

Genetic profile of *Mycobacterium tuberculosis* and treatment outcomes in human pulmonary tuberculosis in Tanzania

SAYOKI G.M. MFINANGA^{1*}, ROB M. WARREN⁴, RUDOVICK KAZWALA³, AMOS KAHWA¹, THECLA KAZIMOTO¹, GODFATHER KIMARO¹, SAID MFAUME¹, TIMOTHY CHONDE², ESTHER NGADAYA¹, SAID EGWAGA³, ELIZABETH M. STREICHER⁴, GEY N.C. VAN PITTIUS⁴, ODD MORKVE⁵ & SARAH CLEAVELAND⁶

¹National Institute for Medical Research, Muhimbili Medical Research Centre, Tanzania

²Department of Medical Biochemistry, Stellenbosch University, Stellenbosch, South Africa

³Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania

⁴National Tuberculosis Control Program, Ministry of Health and Social Welfare, Dar es Salaam, Tanzania

⁵University of Bergen, Bergen, Norway

⁶College of Medicine, Veterinary Medicine and Life Sciences, University of Glasgow, United Kingdom

Abstract: Information on the different spoligotype families of *Mycobacterium tuberculosis* in Tanzania is limited, and where available, restricted to small geographical areas. This article describes the genetic profile of *M. tuberculosis* across Tanzania and suggests how spoligotype families might affect drug resistance and treatment outcomes for smear positive pulmonary tuberculosis patients in Tanzania. We conducted the study from 2006 to 2008, and the isolates were obtained from samples collected under the routine drug resistance surveillance system. The isolates were from specimens collected from 2001 to 2007, and stored at the Central and Reference Tuberculosis Laboratory. A total of 487 isolates from 23 regions in the country were spoligotyped. We were able to retrieve clinical information for 446 isolates only. Out of the 487 isolates spoligotyped, 195 (40.0%) belonged to the Central Asian (CAS) family, 84 (17.5%) to the Latin American Mediterranean (LAM) family, 49 (10.1%) to the East-African Indian (EAI) family, and 33 (6.8%) to the Beijing family. Other isolates included 1 (0.2%) for H37Rv, 10 (2.1%) for Haarlem, 4 (0.8%) for S family, 58 (11.9%) for T family and 52 (10.7%) for unclassified. No spoligotype patterns were consistent with *M. bovis*. Regarding treatment outcomes, the cure rate was 80% with no significant variation among the spoligotype families. The overall level of MDR TB was 2.5% (3/121), with no significant difference among the spoligotype families. All Beijing strains (11.8%, 30/254) originated from the Eastern and Southern zones of the country, of which 80% were from Dar es Salaam. Isolates from the CAS and T families were reported disproportionately from the Eastern-Southern zone, and EAI and LAM families from the Northern-Lake zones but the difference was not statistically significant. Five isolates were identified as non-tuberculous *Mycobacteria*. In conclusion, *M. tuberculosis* isolates from pulmonary tuberculosis cases in Tanzania were classified mostly within the CAS, LAM, and EAI and T families, while the Beijing family comprised about 7% isolates only. Consistently good treatment outcomes were recorded across these spoligotype families. The proportion of drug resistance strains was low. The findings also suggest variation of spoligotype families with varying geographical localities within the country, and identify this area for further research to confirm this finding.

Key words: Genetic profile, *Mycobacterium tuberculosis*, spoligotype, treatment, Tanzania

Introduction

Tuberculosis (TB) is a major public health problem and leading cause of human deaths worldwide (WHO, 2008). In Tanzania, the notification of TB cases has increased dramatically from about 10 000 cases in 1979 to 64,267 in 2011 (NTLP annual reports, 1980-2011). Information on the different *M. tuberculosis* spoligotype families is limited, and, where available, is restricted to small geographical areas of the country (Eldholm *et al.*, 2006; Kibiki *et al.*, 2007).

Different *M. tuberculosis* strains have distinctive epidemiological and clinical characteristics, with virulence, and clinical presentation, appearing to be strain dependent (Lopez *et al.*, 2003).

* Correspondence: Sayoki Mfinanga; E-mail: gsmfinanga@yahoo.com

Recent findings indicate that different strains of *M. tuberculosis* have adapted to specific human populations and such strains have a higher probability of transmission in these communities than others (Gagneux *et al.*, 2006; Gagneux & Small, 2007; Hanekom *et al.*, 2008). The Beijing family is thought to be hyper-virulent and to have high probability for developing drug resistance (Bifani *et al.*, 2002). Some *M. tuberculosis* strains are noted for their wide dissemination and acquisition of drug resistance (Glynn *et al.*, 2002). As a result, there has been increased recent attention paid to *M. tuberculosis* strain identification.

This study presents data on the patterns of strain variation of *M. tuberculosis* in human pulmonary cases in Tanzania, their resistance patterns and treatment outcomes, with the aim of improving our understanding of the epidemiology and management of tuberculosis, particularly with respect to monitoring molecular epidemiological trends and strengthening management and control policy for communities in different parts of the country. It is envisaged that this information will guide direction for future research and highlight recommendations for improvement or assurance of the management for tuberculosis, and thus contribute to control of TB in the country. This study also extends the baseline molecular epidemiological data for Tanzania.

