Snapshot of the genetic diversity of Mycobacterium tuberculosis isolates in Iraq

Mohanad Mohsin Ahmed, Suhad Hadi Mohammed, Hasan A. Abood Nasurallah, Mousa M Ali, David Couvin, Nalin Rastogi

Available at www.sciencedirect.com

ScienceDirect

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

Snapshot of the genetic diversity of *Mycobacterium tuberculosis* isolates in Iraq

Mohanad Mohsin Ahmed [a,*], *Suhad Hadi Mohammed* [b], *Hasan A. Abood Nasurallah* [c], *Mousa M. Ali* [d], *David Couvin* [e], *Nalin Rastogi* [e]

[a] *Department of Microbiology, College of Medicine, University of Kerbala, Iraq*
[b] *Department of Clinical Laboratory, College of Applied Medical Sciences, University of Kerbala, Iraq*
[c] *Department of Internal Medicine, College of Medicine, University of Kerbala, Iraq*
[d] *Medical Research Unit, College of Medicine, University of Kerbala, Iraq*
[e] *WHO Supranational TB Reference Laboratory, TB and Mycobacteria Unit, Institut Pasteur de la Guadeloupe, Guadeloupe, France*

ARTICLE INFO

Article history:
Received 14 July 2014
Accepted 17 July 2014
Available online 10 August 2014

Keywords:
TB
Genetic diversity
Spoligotyping
MIRU-VNTR

ABSTRACT

This study explored the genetic diversity of *Mycobacterium tuberculosis* isolates in Iraq by spoligotyping and 15-locus-based mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR) methods. Initially, 270 isolates from 134 patients were collected and then 134 non-duplicating isolates (1 isolate/patient) were subjected to the study analyses, 70 isolates were found to be multidrug resistant (MDR) upon testing by proportion method on Löwenstein–Jensen medium.

Spoligotyping yielded 39 patterns; 111/134 (82.2%) isolates being grouped in 16 clusters vs. 23/134 (17.2%) isolates being unique. SIT1144/T1 represented the largest cluster ($n = 20$, 14.9%), followed by SIT25/CAS1_Delhi ($n = 19$, 14.2), SIT22/CAS1_Delhi ($n = 12$, 9%); the other clusters ranged from 2 to 8 isolates. The SIT1144 is not reported in neighboring countries and only 4 isolates were reported worldwide (2 in USA, 1 in Venezuela, and 1 in Greece). This study reported 4 isolates belonging to SIT41/Turkey family, and thus it seems that this family is not exclusive to Turkey as previously thought. CAS lineage was predominant in this study (42.5%), followed by ill-defined T (29.9%).

Highly diverse MIRU-VNTR genotypes were displayed; 100 distinct MIRU-VNTR genotypes were detected (8 clusters with 2–8 strains/cluster and 92 unique). The clustering rate was 18.03%. The discriminatory efficiency of MIRU-VNTR was high (Hunter-Gaston discriminatory index [HGDI] = 0.992); it was higher than that of spoligotyping (HGDI; 0.930). However, the highest discriminatory power was provided by spoligotyping and MIRUs together. Owing to the low clustering rate by MIRU-VNTR, these results suggest that drug-resistance TB in Iraq is due to acquired resistance as opposed to transmission.

Conclusion: Iraq is specific in having its own most predominant lineage (SIT1144/T1) which is not found among neighboring countries. The 15-locus MIRU-VNTR can be useful in discriminating *M. tuberculosis* isolates in Iraq.

* Corresponding author.
E-mail address: dr.mohma.med.school@gmail.com (M.M. Ahmed).

http://dx.doi.org/10.1016/j.ijmyco.2014.07.006
2212-5531/

Introduction

Infection with *Mycobacterium tuberculosis* is responsible for extensive morbidity and mortality worldwide with approximately 8.8 million new cases and 1.1 million deaths reported in the year 2010 [1]. The importance of TB as a major public health problem has dramatically escalated with the emergence of the multi-drug resistant tuberculosis (MDR-TB), defined as a combined resistance to Rifampicin and Isoniazid [2].

Iraq is one of the countries with the highest incidence among the Eastern Mediterranean region with TB accounting for 56/100,000 population, and the estimated MDR-TB cases among new pulmonary TB cases notified in 2010 is 210 (50–380), whereas the estimated number among retreated pulmonary cases is 160 (57–260) [3]. These data are being extracted from the cases registration system that is based on clinical findings and microscopy. In the last few decades, Iraq has witnessed political instabilities and consecutive wars that left their marks on the health system. Unfortunately, there is no reliable study yet conducted to figure out the impact of those conditions on TB epidemiology in this country. Although genotyping of *M. tuberculosis* is currently appreciated as a significant complement for TB control programs [4–8], very limited data are available on the genotypes of MTB strains circulating in Iraq [9]. In contrast, in the last few years, several studies were published describing the genotypes of MTB strains isolated in neighboring countries such as Turkey, Iran and Saudi Arabia [10–15].

This study sought to conduct a molecular study to explore the genetic diversity and *M. tuberculosis* isolates in Iraq. Its aim was to shed some light on the population structure of MTB strains in Iraq and its comparison with that in the neighboring countries. It is believed that conducting such a study would generate a sort of data that may eventually lead to a better understanding of tuberculosis epidemiology in Iraq, especially the transmission dynamic of this microorganism. Although *IS6110*-RFLP has been considered the golden standard for studying the molecular epidemiology of *M. tuberculosis*, its use has been limited [16]. Alternatively, use of PCR-based genotyping methods like spoligotyping and Mycobacterial Interspersed Repetitive Units-Variable Tandem Repeats (MIRU-VNTR) have gained wide appreciation due to simplicity, rapid availability of results, reproducibility and possibility of digitizing data and exchange for inter-laboratories communications and comparisons [17].

In this study it is believed that spoligotyping and MIRU-VNTR genotyping methods present the largest amount of molecular epidemiology data on *M. tuberculosis* from Iraq to date.

Materials and methods

Ethics statement

This study was approved by the Ethics Committee of the Iraqi Ministry of Health and was performed in accordance with all national regulations. Nonetheless, since the mycobacterial isolates were collected from patients' routine samples, this study was considered a laboratory study and ethics approval was not required. Although it was a retrospective study, anonymity of the patients were maintained through using a special coding system based on numbers to ensure that this study would have no chance to affect the patient's welfare.

Mycobacterial clinical isolates and DNA extracts

During the period from January 2011 to July 2012, initially all culture positive *M. tuberculosis* isolates (more than 270 isolates) were collected and isolated from Iraqi patients with active pulmonary tuberculosis that attended the National Reference Laboratory of the National Center of Tuberculosis and chest illnesses (NTP – Iraq). Because of that this laboratory is the only lab across Iraq accredited with performing culture and drug susceptibility testing; there was a referral bias towards treatment failure. The isolates were identified as *M. tuberculosis* according to standard phenotypic criteria. Testing of drug susceptibility of these isolates to four first-line anti-TB drugs (Rifampin, Isoniazid, Streptomycin, and Ethambutol) was performed by agar proportional method on Lowenstein–Jensen media as described elsewhere (WHO, 2008). Among the 270 isolates, there were multiple isolates from the same patients, recovered on different visits to the laboratory. As all of the multiple isolates/patients gave the same profile on Spoligotyping and MIRU-VNTR, one isolate per patient was selected. Thus, 134 non-duplicated *M. tuberculosis* culture isolates (1 isolate/patient) were obtained that were subjected for the study analyses later on. Mycobacterial genomic DNA was extracted from cultured cells as described previously (13, 14, 15).

Genotyping of the isolates

The isolates were characterized by two genotyping methods, spoligotyping and MIRU-VNTR. Spoligotyping was performed as previously described [18]. MIRU-VNTR genotyping was performed by PCR-amplification of a panel of 15 MTB MIRU loci using primers described in the MIRU-VNTR standard protocol [17]. *Mycobacterium bovis* P3 and H37Rv were used as positive controls and autoclaved with double distilled water as a negative control. The results were compared with updates of the international spoligotype database of the Pasteur Institute of Guadeloupe which provides information on the shared-type distribution of *M. tuberculosis* spoligotypes at the worldwide level [18]. At the time of the comparison, the updated SITVIT2 version contained more than 90,000 patterns from more than 160 countries of patient origin.

MIRU-VNTR profile matching a pre-existing profile was classified as MIRU-VNTR international type (MIT). Newly created MIT results either after matching of an orphan profile in the study with another orphan profile present in the database or when two or more isolates within the study had the same new MIRU-VNTR profile. Additionally, MIRU-VNTR clusters arise when two or more isolates within the study had an identical profile.

