG, Wieler LH, Schaible UE, Bakheit MA, Mohamed-Noor SE, et al. Mycobacterium tuberculosis complex lineage 3 as causative agent of pulmonary tuberculosis, eastern Sudan. *Emerg Infect Dis.* 2020;26(3):427–36.
40. Takahashi R, Wilunda C, Magutah K, Mwaura-Tenambergen W, Wilunda B, Perngpan U. Correlates of alcohol consumption in rural western Kenya: a cross-sectional study. *BMC Psychiatry.* 2017;17:1–10.
41. Imtiazi S, Shield KD, Roerecke M, Samokhvalov AV, Lönnroth K, Rehm J. Alcohol consumption as a risk factor for tuberculosis: meta-analyses and burden of disease. *Eur Respir J.* 2017;50(1)
42. Walter KS, Tataru MB, da Silva KE, Moreira FMF, Dos Santos PCP, de Melo Ferrari DD, et al. Local and Travel-Associated Transmission of Tuberculosis at Central Western Border of Brazil 2014–2017. *Emerg Infect Dis.* 2021;27(3):905.
43. Wen Y, Xie Q, Zhang R, Zhao J, Liu X, Wu J, et al. Pulmonary tuberculosis screening in emigrants and travellers: a retrospective analysis of Guangzhou port in China. *Travel Med Infect Dis.* 2022;49:102357.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com