Materials and Methods

Study setting and design

Source of Isolates: Isolates were obtained from the existing health facility system for routine surveillance of TB drug resistance in Tanzania. The National Tuberculosis and Leprosy Program (NTLP) runs the surveillance system in collaboration with the National Institute for Medical Research, and covers health facilities in all regions of the country. The system comprises about 700 local facilities for routine TB diagnosis using smear microscopy, with 25% of sputum samples from new smear positive pulmonary tuberculosis cases and from all re-treatment cases submitted to zonal laboratories for culture, on annual basis. To collect the 25% of the samples from all new smear-positive cases, each region is assigned four months in a year, and each diagnostic laboratory, under coordination of District Tuberculosis Coordinators (DTLCs) and Regional Tuberculosis Coordinators (RTLCS) in the region, submits sputum samples to the nearest zonal reference laboratory for culture.

The zonal reference laboratory included Central Tuberculosis Reference Laboratory (CTRL) in Dar es Salaam (serving the Eastern and Southern zone), Bugando Medical Centre (BMC) in Mwanza (serving the Lake zone), and Kilimanjaro Christian Medical Centre (KCMC) in Moshi (serving the Northern zone). The two zonal laboratories, the BMC and KCMC, send isolates to CTRL for drug susceptibility testing (DST). The CTRL also serves as a central reference laboratory for the country and is the only laboratory that is capable of performing drug susceptibility testing. The CTRL undergoes proficiency testing and external quality assessment by sending strains to a supranational laboratory at the Institute of Tropical Medicine in Belgium.

Study inclusion and exclusion criteria

The study inclusion criteria were based on old culture stocks, each correlating to individual patient and originated from specimens brought to the CTRL for routine TB drug surveillance, during the period 2001 to 2007. The exclusion criteria were based on subculture results and these included culture negative and contamination. Isolates that did not have information on drug resistance and treatment outcomes were further excluded during respective analysis.

Culture and drug resistance testing

A total of 1450 culture stocks were received from the reference laboratories and sub-cultured at CTRL. Out of the total culture stocks, 87% originated from the specimens collected from 2005-2006. A

total of 154 culture stocks were from the Northern Zone, 279 were from the Lake Zone, and 1017 were from Southern and Eastern zone. The culturing process was carried out using three slopes of Lowenstein Jensen (LJ) media, containing 0.6% pyruvate in the first slope, 0.75% glycerol in the second slope, and 500µg/ml 4 Para-nitrobenzoic acid in the third slope; and then incubated at 37°C. A culture slope was considered negative when no colonies were seen after 12 weeks. Of 1450 old culture stocks processed for sub-culturing, 582 were culture positive, 362 were negative and 506 were culture contaminated.

Drug susceptibility testing (DST) at CTRL is routinely carried out using the proportion method with the following drug concentrations: 0.2µg/ml and 1.0µg/ml for isoniazid, 40µg/ml for rifampicin, 5 µg/ml for dihydrostreptomycin sulphate and 2 µg/ml for ethambutol. The control media without drugs are prepared at the same time as the drug-containing media, following the WHO guidelines (Aziz *et al.*, 2003). Multi-drug resistance was defined according to standard criteria of resistance to at least isoniazid and rifampicin (Aziz *et al.*, 2003).

Spoligotyping and Sequencing

Heat-killed isolates were sent to Stellenbosch University, South Africa where DNA was extracted using a standardised protocol (Streicher *et al.*, 2007). Ninety-five culture isolates were unsuitable for DNA extraction and therefore they were not genotyped. Four hundred and eighty seven culture-positive isolates were genotyped, using an internationally standardised method for spoligotyping (Kamerbeck *et al.*, 1997; Demay, 2012). Isolates were classified into strain families of the *M. tuberculosis* complex according to their spoligotype signature. Ten isolates, which failed to give a spoligotype, were speciated by sequencing of the 16S rRNA gene (Harmsen *et al.*, 2003).

Clinical and Demographic data

Study forms for all culture positive cases were sent to the DTLCs to provide information on gender, DOT treatment, health facility, type of patient (new or retreatment), date of sample collection, date of completing treatment and treatment outcomes. All patients included in this study were treated under DOTS strategy.

Ethical Consideration

Medical Research Coordinating Committee of the National Institute for Medical Research provided ethical clearance and the National Tuberculosis Control Programme of the Ministry of Health and Social Welfare gave permission to conduct the study.

Data analysis

Genotypic, clinical and demographic records were entered into the study forms, double entered into EpiData database and analyzed using SPSS version 14 for Windows (SPSS Inc, Chicago, IL, USA). The variables used for treatment outcomes were as classified in the Manual of National Tuberculosis Control (NTLP Manual, 2003) as to whether a patient (i) was 'cured', (ii) 'died during treatment', (iii) was 'interrupted treatment (out of control)', or (iv) was 'transferred out' or 'failed'. The 'cure' outcome is defined as sputum smear positive patient who becomes smear negative at the end of treatment, and on at least one previous occasion. The 'failed' outcome is defined as a patient who remains or becomes smear positive again at 5 months or later. The 'died during treatment' outcome is defined as a patient who dies for any reason during TB treatment. The 'out of control' outcome is defined as a patient who interrupts treatment for two consecutive months or more. 'Transferred out' is defined as a patient who has been transferred to a treatment centre in another region or treatment centre and whose treatment result is not known.