Phylogenetic tree and data analysis

Data obtained from spoligotyping and MIRU-VNTR were used to construct a minimum spanning trees (MSTs) with the BioNumerics software version 3.5 (Applied Maths, Sint-Martens-Latem, Belgium), and Spoligoforests with SpolTools software (available at: http://www.emi.unsw.edu.au/spolTools). Spoligoforests were drawn as directed and not necessarily connected graphs using a hierarchical layout and a Fruchterman-Reingold algorithm to determine the parent to descendant links between the spoligotype patterns. Contrarily to the Spoligoforests, MSTs are connected undirected graphs which link all the strains together with the fewest possible linkages between nearest neighbors. MSTs have been constructed with spoligotype patterns as well as MIRU-VNTR patterns.

The discriminatory power and allelic diversity were calculated using h value and Hunter Gaston Discriminatory Index (HGDI) using the equations: $h = 1 - \Sigma xi^2$ where xi is the frequency of the ith allele at each locus [19,20].

$\text{HGDI} = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} n_j(n_j - 1)$ where N: is the total number of isolates, s: is the total number of different patterns, and nj: is the number of isolates belonging to the jth pattern [21].

The mean of allelic diversity of the MIRU-VNTR method was calculated from the formula: $H = \Sigma h/n$ where the h is the allelic diversity and n: is the number of loci [22].

Clustering rate was calculated using the formula $(n_c - c)/n$ where the n_c: is the total number of clustered cases, c: is the number of clusters, n: is the total number of cases in the study [23].

MIRU-VNTR profile of MDR, CAS and ill-defined T lineages were further analyzed separately using non-parametric Mann–Whitney test [24].

Results

Patients

Spoligotypes and the 15-locus MIRU-VNTR genotypes were fully conserved among the multiple isolates per single patient. In contrast, different patients displayed diverse spoligotypes and MIRU-VNTR genotypes. Therefore, in case of availability of multiple isolates from a single patient, only one isolate per patient was selected and considered in the study analyses. Thus, a total of 134 non-duplicating *M. tuberculosis* isolates from conveniently-selected different patients were included; 77 patients (57.9%) were resident in Baghdad, while the rest (42.1%) were living in different provinces and regions of Iraq. More males (91) than females (43) were seen in this study (male to female ratio was 2.1 to 1). The mean age of patients was 36.29 years. Nevertheless, males were slightly older than females (mean age of males and females were 38.89 and 31.55 years, respectively). Phenotypic Drug susceptibility testing (DST) based on agar proportional method detected multi-drug resistance in 70 (52.2%) cases. A significant association was found between gender ($P = 0.02$) and drug resistance, females being associated with MDR. Because the isolates were collected from a reference laboratory with a relatively higher proportion of previously treated patients (chronic TB cases), sampling bias cannot be ruled out, and this data may overestimate the percentage of MDR-TB in Iraq.

Spoligotyping analysis

Reproducible spoligotype profiles were obtained from all the 134 isolates. The spoligotype profiles obtained were compared using the SITVIT2 database, and the strain lineages were identified. In this database, SIT (shared international type) designate spoligotype patterns shared by two or more patients' isolates, whereas "orphan" designates patterns reported for a single isolate.

The description of spoligotype patterns observed among the 134 isolates, corresponding Spoligotype International Type (SIT) designations, genotypic lineages, and their worldwide distribution in the SITVIT2 database is summarized in Table 1. Spoligotyping yielded 39 patterns; 111/134 (82.2%) were grouped in 16 clusters vs. 23/134 (17.2%) isolates that were unique. Comparing results with the international database categorized 128 (95.5%) isolates in 34 shared types (33 were preexisting and one was newly created after matching with an orphan strain from Iran) and 5 isolates as orphans; 82.8% of the strains were grouped into 16 SITs: SIT1144/T1 represented the largest cluster ($n = 20$, 14.9%), followed by SIT25/CAS1_Delhi ($n = 19$, 14.2), SIT22/CAS1_Delhi ($n = 12$, 9%), SIT26/CAS1_Delhi ($n = 8$, 6%), SIT284/T1 ($n = 7$, 5.2%), 3 clusters each contained 6 strains (4.5%); SIT1198/CAS1_Delhi, SIT50/H3, SIT54/MANU2, one cluster contained 5 strains (3.7%); (SIT247/CAS1_Delhi, 4 clusters each contained 4 strains (3%); SIT53/T1, SIT47/H1, SIT127/H4(Ural-2), SIT41/Turkey, and 2 clusters each contained 2 strains (1.5%); SIT29/LAM and SIT 3371/CAS1_Delhi. Around 60%, 83%, and 75% of strains belong to SIT1144, SIT25, SIT53, respectively, and were from young patients (below 35 years). No association was found between gender and shared types or clusters.

Comparing the distribution of the spoligotype patterns reported in this study with strain isolates in neighboring countries revealed several notable findings (Supplement Table S2); SIT1144/T1, the predominant Spoligotype in Iraq ($n = 20$, 14.93% of all cases) is not recorded in any of the neighboring countries, and it showed the following distribution pattern in SITVIT2 database ($n = 24$ strains in total: Iraq $n = 20$, USA $n = 2$, Venezuela $n = 1$, Greece $n = 1$), the distribution proportion of SIT127/(H4/Ural-2) sublineage is 17% in Iran, 3% in Iraq, and 0.7% in Saudi Arabia (0.66%), thus, SIT127 may have dispersed first in Iran, followed by Iraq, and Saudi Arabia. Considering the total number of SIT284 strains in SITVIT2 database ($n = 167$), the global proportion of those strains is more predominant in Bulgaria (23.95%), followed by Turkey (20.96%) and Saudi Arabia (13.77%). However, this new Iraqi study revealed that, considering the proportion of SIT number per country, SIT284 is present in Iraq in a greater proportion than its neighboring countries. In addition, this study showed that there are other similar lineages in circulation among Iraq, Iran, Saudi Arabia and Turkey, such as CAS1_delhi.

Distribution of spoligotype lineages and sublineages in this study is summarized in Table 2. It is noteworthy that 92.6% of all *M. tuberculosis* were attributed to 4 lineages,

Table 1 – The description of spoligotype patterns observed among the 134 isolates, corresponding Spoligotype International Type (SIT) designations, genotypic lineages, and their worldwide distribution in the SITVIT2 database.