The variable for spoligotype families was recorded into new 10 binomial variables, each variable contained two categories: one for a family of interest and another for the rest of the families regrouped into 'others' category, which was used as a reference category.

The binomial variables including age group, gender and the 10 new variables for spoligotype families were cross tabulated with categorical variables namely zones (Eastern and southern Vs. Lake and northern zone) and TB types (new and relapse cases). The odds ratios with their respective 95% confidence interval, Chi-square tests with their respective *p* values were recorded. Pearson Chi-square tests were used to measure statistical significance, and in case of differences where a *p* value was ≤ 0.05 , the difference was considered to be statistically significant.

Results

Geographic distribution of isolates

We were able to retrieve information on 446 isolates from 20 regions of the Tanzania mainland and two isolates from Zanzibar (Table 1 and Figure 1). The majority (66.8%) of isolates were from four regions: Dar es Salaam (29.8%), Kilimanjaro (18.8%), Mwanza (9.2%), and Iringa (9.0%), with most from the Eastern and Southern zone ($n=253$, 56.9%), followed by the Northern zone ($n=115$, 25.8%), and the Lake Zone ($n=77$, 17.3%). We did not observe significant differences in the zonal distribution of culture negative stocks when compared with culture positive isolates.

Table 1: Regional distributions of isolates

Zones	Regions	Isolates n(%)	Relapse n (%)*
Northern and Lake zones	Kilimanjaro	87(19.5)	32(11.1)
	Arusha	3(0.7)	2(0.7)
	Manyara	9(2.0)	9(3.1)
	Singida	5(1.1)	5(1.7)
	Tanga	2(0.4)	2(0.7)
	Mwanza	41(9.2)	41(14.2)
	Shinyanga	10(2.2)	10(3.5)
	Tabora	5(1.1)	5(1.7)
	Kagera	3(0.6)	3(1.0)
	Kigoma	15(3.4)	8(2.8)
	Mara	3(0.7)	2(0.7)
	Sub-total	193(40.9)	119(41.2)
Eastern and Southern	Dar es Salaam	137(30.7)	78(27.0)
	Coast	6(1.3)	3(1.0)
	Morogoro	2(0.4)	2(0.7)
	Zanzibar	12(2.7)	8(2.8)
	Unguja	1(0.2)	0(0.0)
	Iringa	40(9.0)	37(12.8)
	Lindi	16(3.6)	9(3.1)
	Mbeya	13(2.9)	11(3.8)
	Dodoma	15(3.4)	4(1.4)
	Ruvuma	5(1.1)	4(1.4)
	Mtwara	16(3.6)	14(4.8)
	Sub-total	263(58.9)	170(58.8)
Grand Total	446**	289	

*Relapse cases between the two zones; $\chi^2 = 2$, $p = 0.158$; OR = 0.7 (95CI = 0.5 - 1.1); **Missing = 41

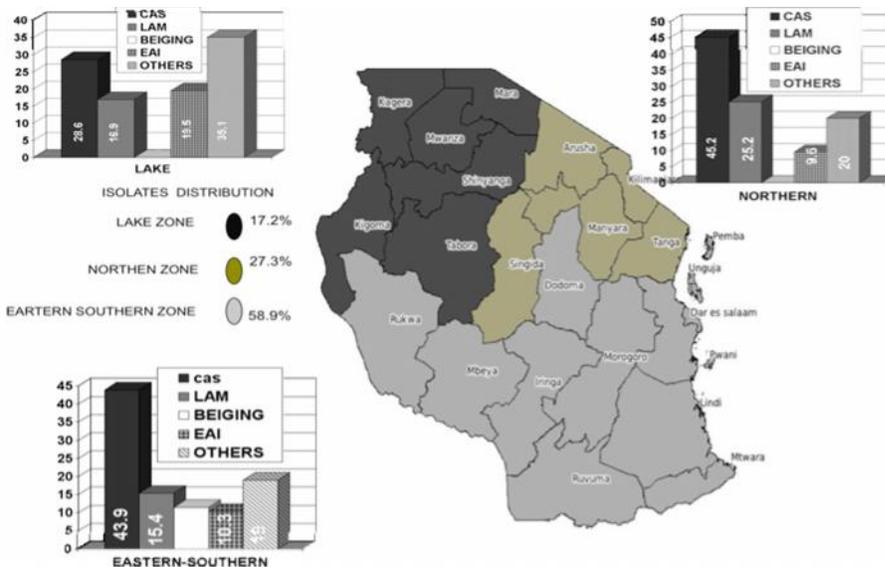


Figure 1: Zonal distribution of isolates

Spoligotype family assignment):

A total of 487 isolates were spoligotyped, of which 195 (40.0%) were classified as the Central Asian (CAS) family, 84 (17.5%) as the Latin American Mediterranean (LAM) family, 49 (10.1%) as the East-African Indian (EAI) family, 58 (11.9%) as T family and 33 (6.8%) as the Beijing family. Other spoligotypes included H37Rv (n=1, 0.2%), Haarlem (n=2, 0.45%), S family (n=21, 4.3%), and 52 (10.7%) isolates could not be assigned to a known strain family. No spoligotype patterns were consistent with *M. bovis* (Table 1).

Sequencing of isolates not classified as *M. tuberculosis*

Of 10 isolates sequenced, one was identified as *Rhodococcus sp*, two as *M. abscessus* or *M. chelonae*, two as *M. fortuitum*, one as *M. intracellulare*, one as *Nocardia farcinica*. Of these, only five isolates were identified as Non-Tuberculous Mycobacteria. Three isolates could not be identified.