SIT*	Spoligotype Description	Octal Number	Clade**	Number (%) in study	% in study vs. database	Clustered vs. Unique patterns***
22	■■■□□□□■■■■■■■■■■■□□□□□□□□□□□□□□□■■■■■■■■	703777400001771	CAS1-Delhi	12 (8.96)	17.14	Clustered
25	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□■■□□■■■■■	703777740003171	CAS1-Delhi	19 (14.18)	3.49	Clustered
26	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□■■■■■■■■■	703777740003771	CAS1-Delhi	8 (5.97)	0.6	Clustered
29	■■■■■□□□□□□□□□□□□■■□□□□□□□□□□□□□□□□□□■■■■■	760001400000171	LAM	2 (1.49)	4.65	Clustered
34	■■■■■■■■■□□■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■	776377777760771	S	1 (0.75)	0.12	Unique
36	■■■■■■■■■■■■□■■■■■■■■■■■■■■■□■□□□□■■■■■■■	777737777720771	H3	1 (0.75)	0.78	Unique
37	■■■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■■■■■■	777737777760771	T3	2 (1.49)	0.44	Clustered
41	■■■■■■■■■■■■■■■■■■■□□□□□■□□■■■■■□□□□■■■■■■■■	777777404760771	Turkey	4 (2.99)	0.84	Clustered
46	■■■■■■■■■■■■■■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□	777777770000000	Unknown	1 (0.75)	0.5	Unique
47	■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■■■■■■	777777774020771	H1	4 (2.99)	0.27	Clustered
49	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■□■■■	777777777720731	H3	1 (0.75)	0.58	Unique
50	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■■	777777777720771	H3	6 (4.48)	0.18	Clustered
53	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■	777777777760771	T1	4 (2.99)	0.07	Clustered
54	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□■■■■■■■■■	777777777763771	Manu2	6 (4.48)	2.64	Clustered
62	■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□□■■■□■■■	777777774020731	H1	1 (0.75)	0.18	Unique
64	■■■■■■■■■■■■■■■■■■■■■□□□□■■■■□■■■□□□□■■■■■■■■	777777607560771	LAM6	1 (0.75)	0.28	Unique
86	■■■■■■■■■■■■■■■■■■■■□■■■■■■■■■■□□□□■■■■■■■■	777777737760771	T1	1 (0.75)	1.09	Unique
102	■■■■■■■■■■■■□□□□■■■■■■■■■■■■■■■■□□□□■■■■■■■■	777703777760771	T	1 (0.75)	1.39	Unique
127	■□■■■■■■■■■■■■■■■■■■■■■■■■■□□□■□□□□■■■■■■■■	577777777420771	H4 (Ural-2)	4 (2.99)	2.13	Clustered
247	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□■■■□□■■■■	703777740003471	CAS1-Delhi	5 (3.73)	14.71	Clustered
284	□□□□■■■■■■■■□□■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■	037637777760771	T1	7 (5.22)	4.19	Clustered
516	■■■■■■■■■■■■■■■■■■■■■■■■■■■□■■■■■□□□□■■■■■■■■	777777777360771	T1	1 (0.75)	5.88	Unique
520	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■□■■■■■	777777777760571	T1	1 (0.75)	5.	Unique
704	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□■■□□■□□■■	703777740003111	CAS1-Delhi	1 (0.75)	16.67	Unique
777	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□■□□□□■■■■■■■■	777777777420771	H3/H4 (Ural-1)	1 (0.75)	2.04	Unique
878	■■■■■■■■■■■■■■■■■□■■■■■■■■■■■■■□□□□■□■■■■■	777776777760571	X1	1 (0.75)	4.17	Unique
954	■■■□□□□■■■■□■■■■■■■■■■□□□□□□□□□□□□■■■■■■■■■	703677740003771	CAS1-Delhi	1 (0.75)	7.14	Unique
1144	■■■■■■■■■■■■■■■■■■■■■□□□□□□□■■■■■□□□□■■■■■■■■	777777600760771	T1	20 (14.93)	83.33	Clustered
1198	■■■□□□□■■■■■■□■■■■■■■■■□□□□□□□□□□□□■■□□■■■■■	703737740003171	CAS1-Delhi	6 (4.48)	17.65	Clustered
1789	■■■□□□□■■■■■■■■■■■■■■■□□□□□□□□□□□□□□□□□■■■■■	703777740000171	CAS	1 (0.75)	4.76	Unique
2145	■■■□□□□■■■■■■■■■■■■■■□■□□□□□□□□□□□□■■■■■■■■■	703777640003771	CAS1-Delhi	1 (0.75)	10.0	Unique
2230	■■■■■■■■■■■■□□□□□□□□□■■■■■■■■■■■□□□□■■■■■■■■	777700077760771	Unknown	1 (0.75)	20.0	Unique
3371	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□■■■■■■■□■	703777740003761	CAS1-Delhi	2 (1.49)	50.0	Clustered
3776*	□□□□■■■■■■■■□□■■■■■■■■■■■■■■■■■■□□□□■■■□■■■	037637777760731	T2	1 (0.75)	50.0	Unique
Orph 1	■■■■■■■■■■■■□□■■■■■■■■■■■■■■■■■■□□□□□■■■■■■	777717777760371	T1	1 (0.75)	100.0	Orphan
Orph 2	■■■□□■□■■■■■■■■■■■■■■■■■■■■■■■■□□■■□□■■■■■	713777777763171	Manu2	1 (0.75)	100.0	Orphan
Orph 3	■■■■■■■■■■■■■■■■■■■■■■□□□□□■■■■■□□■■■■■■■■■	777777740763771	Manu2	1 (0.75)	100.0	Orphan
Orph 4	□□■■■■■■■■■■■■■■■■■■■■■■■■■■□□□■□□□□■■□■■■■	177777777420671	H4 (Ural-2)	1 (0.75)	100.0	Orphan
Orph 5	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□■■□□■■■□■	703777740003161	CAS1-Delhi	1 (0.75)	100.0	Orphan

A total of 33/34 SITs matched a preexisting shared-type in the database, whereas 1 SIT (SIT3776, $n = 1$ isolate) was newly created after a match with an orphan strain from Iran. Note that orphan patterns ($n = 5$) are shown at the bottom of the table, and mentioned as Orph 1–5, which correspond to the following strain numbers: Orph 1 (3284); Orph 2 (5760); Orph 3 (5953); Orph 4 (6005); Orph 5 (6320).

**Clade designations according to SITVIT2 using revised SpolDB4 rules; "Unknown" designates patterns with signatures that do not belong to any of the major clades described in the database.

***Clustered strains correspond to a similar spoligotype pattern shared by 2 or more strains "within this study"; as opposed to unique strains harboring a spoligotype pattern that does not match with another strain from this study. A unique strain matching a preexisting pattern in the SITVIT2 database is classified as SIT, whereas in case of no match, it is designated as "Orphan" ($n = 5$, shown at the bottom part of the Table).

Table 2 – Distribution of spoligotype lineage and sublineages in this study.

M. tuberculosis lineages	Sublineage	No. of isolates (%)
CAS	CAS	1 (0.75)
	CAS1-Delhi	56 (41.80)
Haarlem	H1	5 (3.73)
	H3	8 (5.97)
	H3/H4 (Ural-1)*	1 (0.75)
	H4/Ural-2*	5 (3.73)
LAM	LAM	2 (1.49)
	LAM6	1 (0.75)
MANU	MANU2	8 (5.975)
	S	1 (0.75)
T	T	1 (0.75)
	T1	36 (26.8)
	T2	1 (0.75)
	T3	2 (1.49)
Turkey	Turkey*	4 (2.98)
Unknown	Unknown	2 (1.49)
X	X1	1 (0.75)
Total		134 (100)

* Note that, (i) H3/H4-Ural-1 and H4/Ural-2 are now considered as "Ural lineage"; (ii) the Turkey lineage was previously referred to as LAM7-TUR.

namely; CAS (42.5%), ill-defined T (29.9%), Haarlem (14.2%, that could be split in Haarlem sensu-stricto, 9.7%; and Ural 4.48%, since H3/H4-Ural-1 and H4/Ural-2 are now considered as "Ural lineage"), and MANU (6%).

Table 3 summarizes the distribution of the MDR strains according to shared types. Amongst the 70 MDR strains, 24 spoligotyping patterns were identified; 56 isolates (80%) were grouped into 10 clusters (2–18 isolates/clusters), and 14 were unique. The clustering rate was 65.7%, which was comparable to that in the whole set of isolates in this study. Amongst the 70 MDR strains, the clustering rate was 65.7%, which was comparable to the whole study set of isolates. The SIT1144/T1 was over-represented amongst the MDR isolates comprising around one quarter (25.7%, $n = 18$) of the total MDR cases and thus 90% of all SIT1144 strains were MDR. Therefore, it seems that this SIT has a great propensity to be MDR. In addition, 8/12 of SIT22/CAS1_Delhi strains (66.7%) were MDR comprising 11.4% of total MDR cases. Furthermore, 6/6 (100%) of SIT50/H3 and 4/5 (80%) of SIT247/CAS1_Delhi strains were MDR. The remaining MDR strains were as follows: 4/7 SIT284 (5.7%), 3/8 SIT26 (4.2%), 3/4 SIT41 (4.2%), 2/2 SIT29 (2.8%) and 2/4 SIT127 (2.8%). On the contrary, certain SITs showed the opposite trend, where 4/4 of the SIT47/H1 strains were non-MDR.

Construction of spoligoforest graphs

For visualizing the relationships among genotypes of tuberculosis strains in this study, a spoligoforest was constructed using the application available in the spolTools website http://www.emi.unsw.edu.au/spolTools (Fig. 1). Information concerning the age of a spoligotype is manifested in three attributes of a spoligoforest. First, the larger the node size, the longer the time over which strain with that spoligotype has been extensively transmitted. Second, the larger number of descendant (outbound edges), the longer period the strain has to generate mutations. Third, the position of the node gives evidence on the age of the spoligotype–the closer it is to the root node, the older it is.

Note that the meaning/significance of the links/edges is the same for Hierarchical Layout and Fruchterman–Reingold trees. Solid black lines link patterns with a maximum weight of distance (very similar: loss of one spacer). Dashed lines represent a link of weight comprised between 0.5 and 1, and dotted lines represent a link of weight less than 0.5.

In both trees, one can denote that SIT1144 (T1) is the biggest and more visible node ($n = 20$). SIT25 (CAS1-Delhi) represents the second biggest node ($n = 19$). The 3 following predominant nodes are constituted by SIT22/CAS1_Delhi ($n = 12$), SIT26/CAS1_Delhi ($n = 8$), and SIT284/T1 ($n = 7$).

Table 3 – Distribution of MDR cases according to the SITs.