Comparison of profile and spoligotype families according to zones:

The mean age of patients from the Eastern and Southern zones (mean age 35.1, SD 11.40) was significantly lower ($t = 4.99, p = 0.001$) than those from the Lake and Northern zones (mean age 42.7, SD 13.9) (Table 3). In addition, the proportion of isolates from the younger age group (9-40 years) was higher (71.8%, 148/208) in Eastern and Southern zones than in the Lake and Northern zones (48.3% 56/116), and the proportional difference was statistically significant ($\chi^2 = 17.8, p = 0.001$). Comparing the proportion of young patients, aged 9-40 years (71.0%, $n = 107$) with Dar es Salaam census data (2002 National census data for Dar es Salaam, projection for 2005) for the same age group (62.8%, $n = 2721926$), we found no statistical difference ($\chi^2 = 3.10, p = 0.078$) between the two groups. All Beijing strains (12.1%, 29/239) were isolated from patients coming from the Eastern and Southern zones, of which 80% of the strains were from Dar es Salaam only. We observed proportional difference in the distribution of the CAS and T families, between the Eastern-Southern (46.4% and 11.3% for CAS and T family respectively) and Lake-Northern zones (41.6% and 15.7% for CAS and T family respectively), but proportionately the differences were not statistically different between CAS and T family isolates for the two zones ($\chi^2: 1.0, p = 0.322$ for CAS family; and $\chi^2: 1.7; p = 0.187$ for T family). Similarly, the proportions of EAI and LAM families, reported in Northern-Lake

zones (13.6% and 24.7%, respectively) were higher as compared to the Eastern and Southern zones (10.5 and 16.3% respectively), but the statistical differences was only observed for LAM family (χ^2 : 1.3; $p= 0.254$ for EAI and χ^2 : 4.5; $p= 0.034$ for LAM).

Table 2: Mycobacterium tuberculosis spoligotyping patterns

Major family	No. of Isolates, N = 487 n (%)	Sub-family	Number of isolates n (%)
BEIJING	33 (6.8)	-	-
CAS	195 (40.0)	CAS	11 (5.6)
		CAS1_DELHI	27 (13.8)
		CAS1_KILI	152(76.0)
		CAS2	4 (2.0)
		CAS not in spolDB	12 (6.0)
EAI	49 (10.1)	EAI1_SOM	6 (12.2)
		EAI5	37 (75.5)
		EAI6 BGD1	2 (4.1)
		MANU1*	1(2.0)
		MANU2	1(2.0)
H37Rv	1(0.2)	-	-
Haarlem	10(2.1)	H1	3 (30.0)
		H3	7 (70.0)
LAM	84 (17.5)	LAM11_ZWE	53(63.1)
		LAM10 CAM	2 (2.4)
		LAM3	8 (9.5)
		LAM4	2 (2.4)
		LAM4-LAM5	2 (3.4)
		LAM9	18 (21.4)
S family	4(0.8)	-	-
T family *****	58(11.9)	T1	34 (58.6)
		T2	3 (5.2)
		T2_Uganda**	6 (10.3)
		TAN/T2_Uganda***	13 (22.4)
		T2-T3	6 (10.3)
		T3	3 (5.2)
		T3_ETH	1(1.7)
		T4_CEU1	2 (3.4)
Unclassified	52(10.7)	-	-

*MANU family is considered as a subfamily of the EAI Major family; ** T2-uganda according to SITVIT database - deletion of DVR 40 and 43 characteristic to TAN; ***TAN/T2_uganda identified in this study – characteristic deletion of DVR40 and DVR43 thus could be considered TAN/T2_uganda, but not found in SITVIT database

Comparison of spoligotype families between new and relapse/retreatment PTB

The Beijing and LAM families were more common in new smear positive cases, where we reported 14.8% and 22.7%, of 128 isolates respectively, than in retreatment cases, where were reported in 3.9%, and 16.8% of 285 isolates, respectively (Table 4). The statistical proportional difference was observed only in Beijing family with chi-square of 15.8 ($p = 0.001$), and not in LAM family ($\chi^2 = 1.5$ ($p = 0.222$)). In contrast, the CAS family was more common in retreatment (46.3, 132/285) than in new smear

positive (41.4%, 53/ 128) pulmonary tuberculosis isolates, but the proportional differences was not statistically different ($\chi^2 = 0.9$, $p = 0.353$).

Table 3: Comparison of profile and spoligotype families according to zones

Factors	Eastern and Southern Zone n (%)	Lake Zone n (%)	and Northern Zones	Statistical Tests	P-Value	OR	OR 95%CI
Gender							
Male	177 (70.0)	130 (69.5)		$\chi^2 = 0.01$	0.920	1.0	0.7 – 1.5
Female	76 (30.0)	57 (30.5)					
Means and Age Groups (Years)							
Mean age(SD, n)	35.1(11.40, 207)	42.7 (13.9, 116)		t = 4.99	0.001	-	
9 – 40 years	148 (71.8)	56 (48.3)		$\chi^2 = 17.8$	0.001	2.7	1.7 – 4.4
41 – 84 years	58 (28.2)	60 (51.7)					
Major M. tuberculosis family							
				χ^2			
BEIJING	29 (12.1)	0 (0)		23.2	0.001	-	-
CAS	111(46.4)	74 (41.6)		1.0	0.322	1.2	0.8 – 1.8
EAI	25 (10.5)	24 (13.6)		0.9	0.329	0.7	0.4 – 1.3
H37Rv	0(0)	1 (0.6)		1.3	0.246	-	-
Haarlem	6 (2.5)	4 (2.2)		0.0	0.862	1.1	0.3 – 4.0
LAM	39 (16.3)	44 (24.7)		4.5	0.034	0.6	0.4 – 0.9
S family	1 (0.4)	1 (0.6)		0.0	0.834	0.7	0.5 – 11.9
T family	27 (11.3)	28(15.7)		1.7	0.186	0.7	0.3 – 1.2
Total#	239	178					

#The overall total do not add up to 487 because of missing values

Comparison of other outcomes among spoligotype families

Overall, resistance to any drug was 15.9% (19/121), and the level of MDR TB was 2.5% (3/121). There were no statistical significant differences for the resistant isolates among the spoligotype families. Out of 228 patients who had complete information on treatment outcomes, 81% of patients were cured, 1.3% completed treatment but not cured, 6.6% interrupted treatment (out of control) and 4.4% were transferred. In addition, 5.3% of patients died during TB treatment, slightly fewer deaths than reported in the NTLR reports (6-10%) (MoH, 2001-2008). However, this difference was not significant ($\chi^2 = 2.16$, $p = 0.141$) when comparing the proportion of deaths for the year 2007 (7.9%, $n = 24724$) and 2008 (6.3%, $n = 24520$).

Table 4: Spoligotype families according to TB types

Family	New smear positive	Relapse	Chi square	P-Value	OR	OR, 95%CI
BEIJING	19(14.8)	11(3.9)	15.8	0.001	4.3	2.0 – 9.4
CAS	53(41.4)	132(46.3)	0.9	0.353	0.8	0.5 – 1.3
EAI	14(10.9)	35(12.4)	0.2	0.679	1.2	0.6 – 2.2
H37Rv	0(0.0)	1(0.4)	0.5	0.502	-	-
Haarlem	0(0.0)	9(3.2)	4.1	0.042	-	-
LAM	29(22.7)	50(17.5)	1.5	0.222	1.4	0.8 – 2.3
S family	1(0.8)	1(0.4)	0.3	0.560	2.2	0.1 – 36.0
T family	12(9.4)	44(15.4)	2.8	0.096	0.6	0.3 – 1.1
Total	128	285				

#The overall total do not add up to 487 because of missing values

Comparison of spoligotype families according to treatment outcomes, drug susceptibility testing, regions (Dar es Salaam against others), and age groups (9 – 40 against 41 – 84 years) did not reveal any statistically significant differences among the respective spoligotype families.

Discussion

The findings of this study provide an overview of the genetic profile of *M. tuberculosis* families in relation to geographic distribution, drug resistance patterns and the treatment outcomes for human pulmonary tuberculosis in Tanzania. We report a high diversity of spoligotype patterns with a predominance of the Central Asian (CAS), Latin American Mediterranean (LAM), and East-African Indian (EAI) families. Overall, this pattern is consistent with earlier studies in Dar es Salaam (Eldholm *et al.*, 2006) and in a referral hospital in northern Tanzania (Kibiki *et al.*, 2007), as well as results from Kenya (Githui *et al.*, 2004), indicating that the predominance of this group of spoligotypes occurs across a wide geographic scale. The consistency of results also indicates relative temporal stability of this pattern, as isolates from this study were collected over a much longer time period (7 years) than previous studies. The LAM family has also been reported to predominate in Zimbabwe and Zambia (Easterbrook *et al.*, 2004; Chihota *et al.*, 2007). In contrast, one study in Kampala, Uganda showed predominance of the T2 family (Asiimwe *et al.*, 2008). In this study, the T2 family is also reported, and although the prevalence recorded here (8.8%) is lower than the Kampala study (Asiimwe *et al.*, 2008), similar prevalence have been reported in Kenya (11%) (Githui *et al.*, 2004).

Within the three dominant families, slightly zonal variation was observed without any statistical significance, which is suggestive of localised epidemics. Similar trends have been observed in other African countries where local genotypes tend to form a greater proportion of the circulating strains (Kallenius *et al.*, 1999; Niobe-Eyangoh *et al.*, 2004; Easterbrook *et al.*, 2004; Chihota *et al.*, 2007). In recent findings, a study showed that different strains of *M. tuberculosis* have adapted to specific human populations and such strains have a higher probability of transmission than others (Gagneux *et al.*, 2006; Gagneux & Small, 2007; Hanekom *et al.* 2008).

The occurrence of EAI and LAM families is consistent with earlier studies in Tanzania (Kibiki *et al.*, 2007), and suggests that the TB epidemic could reflect an ancient lineage as EAI is considered to be the oldest TB lineage (Wirth *et al.*, 2008). The most common ancestor of the bacterial complex emerged some 40,000 years ago from its progenitor in East Africa; it is believed that the modern human population disseminated from the region around the same time (Wirth *et al.*, 2008).