SIT		MDR		Total
		Non-MDR	MDR	
SIT	Orphan	3	1	4
	22	4	8	12
	25	13	6	18
	26	5	3	8
	29	0	2	2
	34	0	1	1
	36	1	0	1
	37	1	1	2
	41	1	3	4
	46	0	1	1
	47	4	0	4
	49	0	1	1
	50	0	6	6
	53	3	1	4
	54	7	1	8
	62	1	0	1
	64	1	0	1
	86	0	1	1
	102	0	1	1
	127	2	2	4
	247	1	4	5
	284	3	4	7
	516	1	0	1
	520	0	1	1
	704	1	0	1
	777	0	1	1
	878	1	0	1
	954	0	1	1
	1144	2	18	20
	1198	6	0	6
	1789	1	0	1
	2145	0	1	1
	2230	1	0	1
	3371	1	1	2
	3776	1	0	1
Total		64	70	134

MIRU-VNTR

Among the 134 *M. tuberculosis* non-duplicating isolates, 122 showed amplification products of the 15 loci MIRU-VNTR, whereas 12 isolates had incomplete or no MIRU-VNTR profiles even after repeating the test. Highly diverse MIRU-VNTR genotypes were displayed; 100 distinct MIRU-VNTR genotypes

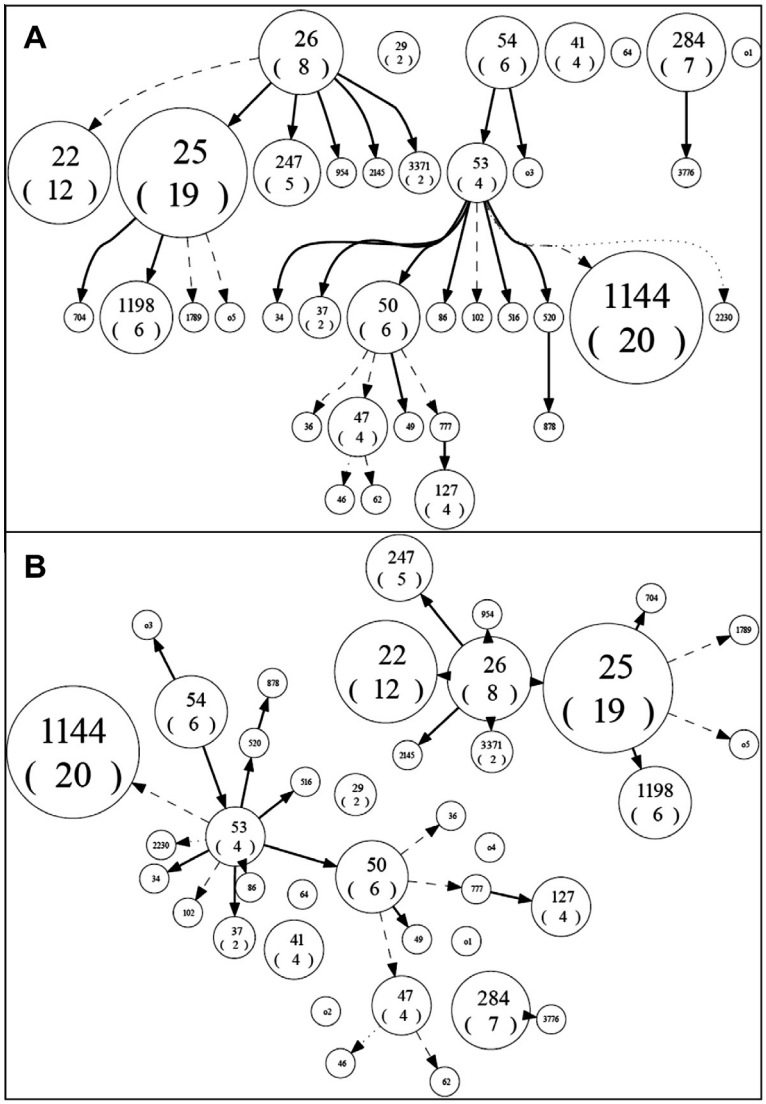

Fig. 1 – The spoligoforest trees drawn using http://www.emi.unsw.edu.au/spolTools/. The trees represent each spoligotype pattern by a node with area proportional to the number of isolates with that spoligotype pattern. Changes (loss of spacers) are represented by directed edges between nodes, with the arrowheads pointing to descendant spoligotypes. (A). The Hierarchical Layout represents hierarchically the changes between MTBC spoligotypes. The more the strains evolve (lose spacers), the more they are present in the down layouts. Thus the patterns which have not lost many alleles/spacers are located in the upper layouts as opposed to those which have lost more spacers (and are located in the down layouts). When there are too many changes between the patterns of spoligotype, there are no links/edges linking them. (B). The Fruchterman–Reingold trees use a force-based algorithm to position the nodes of a graph in a two-dimensional or three-dimensional space so that all the edges are of more or less equal length and there are as few crossing edges as possible. Thus the nodes containing a huge number of strains are centered and are more visible in the figure.

were detected, including 8 clusters (2–8 strains/cluster) and 92 orphan genotypes (unique). The clustering rate was 18.03% (Supplemental Table S1).

The largest cluster defined by 15-MIRU-VNTR ($n = 8$) was MIT230 (342442442273433); followed by MIT221 (522364326235447) ($n = 5$). The MIT222 and MIT231 clusters each included 4 strains; the MIT 223 cluster included 3 strains. The MIT228, MIT229, MIT232, clusters each with 2 strains per cluster. Noteworthy: the majority of strains within every MIRU cluster were isolated in the same geographic region; this may imply an epidemiological link. Of note, in most of the clusters, the strains were from Baghdad. This may be because Baghdad is the most populated province amongst the 18 provinces of Iraq (being inhabited by almost one third of the total Iraqi population).

Allelic diversity

The results of allelic diversity and the Hunter-Gaston Diversity Index (HGDI) for all strains are summarized in Table 4, whereas Table 5 summarizes the HGDIs for CAS and T strains, in addition to the MDR strains. HGDI of all strains revealed that the

Table 4 – Distribution of 122 *M. tuberculosis* containing variable numbers of each MIRU-VNTR locus and allelic diversity.

Loci	Frequency at each locus									Allelic Diversity (HGDI)	Conclusion*
	1	2	3	4	5	6	7	8	9		
MIRU4	5	115	1	1						0.10	Poorly discriminate
MIRU10		10	19	29	33	23	7	1		0.82	Highly discriminate
MIRU16	4	22	37	53	2	4				0.69	Highly discriminate
MIRU26	7		10	3	43	10	44	3	2	0.73	Highly discriminate
MIRU31		6	60	3	50	3				0.59	Moderately discriminate
MIRU40	3	19	57	34	1			1		0.69	Highly discriminate
ETRA	3	47	34	28	2	6				0.74	Highly discriminate
ETRC		56	21	40	3					0.67	Highly discriminate
Qub11b		75	5	23	3	3	6			0.66	Highly discriminate
Qub26	2	12	47	50	8	4				0.68	Highly discriminate
Qub4156		1	26	5	18	25	23	15	9	0.85	Highly discriminate
Mtub04	1	26	40	5	32					0.79	Highly discriminate
Mtub21	6	29	26	39	7	15				0.72	Highly discriminate
Mtub30	1	99	1	21						0.33	Moderately discriminate
Mtub39	2	1	22	76	16	2		1		0.59	Moderately discriminate
H value	0.64										

Abbreviations: MIRU-VNTR, mycobacterial interspersed repetitive units-variable number of tandem repeats; HGDI, Hunter-Gaston Diversity Index.

* Discriminatory power indicated as defined by Sola et al. (2003).

Table 5 – Allelic diversity of the MIRU-VNTR loci in CAS vs. other lineages, T-lineage vs. other lineages and in MDR-TB isolates vs. non MDR-TB.

MIRU-VNTR loci	HGDI value of CAS vs. other lineages			HGDI value of T lineage vs. other lineages			HGDI value of MDR-TB vs. non-MDR TB		
	CAS lineage	Other lineages	*P*-value	T-linage	Other lineages	*P*-value	MDR-TB	Non-MDR-TB	*P*-value
MIRU4	0.01[a]	0.2[a]	0.292	0.06[a]	0.14[a]	0.119	0.2[a]	0.01[a]	0.284
MIRU10	0.81[c]	0.77[c]	0.000*	0.67[c]	0.81[c]	0.000*	0.8[c]	0.77[c]	0.990
MIRU16	0.46[b]	0.7[c]	0.000*	0.6[c]	0.61[c]	0.000*	0.73[c]	0.65[c]	0.008*
MIRU26	0.69[c]	0.76[c]	0.010*	0.71[c]	0.74[c]	0.552	0.74[c]	0.72[c]	0.066
MIRU31	0.2[a]	0.26[a]	0.000*	0.34[b]	0.55[b]	0.000*	0.55[b]	0.62[c]	0.007*
MIRU40	0.45[b]	0.73[c]	0.830	0.56[b]	0.33[b]	0.380	0.72[c]	0.57[b]	0.015*
ETRA	0.71[c]	0.74[c]	0.563	0.62[c]	0.74[c]	0.107	0.67[c]	0.79[c]	0.243
ETRC	0.12[a]	0.6[c]	0.000*	0.46[b]	0.53[b]	0.000*	0.66[c]	0.63[c]	0.055
Qub11b	0.24[a]	0.72[c]	0.000*	0.61[c]	0.47[b]	0.000*	0.62[c]	0.55[b]	0.062
Qub26	0.32[b]	0.54[b]	0.000*	0.32[b]	0.62[c]	0.000*	0.67[c]	0.67[c]	0.031*
Qub4156	0.88[c]	0.78[c]	0.000*	0.58[b]	0.84[c]	0.000*	0.82[c]	0.85[c]	0.004*
Mtub04	0.68[c]	0.64[c]	0.000*	0.45[b]	0.79[c]	0.009*	0.76[c]	0.82[c]	0.929
Mtub21	0.74[c]	0.73[c]	0.000*	0.69[c]	0.8[c]	0.007*	0.78[c]	0.79[c]	0.317
Mtub30	0.04[a]	0.48[b]	0.000*	0.12[a]	0.39[b]	0.034*	0.36[b]	0.31[b]	0.223
Mtub39	0.31[b]	0.68[c]	0.570	0.63[c]	0.53[b]	0.000*	0.43[b]	0.59[b]	0.944
H value									