As expected, Dar es Salaam contributed significantly to all main families we reported, reflecting high connectivity, both nationally and internationally. In particular, many people come to Dar es Salaam to seek better medical services, thus the isolates from the Dar es Salaam referral facilities are likely to represent spoligotypes from across the country.

Our study reported the Beijing family in about 7.0% of isolates, which is higher than that reported previously (Eldholm *et al.*, 2006; Kibiki *et al.*, 2007). The contribution of the Beijing family reported varies widely across East African sites, from 1.2% to 8.3% (Asiimwe *et al.*, 2008; Githui *et al.*, 2004; Glynn *et al.*, 2005), but the family is dominant in most Asian countries, and South Africa (van Soolingen *et al.*, 1995; Glynn *et al.*, 2005). From the fact that all patients from whom the Beijing family were isolated, originated from eastern and southern Tanzania, this may reflect the influence of increasing international trade in the area, especially in recent years China and India both have a long history of trade and investment with modern-day Africa, particularly in East Africa, where there are significant expatriate Asian communities. However, the exact reasons and evidence are uncertain, and this warrants further studies such as monitoring trends of Beijing family and its transmission dynamics, using genetic markers.

The CAS family strains, which were reported more frequently in eastern and southern Tanzania, are predominant in Central and Middle Eastern Asia (Filliol *et al.*, 2002). This is likely to reflect the influence of Asian traders in this part of Tanzania over many centuries (Broadman *et al.*, 2007) as well as the fact that the zone is the driver of TB epidemic in the country. Even though not statistical difference was observed but reason for more frequently reported isolates from relapse cases in eastern and southern zone including Dar es Salaam, Iringa, Mwanza and Kilimanjaro regions could be due to high burden of TB cases reported in the zone and the mentioned regions than in rest part of the country (NTLP annual reports, 2005-2007). The Beijing and LAM families were more often reported in new smear positive than in relapse TB cases. This could indicate recent introduction and transmission, while predominance of the CAS family in retreatment cases could be due to its long-term establishment in the epidemic.

Regarding treatment outcomes, the proportion of patients who died during treatment, and the proportion that were cured, is consistent with proportions reported in NTLP reports (MoH, 2005-2007). It is worth noting that the vast majority of isolates, including those from CAS and Beijing families, and from re-treatment cases, had favourable treatment outcomes. The finding of two out of three isolates of MDR TB in the CAS family calls for further studies on a possible association between MDR TB and the CAS family, particularly given that MDR TB isolates were also reported in the CAS family in Kenya (Githui *et al.*, 2004). However, it should be noted that the level of MDR TB in our study was very low (2.5%). All three MDR TB isolates were from retreatment cases. Overall, the level of MDR TB in our study does not differ much from the national levels, established by routine surveillance (1.0% in new and 2.2% in retreatment cases) and by recent country wide drug resistance survey (1.2% of MDR TB in new cases and 3.1% in re-treatment cases), which are low levels of drug resistance (MoH, 2007; Chonde *et al.*, 2010).

The Beijing family is thought to be hyper-virulent and to have a high probability of developing drug resistance (Bifani *et al.*, 2002). However, we failed to demonstrate any of these characteristics among the Beijing strains identified in Tanzania. Similar findings of relatively low drug resistance were reported for the Beijing genotype strains identified in northern Tanzania (Kibiki *et al.*, 2007), Uganda, Malawi and Indonesia (van Crevel *et al.*, 2001; Glynn *et al.*, 2005; Asiimwe *et al.*, 2008). These findings suggest that the spoligotype families in Tanzania might not affect treatment outcomes and drug resistance, and hence provide no evidence for changing current TB treatment protocols.

The finding of isolates that were not classified as *M. tuberculosis* complex raises concern especially when considering that these specimens were from pulmonary disease, in HIV endemic country. Many questions remain to be addressed about the impact of these isolates and consequences for treatment. The treatment outcomes for these isolates were not available due to the limitations of a retrospective design and ineffective record keeping. Furthermore, potential zoonotic transmission need to be considered since there is evidence that the same species of non-tuberculous mycobacteria have been isolated from human lymph nodes in extra pulmonary TB cases (Mfinanga *et al.*, 2004), milk and tissue specimens taken from cattle and wild animals (DFID, 2002; Kazwala *et al.*, 1998). More studies are needed to examine the impact and implications of non-tuberculous mycobacteria, especially in HIV-positive patients.

The geographic representation of samples from this study is somewhat consistent with current patterns of TB notification in the country, where the highest incidence of diseases occurs in Eastern and Southern Tanzania (with higher contribution of almost one quarter from Dar es Salaam), 7% from Mwanza and 7% from Iringa (MoH, 2007). Results from this study are therefore likely to be broadly representative of TB cases in Tanzania. In support of this, our findings compare well with national data on regional and national proportions of TB case notification, treatment outcomes, and drug resistance levels. However, caution needs to be taken in extrapolating findings to the general population in the country, as this was not a random community population-based study, but an

analysis of samples collected through the routine drug resistance surveillance system. As such, biases could invariably arise through variation among clinics in how samples are collected and cultured. For example, the high proportion of culture negative and contaminated samples might explain the absence of *M. bovis* in our study findings. Previous studies reporting *M. bovis* as a cause of human extra-pulmonary tuberculosis have focussed on rural populations (Kazwala *et al.*, 2001; Mfinanga *et al.*, 2004), whereas this study is likely to be biased towards urban communities with greater access to diagnostic facilities. A targeted study addressing the question of *M. bovis* as a cause of human pulmonary TB, particularly in rural communities, is still needed.