[a] Poorly discriminate.
[b] Moderately discriminate.
[c] Highly discriminate.
* Statistically significant.

discriminatory power of the MIRU-VNTR was higher than that of spoligotyping (0.992 and 0.930, respectively). For the CAS strains, HGDI value was 0.82 for spoligotyping and 0.988 for the MIRU-VNTR, whereas for T strains, HGDI values for spoligotyping and MIRU-VNTR were 0.73 and 0.951, respectively. For the MDR strains, HGDI value for spoligotyping was 0.920, and its value for MIRU-VNTR was 0.980. Furthermore, the allelic diversity of each of the 15 MIRU-VNTR loci was calculated to identify the discriminatory index in each locus [25]. For all strains, 11 loci were highly discriminate ($h > 0.6$), 3 moderately discriminate (Mtub30, Mtub39, MIRU31; $0.3 \leqslant h \leqslant 0.6$) and 1 loci was less polymorphic or poorly discriminate (MIRU4; $h < 0.3$). QUB4156 was the most discriminate loci ($h = 0.85$), whereas MIRU4 was the least discriminate one ($h = 0.1$). The mean of the allelic diversity of 15 loci was 0.65 as measured according to the method described elsewhere [26].

Ten loci were informative in CAS and T strains as they were effective in differentiating them from other strains

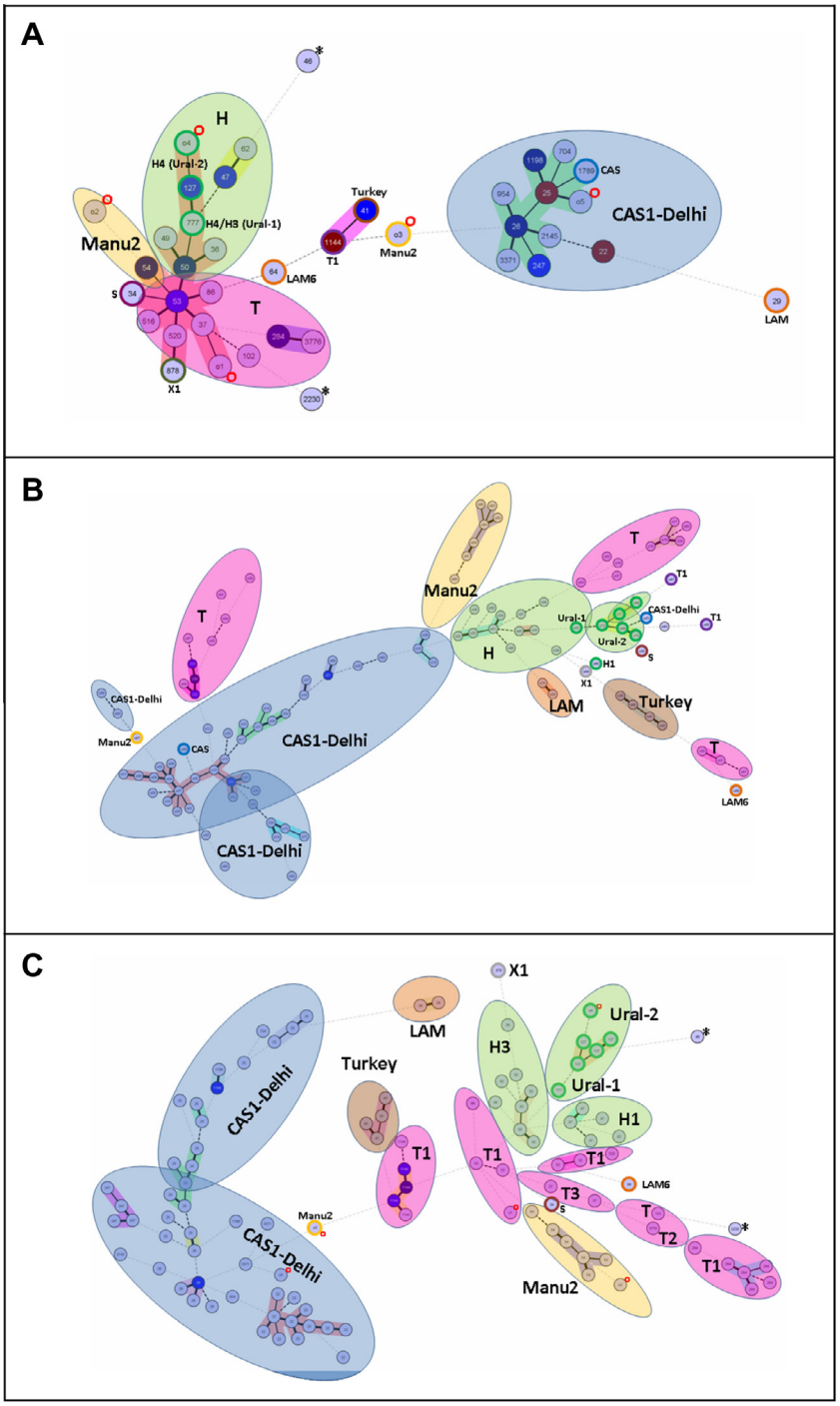

Fig. 2 – A minimum spanning tree (MST) illustrating evolutionary relationships between the Iraqi *M. tuberculosis* complex (MTBC) genotypes: (A) Spoligotypes; (B) MIRU-VNTRs; and (C) Spoligotypes and MIRU-VNTRs pooled together. The tree connects each genotype based on the degree of changes required to go from one allele to another. The structure of the tree is represented by branches (continuous vs. dotted lines) and circles representing each individual pattern. Note that the length of the branches represents the distance between patterns while the complexity of the lines (continuous, black dotted and gray dotted) denotes the number of allele/spacer changes between two patterns: solid lines, 1 or 2 change (thicker ones indicate a single change, while the thinner ones indicate 2 changes); dotted lines, three or more changes (black dotted for 3, and grey dotted for 4 or more changes). The color of the circles is proportional to the number of clinical isolates in this study, illustrating unique isolates (sky blue) vs. clustered isolates (Blue, 2–5 strains; Dark blue, 6–9 strains; Bordeaux, 10–19 strains; Red, 20 and more). Note that: (i) spoligotype orphan patterns ($n = 5$) are marked by the letter "o" in bold red in Fig. 1A; (ii) patterns with an unknown spoligotyping signature (i.e., unclassified in SITVIT2 database) are marked by an asterisk (*) in Fig. 1A and C; (iii) five strains without MIRU typing data (strain ID 5291, 6255, 2478, 4826, and 6223 in Supplemental Table S1) were not included in MSTs shown in Figs. 1B and C.

(MIRU10, MIRU16, MIRU31, ETR-C, QUB11b, QUB26, QUB4156, Mtub04, Mtub21, Mtub30 ($p \geqslant 0.05$). In addition, one more locus was informative in each of CAS and T strains (MIRU26 and Mtub39, respectively) (Table 5). Comparison of the allelic diversity between MDR cases and non-MDR isolates revealed that MIRU16, MIRU31, MIRU40, QUB26, QIB4156 were statistically significant in discriminating the MDR cases from non-MDR cases.

Combination of spoligotyping and MIRU-VNTR data analyses

Fig. 2 shows a minimum spanning tree (MST) illustrating evolutionary relationships between the Iraqi *M. tuberculosis* complex (MTBC) genotypes. MSTs based on the 15 locus MIRU-VNTR (Fig. 2B) showed a better resolution than spoligotyping alone (Fig. 2A). However, the highest discriminatory power was provided by spoligotyping and MIRU-VNTR used together (Fig. 2C).

By using MIRU-VNTR the clustering rate was decreased from 70.8% on spoligotyping to 18.03% on MIRU-VNTR. Many spoligotyping clusters were subdivided or split into smaller clusters or into unique patterns; clusters belong to CAS1-DELHI lineage ($n = 50$) and ill-defined T lineage ($n = 33$) were further subdivided by MIRU-VNTR to 4 clusters with 13 isolates belonging to CAS1-DELHI and 4 clusters with 17 isolates belonging to ill-defined T genotypes. Notably, SIT1144/T1 cluster ($n = 20$) was subdivided into 3 smaller clusters containing 8, 4, 3 isolates/cluster. Similarly, SIT25/CAS1_DELHI cluster ($n = 19$) was subdivided to one cluster of 2 strains and 17 unique profiles.

Discussion

This report demonstrates the largest amount of molecular epidemiology data on *M. tuberculosis* from Iraq to date. Studying the molecular epidemiology of *M. tuberculosis* clinical isolates may feed a successful TB control program, as this may unravel the disease transmission dynamics and may highlight the most successful strains circulating within a population in a specific geographical region. In Iraq, data on *M. tuberculosis* genotypes are very limited. Thus, genotyping of *M. tuberculosis* isolates and determination of predominating genotypes using spoligotyping and MIRU-VNTR methods were the aims of this study.

In the current study, spoligotyping of 134 *M. tuberculosis* isolates yielded high clustering rates (70.8%). High clustering rates were also reported in the neighboring countries, such as Turkey [11,27] and Saudi Arabia [28]. The basic idea behind identification of a cluster is the notion that clustering indicates an ongoing or recent transmission, while unique patterns indicate reactivation events [23,29]. Nonetheless, spoligotyping is known to overestimate clustering, and only a systemic second-line typing, such as MIRUs, couples with a better tuberculosis registry and epidemiological investigations would allow recognizing the rate of ongoing transmission.

Among 39 spoligotyping patterns observed, 33 patterns matched preexisting shared types in the international database SITVIT2. One SIT was newly created after matching with an orphan strain from Iran; this strain was recovered from a patient from Misan Province, which is close to the Iranian borders. Additionally, 5 patterns were new and unique (orphan). Although orphan strains did not match any preexisting SIT, nevertheless, a comparison of their patterns based on the absence and presence of spacers with that found in other SITs revealed that these orphans were closely related to major lineages. These results may suggest that evolving events have occurred and resulted in the emerging of these orphan spoligotyping patterns. In contrast to many countries where extreme dominance of a single lineage is seen, this study showed that several lineages are in circulation within Iraqi patients. A specific comparison was also made with the three "big" neighboring countries of Iraq: Iran, Saudi Arabia, and Turkey. Results of comparisons showed that even if more or less of the same lineages are in circulation among these four countries, each country has its own most predominant lineage(s). Iraq is specific in having its own most predominant lineage (SIT1144/T1, see below), which is not found among neighboring countries.

The predominant lineage in this study was the CAS (42.5%). The CAS lineage is characterized by the absence of spacer 4–7 and 23–34 [30]. This lineage is reported in many countries [31,32]; however, the "high incidence" areas of CAS1 family are the Middle East and Central and Southern Asia [33–36]. It is identified as a predominant strain in Pakistan [32], Delhi [33] and Mumbai, India [37]. In China, CAS1 type is almost exclusively detected in Xinjiang Uygur Autonomous Region [38]. Geographically, the Xinjiang region is adjacent to several Central Asian countries, and they share similar culture traditions. Historically, those countries, along with Iraq, are all located along what was called the "Silk Road" that was the backbone of trading between the Far East and the Middle East in history [39]. Movements of people and merchants along this road may have played a major role in transmission of this strain type between those countries [40]. Indeed, six clusters belonging to the CAS lineage were reported in this study. These clusters were of variable sizes, ranging from 19 isolates (SIT25) to 2 isolates (SIT3371). The presence of the CAS lineage in big clusters may indicate successful circulation of this lineage within the population and also may implicate inefficiency of the current TB control program. Moreover, detecting clusters alongside with unique patterns of strains belonging to the CAS lineage in the Iraqi population may reflect ongoing evolution events. Furthermore, the majority of SIT25/CAS1_Delhi strains were isolated from a younger age group–a result that may reflect the recent spread of this shared type.

The ill-defined T lineage which can be subdivided into 5 sublineages [33], was the 2nd most predominant lineage in our study ($n = 40$, 29.9%). The T family genotype prevails in Africa, Central and South America and Europe [36]. Although there is no geographic link between Iraq and these regions, both of the two largest surrounding countries to Iraq, Turkey and Iran are known to share active borders and historic links with European countries; therefore, it may be suggested that the T–family (except SIT1144, see below) may have first spread in Turkey and/or Iran and then transmitted to Iraq. Several SITs (clusters) belonging to the T-family were recorded in this study; notably, SIT1144 was the biggest

cluster recorded in this study ($n = 20$, 14.93% of all cases). SIT1144/T1 showed the following distribution pattern in the SITVIT2 database ($n = 24$ strains in total: Iraq $n = 20$, USA $n = 2$, Venezuela $n = 1$, Greece $n = 1$). It is not known whether there is a link between political displacements and/or war and the presence of those sporadic strains in USA, Venezuela and Greece. In this study, SIT1144 is mostly isolated from younger age groups; A finding that may indicate the recent spread of this shared type. In addition, 18/20 SIT1144 strains were MDR, suggesting that this shared type has more propensity to develop multi-drug resistance. SIT284/T1 was represented in this study by 7 isolates (5.22%). This SIT is remarkable by the absence of spacers 1–4, 12–13, and 33–36. Considering the total number of strains in the SITVIT2 database ($n = 167$), the global proportion of SIT284 strains is more predominant in Bulgaria (23.95%), followed by Turkey (20.96%) and Saudi Arabia (13.77%). However, this new Iraqi study revealed that considering the proportion of the SIT number per country, SIT284 is present in Iraq in a greater proportion than the neighboring countries. SIT53 had a significantly higher proportion in Turkey. This might reflect the transmission of these isolates between the two countries.

Characterized by the absence of spacers 29-31 and 33-36, Haarlem and Ural group taken collectively ranked 3rd in our study ($n = 19$, 14.2%). Isolates belonging to these lineages have been reported in several countries, including Iran, Armenia, Finland, Georgia, and Russia [18]. Indeed, in one study from Iran, this lineage was the most frequent among the Iranian population [41]. Looking at the spread of SITs belonging to the Ural and H families in Iraq and its neighboring countries, it can be noted that the SIT127 (Ural-2) was significantly more prevalent in Iran, whereas SIT50 was significantly more prevalent in Turkey, and SIT47 was significantly more prevalent in Iraq. MANU lineage was ranked as the fourth lineage spreading in Iraq ($n = 8$, 6%), corresponding to SIT54/MANU2 (6 isolates) and 2 orphans. The MANU2 family has a low prevalence worldwide. Indeed, the MANU family is divided into MANU1 (deletion of spacer 34), MANU2 (deletion of spacers 33–34), and MANU3 (deletion of spacers 34–36). In addition, a high proportion of drug resistance was found in ST54/MANU2 genotypes. A similar result was previously reported in China [42].

The SIT41/Turkey family (previously classified as LAM7-TUR, but recently shown to be a lineage that does not belong to the LAM lineage), is probably phylogenetically close to the X-family of Anglo-Saxon descent as well as the ill-defined T lineage. Previous studies demonstrated that this shared type was exclusively isolated in Turkey. [10,11,43–45]. However, in this study, it was reported that 4 isolates belong to this family. Several factors may contribute to the spread of this genotype in Iraq; Turkey had strong historic relationships with Iraq; indeed, Iraq was a part of the Ottoman Empire. There is a Turkish ethnic minority in Iraq, and there is continuous population movement between those two countries for trading, tourism and immigration. Moreover, three fourths (75%) of the SIT41 strains were MDR; this is inconsistent with the study of Kisa et al. [10] who found that this shared type is associated with drug resistance. LAM lineage was represented in this study by three isolates (2.2%) corresponding to SIT 29 ($n = 12$) and SIT64 ($n = 1$). LAM lineage prevails in Southern Europe, Africa [18] and South America [46]. In Saudi Arabia and Turkey, 7.2% and 5.3%, respectively, of the strains were attributed to LAM lineages [10,28]; therefore, it is not surprising to report this lineage in Iraq.

Regarding age and gender of patients, two notable features were emerging within the two largest clusters (SIT1144 and SIT25); first, the majority of isolates were from younger age groups (<35 years old); second, in contrast to the whole set of isolates, there was almost an equal male to female ratio seen in those 2 clusters. By assuming that clusters arise from recent transmission rather than reactivation, this may explain the discrepancy in male to female ratio as a proportion of disease from reactivation that is occurring in males more than females.