In conclusion, a diverse range of spoligotype families was found from all zones of the country. The most common families were Central Asian (CAS) (40.9%), the Latin American Mediterranean (LAM) (18%), and the East-African Indian families (EAI) (11.5%), with the Beijing family found in about 7% of the isolates. Dar es Salaam contributed significantly to all main reported families. All patients infected with *M. tuberculosis* from the Beijing family were from eastern and southern Tanzania with 80% of these cases reported from Dar es Salaam. There was no evidence that drug resistance and treatment outcomes were associated with spoligotype families.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SGM: participated in conception and design, acquisition of data, analysis and interpretation of data, drafting and revising the manuscript, and in final approval of the version to be published. RK, OM and SC: participated in conception and design, interpretation of data, drafting and revising the manuscript, final approval of the version to be published. SE, AK, GK, EN: participated in acquisition of data, interpretation of data, drafting and revising the manuscript, final approval of the version to be published. EMS, TC, and TK, SM: acquisition of data, culturing, drug susceptibility testing and DNA separation, reviewing the manuscript, and approval of final the version. RMW, GNCP: participated molecular testing, interpretation of finding, revising and approval of final version of the manuscript.

Acknowledgements

We are grateful to the Department for International Development, Animal Health Programme, UK for funding the project. We are also grateful to Wellcome Trust for supporting some of authors through Afrique one consortium during preparation of this manuscript. We would like to honour all study participants as the source of the specimens that are now contributing to the knowledge of the world on tuberculosis. We would like to express our gratitude to the Central Unit of NTLP, DTLCs, RTLCs and all other laboratory technicians at district, regional, Zonal and Central Tuberculosis Reference laboratories for their contribution in collection and processing of the specimens for the isolates we spoligotyped.

References

Aziz, M.A., Laszlo, A., Raviglione, M., Rieder, H., Espinal, M. & Wright A. (2003) Guidelines for surveillance of drug resistance in tuberculosis. 2nd ed. World Health Organization 2003. WHO/CDS/CSR/RMD/2003.3(<http://whqlibdoc.who.int/publications/2003/9241546336.pdf>)

- Bifani, P.J., Mathema, B., Kurepina, N.E., & Kreiswirth, B.N. (2002) Global dissemination of the Mycobacterium tuberculosis W-Beijing family strains. *Trends in Microbiology* 10, 45–52.
- Chihota, V., Apers, L., Mungofa, S., Kasongo, W., Nyoni, I.M., Tembwe, R., Mbulo, G., Tembo, M., Streicher, E.M., Spuy, G.D., van der, Victor, T.C., van Helden P., & Warren, R.M. (2007) Predominance of a single genotype of Mycobacterium tuberculosis in regions of Southern Africa. *International Journal of Tuberculosis and Lung Disease* 11, 311-318.
- Chonde, T., M., Basra, D., Mfinanga, S. G. M., Range N., Lwilla F., Shirima R. P., van Deun A., Zignol, M., Cobelens F. G., Egwaga S. M., & van Leth, F. (2010) National anti-tuberculosis drug resistance study in Tanzania. *International Journal of Tuberculosis and Lung Disease* 14, 967–972.
- van Crevel, R., Nelwan, R.H., de Lenne, W., Veeraragu, Y., van der Zanden, A.G., Amin, Z., van der Meer, J.W. & van Soolingen, D. (2001) Mycobacterium tuberculosis Beijing genotype associate with febrile response to treatment. *Emerging Infectious Disease* 7, 880-883.
- Demay, C., Liens, B., Burguière, T., Hill, V., Couvin, D., Millet, J., Mokrousov, I., Sola, C., Zozio, T. & Rastogi, N. (2012) SITVITWEB—a publicly available international multimarker database for studying Mycobacterium tuberculosis genetic diversity and molecular epidemiology. *Infection, Genetics and Evolution* 12, 755-766.
- Easterbrook, P.J., Gibson A., Murad, S., Lamprecht, D., Ives, N., Ferguson, A., Lowe, O., Mason, P., Ndudzo, A., Taziwa, A., Makombe, R., Mbengeranwa, L., Sola, C., Rastogi, N. & Drobniowski, F. (2004) High rates of clustering of strains causing tuberculosis in Harare, Zimbabwe: a molecular epidemiological study. *Journal of Clinical Microbiology* 42, 4536-4544.
- Eldholm, V., Matee, M., Mfinanga, S.G., Heun, M., & Dahle, U.R. (2006) A first insight into the genetic diversity of Mycobacterium tuberculosis in Dar es Salaam, Tanzania, assessed by spoligotyping. *BMC Microbiology* 6:76.
- Filliol, I., Driscoll, J.R. & van Soolingen, D. (2002) Global distribution of Mycobacterium tuberculosis spoligotypes. *Emerging Infectious Disease* 8, 1347–1349.
- Gagneux, S. & Small, P.M. (2007) Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. *Lancet Infect Dis*, 7(5): 328-337.
- Gagneux, S., De Riemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C. & Small, P.M. (2006) Variable host-pathogen compatibility in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2869-2873.
- Githui, W.A., Jordaan, A.M., Juma, E.S., Kinyanjui, P., Karimi, F.G., Kimwomi, J., Meme, H., Mumbi, P., Streicher, E.M., Warren, R., Meme, H., Mumbi, P., van Helden P.D., & Victor, T.C. (2004) Identification of MDR-TB Beijing/W and other Mycobacterium tuberculosis genotypes in Nairobi, Kenya. *International Journal of Tuberculosis and Lung Disease* 8, 352-360.
- Glynn, J. R., J. Whiteley, P. J., Bifani, K., Kremer, D. & van Soolingen, D. (2002) Worldwide occurrence of Beijing/W strains of Mycobacterium tuberculosis: a systematic review. *Emerging Infectious Disease* 8, 843–849.
- Glynn, J.R., Crampin, A.C., Traore, H., Yates, M.D., Mwaungulu, F.D., Ngwira, B.M., Chaguluka, S.D., Mwafulirwa, D.T., Floyd, S., Murphy, C., Drobniowski, F.A. & Fine, P.E. (2005) Mycobacterium tuberculosis Beijing genotype, northern Malawi. *Emerging Infectious Disease* 11, 150-153.
- Hanekom, M., van der Spuy, G.D., Gey van Pittius, N. C., McEvoy, C. R. E., Hoek K. G. P., Ndabambi, S. L., Jordaan, A. M., Victor, T. C., van Helden P. D., & Warren R. M. (2008) Discordance between Mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS 6110 restriction fragment length polymorphism genotyping for analysis of Mycobacterium tuberculosis Beijing strains in a setting of high incidence of tuberculosis. *Journal of Clinical Microbiology* 46, 3338.

- Harmsen, D., Dostal, S. & Roth, A. (2003) RIDOM: comprehensive and public sequence database for identification of Mycobacterium species. *BMC Infectious Diseases* 3: 26.
- Kallenius, G., Koivula, T., Ghebremichael, S., Hoffner, S.E., Norberg, R., Svensson, E., Dias, F., Marklund, B.I., & Svensson, S.B. (1999) Evolution and clonal traits of Mycobacterium tuberculosis complex in Guinea-Bissau. *Journal of Clinical Microbiology* 37, 3872-3878.
- Kazwala, R.R., Daborn, C.J., Kusiluka, L.J., Jiwa, S.F., Sharp, J.M. & Kambarage, D.M. (1998) Isolation of Mycobacterium species from raw milk of pastoral cattle of the Southern Highlands of Tanzania. *Tropical Animal Health and Production* 30, 233-239.
- Kazwala, R. R., Daborn, C. J., Sharp, J. M., Kambarage, D. M., Jiwa, S. F., & Mbembati, N. A. (2001) Isolation of Mycobacterium bovis from human cases of cervical adenitis in Tanzania: a cause for concern? *International Journal of Tuberculosis and Lung Disease* 5, 87-91.
- Kibiki, G.S., Mulder, B., Dolmans, W.M., de Beer, J.L., Boeree, M., Sam, N., van Soolingen, D., Sola, C. & Zanden, A.G. (2007) M. tuberculosis genotypic diversity and drug susceptibility pattern in HIV infected and non-HIV-infected patients in northern Tanzania. *BMC Microbiology* 7:51.
- Lopez, B., D., Aguila, H., Orozco, M., Burger, C., Espitia, V., Ritacco, L., Barrera, K., Kremer, R., Hernandez-Pando, K., Huygen, D., & Soolingen, D. (2003) A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes. *Clinical & Experimental Immunology* 133, 30-37.
- Mfinanga, S., Morkve, O., Kazwala, R., Cleaveland, S., Sharp, M., Kunda, J., & Nilsen, R. (2004). Mycobacterial adenitis: role of Mycobacterium bovis, non-tuberculous mycobacteria, HIV infection, and risk factors in Arusha, Tanzania. *East African Medical Journal* 81, 171-178.
- MoH (2003) Ministry of Health, Department of Preventive Medicine, National Tuberculosis, and Leprosy Control Programme, Tanzania, The Fourth Edition Manual.
- MoH (1980-2011) *National Tuberculosis, and Leprosy Control Programme, Tanzania, Annual Report, 1980 - 2011*. Ministry of Health and Social Welfare, Dar es Salaam, Tanzania
- Niobe-Eyangoh, S.N., Kuaban, C., Sorlin, P., Thonnon, J., Vincent, V., & Gutierrez, M.C. (2004) Molecular characteristics of strains of the Cameroon family, the major group of Mycobacterium tuberculosis in a country with a high prevalence of tuberculosis. *Journal of Clinical Microbiology* 42, 5029-5035.
- Soolingen, D., Qian, L., Haas, P.E., Douglas, J.T., Traore, H., Portaels, F., Qing, H.Z., Enkhsaikan, D., Nymadawa, P. & Embden, J.D. (1995) Predominance of a single genotype of Mycobacterium tuberculosis in countries of east Asia. *Journal of Clinical Microbiology* 33, 3234-3238.
- Streicher, E.M., Victor, T.C., van der, G., Sola, C., Rastogi, N., van Helden, P. D. & Warren, R. M. (2007) Spoligotype signatures in the Mycobacterium tuberculosis complex. *Journal of Clinical Microbiology* 45, 237-240.
- WHO (2008) Global Tuberculosis Control, Surveillance, Planning, Financing; WHO Report/HTM/TB/2008.393. World Health Organization, Geneva, Switzerland.