The clustering rate among the MDR strains was 65.7%, which is comparable with the full set of isolates. Around one quarter (25.7%, $n = 18$) of the MDR cases was attributed to the SIT1144/T1; therefore, it seems that this SIT may have a propensity to be MDR. This is an alarming result, and further studying of this SIT is required, as well as the monitoring of its spread is recommended. In addition, other strains have also showed a propensity toward MDR, such as SIT50/H3 (6/6,100%), SIT247 (4/5, 80), SIT41 (3/4, 75%), and SIT29 (2/2, 100%). Drug resistance was also reported in these strains in Turkey [47]. High percentages of MDR within certain genotypic lineages were also reported in several studies [48]. However, the proportions of MDR cases were higher in this study, and this may be due to the fact that these cases are biased toward treatment failure. In contrary, certain SITs showed the opposite trend; for instance, all SIT1198 (6/6) and SIT47/H1 (4/4) strains, in addition to the majority of SIT54 (7/8) strains were non-MDR. Furthermore, a significant association was found between DST results and gender (females associated with MDR [$P = 0.02$]). Previous studies reported an association between TB infection and other demographic factors like young age [23,49], and drug resistance [49].

A spoligoforest was constructed to visualize the relationships among spoligotypes reported in this study. Spoligoforest is a tool to recover the history of transmission and mutation events [50]. As shown in Fig. 2, several large spoligotype nodes were reported in this study, including SIT1144, SIT25, SIT22, SIT284...etc. These shared types may have been transmitted extensively over a long period of time. The SIT53 has the largest number of descendent/outbounds ($n = 9$), followed by the SIT26 ($n = 6$); this implies that those strains have had a long period of time over which they generated mutations. In addition, as shown in Fig. 2A, certain shared types (such as SIT26, SIT54, SIT284) are located in the upper layouts and thus they potentially represent the oldest spoligotype in this forest. Furthermore, the SIT26 has a large node size yet positioned at a tip of the spoligoforest; this may indicate that the strain within this spoligotype are being transmitted faster than other strains in this data set [51]. The SIT1144, SIT50, and SIT1198 are present in the down layouts and could therefore be "emerging strains". In line with this finding, MIRU-VNTR showed a high strain homogeneity among the strains of the SIT1144; this suggests once again that SIT1144 may be associated with a higher transmission rate than other strains. However, further quantitative analyses are required to verify this point.

Although a link between SIT1144/T1 and SIT41/Turkey is not highlighted in spoligoforests (Fig. 1), one may notice some similarities between these two patterns. Indeed, only two changes differentiate these spoligotypes: SIT1144 is defined by the presence of spacer 20 and the absence of spacer 25 as opposed to SIT41, which is defined by the absence of spacer 20 and the presence of spacer 25. Like LAM lineage's signature, SIT41 is characterized by the absence of spacers 21–24 and spacers 33–36. However, contrary to the traditional LAM signature, spacer 25 is absent in the SIT1144 signature. Hence, further studies on a finer characterization of SIT41 and SIT1144 might help understand the reasons for its predominance in this study vs. its absence in neighboring countries, as well as the naming/renaming of the LAM7-TUR/Turkey lineage(s). It is noteworthy that the 15-loci MIRUs did not underline any direct link between SIT41 and SIT1144 in this study (Supplemental Table S1).

There was a high diversity of MIRU-VNTR patterns among the isolates in this study; 100 distinct profiles were detected corresponding to 92 orphan profiles and 8 MIT clusters each with 2–8 isolates/cluster. The clustering rate was 18.03% (see Supplemental Table S2), implying that the minimum estimate of disease related to recent transmission was 18.3%. Despite the fact that MIRU-VNTR typing revealed decreased clustering rates compared with spoligotyping, preliminary MIRU-VNTR analyses suggest that all of the most common spoligotype clusters, except SIT25 and SIT50, contain MIRU patterns with high profile homogeneity (less than 3 loci difference, 80% or more homogeneity) even if the strains had been isolated at distinct geographic regions of Iraq, suggesting that those SITs comprise clonal groups of strains. These results implicate that all spoligotype clusters, except SIT25 and SIT50, may represent large clonal clusters. Several studies have reported that some MIRU loci have faster molecular clocks [16,52–54], which could explain the single locus variations between isolates linked to an ongoing transmission. Furthermore, all of the isolates that were classified as unique by spoligotyping were shown to be orphans by MIRU-VNTR. If this is true, the majority of these spoligotype clusters could be linked to the transmission process. Accordingly, SIT25 and SIT50 contained different strain genotypes and thus are less likely to be linked to recent transmission. It is noteworthy that the majority of strains within MIRU clusters were isolated in the same geographic regions; this may represent an epidemiological link. Furthermore, most of the clusters were involving more than two strains (5/8). It is generally assumed that the detection of links in a transmission chain involving a substantial number of cases produce higher yields. These results may indicate that geographic regions could explain the majority of the MIRU-VNTR clusters. Of note, most cluster strains were from Baghdad; this may be because Baghdad is the most populated province amongst the 18 provinces of Iraq (being inhabited by almost one third of the total Iraqi population). The geographic explanation of the MIRU-VNTR clusters in this study may confirm the value of MIRU-VNTR identification of transmission events.

Notably, the SIT1144/T1 was divided into three smaller clusters by MIRU-VNTR typing that only differ in 1 or 2 loci. These results suggest SIT1144 might represent a clonal group of strains. In addition, analysis of the geographic distribution of SIT1144 strains shows that they were isolated in different regions in Iraq. Thus, SIT1144 strains seem to have selective advantages over other strains in their ability to cause disease and be transmitted.

Evidence suggests that the choice of appropriate loci for MIRU-VNTR may require evaluation in diverse MTB lineages in countries with a high TB prevalence [55,56]. HGDI calculations revealed that the discriminatory power of the MIRU-VNTR was higher than that of spoligotyping, 0.992 and 0.93, respectively. This confirms that MIRU-VNTR typing is most suited for the study of molecular diversity of *M. tuberculosis* isolates.

For combined spoligotyping and MIRU-VNTR analysis, isolates with identical SIT and MIT profiles were categorized as "cluster," whereas "unique" profiles were reported for single isolates. Combined spoligotyping and MIRU-VNTR analysis classified 30 strains into 2 major clades: ill-defined T (17 isolates clustered into 4 clusters) and CAS1-DELHI (13 isolates clustered into 4 clusters).

In addition, the allelic diversity was calculated for each locus in 15-MIRU-VNTR within three formats: (a) within whole study strains; (b) within MDR cases; and (c) within the 2 largest lineages (CAS and T). MIRU16, MIRU31, MIRU40, QUB26, and QIB4156 had a significantly higher discriminatory index with P value <0.05 and thus efficient in discriminating the MDR cases from the non-MDR strains; MIRU10, MIRu16, MIRU26, MIRU31, ETR-C, QUB11b, QUB26, QUB4156, Mtub04, Mtub21, and Mtub30 significantly discriminated CAS from non-CAS strains; and MIRU10, MIRU16, MIRU31, ETR-C, QUB11b, QUB26, QUB4156, Mtub04, Mtub21, Mtub30, and Mtub39 were able to significantly discriminate between the T and non-T isolates. HGDI (spoligotyping and MIRU-VNTR) for both T and CAS1-DELHI were (0.82 and 0.988), and (0.73 and 0.951), respectively. The clustering rates within the predominant lineages in Iraq were 80.7% for CAS and 72.5% for T. This might reflect that isolates belonging to these two lineages are more actively transmitted than other isolates. Despite MIRU-VNTR typing splits the CAS and T spoligotype SIT into non-clustered profiles, the differences between MIRU profiles does not exceed 3 loci (except SIT25- which seems to be more polymorphic). This might be a result of intrinsic similarity or due to reflective relatedness within CAS and T lineages.

Conclusions

TB in Iraq is occurring in younger age groups with twice as many males than females. In contrast to many countries where extreme dominance of a single lineage is seen, this study showed that several lineages are in circulation within Iraqi patients. The predominant lineage was CAS (42.5%), followed by the ill-defined T (29.9%). This was expected as those 2 lineages are common isolates in the neighboring countries. In addition, there were other lineages that were well represented among these isolates, such as Haarlem (14.2%; split in Haarlem sensu-stricto 9.7%, and Ural 4.48%), and MANU (6%). It is noteworthy, the Turkey lineage, that previously reported to be exclusively isolated in Turkey, was represented in this study by four isolates (3%). Surprisingly, the Beijing family–the worldwide distributed group of strains–was not

represented in this study; thus, it appears that Iraq is preserved from the transmission of this problematic family of strains. A comparison of the spoligotypes in this study with the three "big" neighboring countries of Iraq–Iran, Saudi Arabia, and Turkey–showed that if more or less the same lineages are in circulation among these four countries, each country has its own most predominant lineage(s). Iraq is unique in having its own most predominant lineage (SIT1144/T1), which is not found among neighboring countries. This shared type has more propensities for multi-drug resistance.

Highly diverse MIRU-VNTR genotypes were displayed; 100 distinct MIRU-VNTR genotypes were detected (8 clusters with 2–8 strains/cluster and 92 unique). The clustering rate was 18.03%. The discriminatory efficiency of MIRU-VNTR was high (HGDI 0.992); it was higher than that of spoligotyping (HGDI 0.930). This high resolution was due to the fact that 14/15 loci were highly to moderately discriminate ($h > 0.6$, and $h > 0.3$, respectively) and the mean of the allelic diversity of the 15 loci was high (0.65). The minimum spanning tree (MST) showed a better resolution by MIRU-VNTR than by spoligotyping alone. However, the highest discriminatory power was provided by spoligotyping and MIRUs together. Owing to the low clustering rate by MIRU-VNTR, these results suggest that drug-resistant TB in Iraq is due to acquired resistance as opposed to transmission.

Conflict of interest

None declared.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijmyco.2014.07.006.

REFERENCES

[1] WHO, Global tuberculosis control. WHO Report 2011. Geneva.

[2] T. Dalton, P. Cegielski, S. Akksilp, L. Asencios, J. Campos Caoili, Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study, Lancet 380 (2012) 1406–1417.

[3] S. Huseynova, D.S. Hashim, M.R. Tbena, R. Harris, A. Bassili, et al, Estimating tuberculosis burden and reporting in resource-limited countries: a capture-recapture study in Iraq, Int. J. Tuberc. Lung Dis. 17 (2013) 462–467.

[4] S. Al-Hajoj, B. Varghese, F. Al-Habobe, M.M. Shoukri, A. Mulder, et al, Current trends of *Mycobacterium tuberculosis* molecular epidemiology in Saudi Arabia and associated demographical factors, Infect. Genet. Evol. 16 (2013) 362–368.

[5] M.W. Borgdorff, D. van Soolingen, The re-emergence of tuberculosis: what have we learnt from molecular epidemiology?, Clin Microbiol. Infect. (2013).

[6] S. Borrell, M. Espanol, A. Orcau, G. Tudo, F. March, et al, Factors associated with differences between conventional contact tracing and molecular epidemiology in study of tuberculosis transmission and analysis in the city of Barcelona, Spain, J. Clin. Microbiol. 47 (2009) 198–204.

[7] C. Allix, K. Walravens, C. Saegerman, J. Godfroid, P. Supply, et al, Evaluation of the epidemiological relevance of variable-number tandem-repeat genotyping of *Mycobacterium bovis* and comparison of the method with IS6110 restriction fragment length polymorphism analysis and spoligotyping, J. Clin. Microbiol. 44 (2006) 1951–1962.

[8] C. Loiez, E. Willery, J.L. Legrand, V. Vincent, M.C. Gutierrez, et al, Against all odds: molecular confirmation of an implausible case of bone tuberculosis, Clin. Infect. Dis. 42 (2006) e86–88.

[9] M.A. Merza, P. Farnia, A.M. Salih, M.R. Masjedi, A.A. Velayati, First insight into the drug resistance pattern of *Mycobacterium tuberculosis* in Dohuk, Iraq: using spoligotyping and MIRU-VNTR to characterize multidrug resistant strains, J. Infect. Public Health 4 (2011) 41–47.

[10] O. Kisa, G. Tarhan, S. Gunal, A. Albay, R. Durmaz, et al, Distribution of spoligotyping defined genotypic lineages among drug-resistant *Mycobacterium tuberculosis* complex clinical isolates in Ankara, Turkey, PLoS One 7 (2012) e30331.

[11] O. Kisa, A. Albay, O. Baylan, E. Tozkoparan, C.H. Acikel, et al, Genetic diversity of *Mycobacterium tuberculosis* isolates at the Military Medical Academy in Ankara, Turkey, Res. Microbiol. 158 (2007) 318–323.

[12] O. Kisa, A. Albay, O. Baylan, A. Balkan, L. Doganci, Drug resistance in *Mycobacterium tuberculosis*: a retrospective study from a 2000-bed teaching hospital in Ankara, Turkey, Int. J. Antimicrob. Agents 22 (2003) 456–457.

[13] M. Rohani, P. Farnia, M.N. Nasab, R. Moniri, M. Torfeh, et al, Beijing genotype and other predominant *Mycobacterium tuberculosis* spoligotypes observed in Mashhad city, Iran, Indian J. Med. Microbiol. 27 (2009) 306–310.

[14] S.Z. Bostanabad, M. Shekarabei, S.A. Nojoumi, E. Jabbarzadeh, M. Ghalami, et al, Study of genetic evolution in *Mycobacterium tuberculosis* isolates from patients with active pulmonary tuberculosis in the Iran and Belarus, Open Microbiol. J. 5 (2011) 32–42.

[15] S. Al Hajoj, N. Rastogi, The emergence of Beijing genotype of *Mycobacterium tuberculosis* in the Kingdom of Saudi Arabia, Ann. Thorac. Med. 5 (2010) 149–152.

[16] P. Supply, C. Allix, S. Lesjean, M. Cardoso-Oelemann, S. Rusch-Gerdes, et al, Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*, J. Clin. Microbiol. 44 (2006) 4498–4510.

[17] N. Alonso-Rodriguez, M. Martinez-Lirola, M. Herranz, M. Sanchez-Benitez, P. Barroso, et al, Evaluation of the new advanced 15-loci MIRU-VNTR genotyping tool in *Mycobacterium tuberculosis* molecular epidemiology studies, BMC Microbiol. 8 (2008) 34.

[18] C. Demay, B. Liens, T. Burguière, V. Hill, D. Couvin, et al, SITVITWEB - a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology, Infect. Genet. Evol. 12 (2012) 755–766.

[19] R.K. Selander, T.K. Korhonen, V. Vaisanen-Rhen, P.H. Williams, P.E. Pattison, et al, Genetic relationships and clonal structure of strains of *Escherichia coli* causing neonatal septicemia and meningitis, Infect. Immun. 52 (1986) 213–222.

[20] P.R. Hunter, M.A. Gaston, Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity, J. Clin. Microbiol. 26 (1988) 2465–2466.

[21] D. Affolabi, G. Anyo, F. Faihun, N. Sanoussi, I.C. Shamputa, et al, First molecular epidemiological study of tuberculosis in Benin, Int. J. Tuberc. Lung Dis. 13 (2009) 317–322.

[22] R.K. Selander, B.R. Levin, Genetic diversity and structure in *Escherichia coli* populations, Science 210 (1980) 545–547.

[23] P.M. Small, P.C. Hopewell, S.P. Singh, A. Paz, J. Parsonnet, et al, The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods, N. Engl. J. Med. 330 (1994) 1703–1709.
[24] A. Ali, Z. Hasan, M. Tanveer, A.R. Siddiqui, S. Ghebremichael, et al, Characterization of *Mycobacterium tuberculosis* Central Asian Strain 1 using mycobacterial interspersed repetitive unit genotyping, BMC Microbiol. 7 (2007) 76.
[25] C. Sola, S. Ferdinand, C. Mammina, A. Nastasi, N. Rastogi, Genetic diversity of *Mycobacterium tuberculosis* in Sicily based on spoligotyping and variable number of tandem DNA repeats and comparison with a spoligotyping database for population-based analysis, J. Clin. Microbiol. 39 (2001) 1559–1565.
[26] C. Sola, I. Filliol, E. Legrand, S. Lesjean, C. Locht, et al, Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics, Infect. Genet. Evol. 3 (2003) 125–133.
[27] B. Gencer, T.M. Shinnick, Molecular genotyping of *Mycobacterium tuberculosis* isolates from Turkey, Am. J. Infect. Dis. 1 (2005) 5–11.
[28] S.A. Al-Hajoj, T. Zozio, F. Al-Rabiah, V. Mohammad, M. Al-Nasser, et al, First insight into the population structure of *Mycobacterium tuberculosis* in Saudi Arabia, J. Clin. Microbiol. 45 (2007) 2467–2473.
[29] T. Zozio, C. Allix, S. Gunal, Z. Saribas, A. Alp, et al, Genotyping of *Mycobacterium tuberculosis* clinical isolates in two cities of Turkey: description of a new family of genotypes that is phylogeographically specific for Asia Minor, BMC Microbiol. 5 (2005) 44.
[30] C. Sola, I. Filliol, M.C. Gutierrez, I. Mokrousov, V. Vincent, et al, Spoligotype database of *Mycobacterium tuberculosis*: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives, Emerg. Infect. Dis. 7 (2001) 390–396.
[31] I. Filliol, J.R. Driscoll, D. van Soolingen, B.N. Kreiswirth, K. Kremer, et al, Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study, J. Clin. Microbiol. 41 (2003) 1963–1970.
[32] Z. Hasan, M. Tanveer, A. Kanji, Q. Hasan, S. Ghebremichael, et al, Spoligotyping of *Mycobacterium tuberculosis* isolates from Pakistan reveals predominance of Central Asian Strain 1 and Beijing isolates, J. Clin. Microbiol. 44 (2006) 1763–1768.
[33] U.B. Singh, N. Suresh, N.V. Bhanu, J. Arora, H. Pant, et al, Predominant tuberculosis spoligotypes, Delhi, India, Emerg. Infect. Dis. 10 (2004) 1138–1142.
[34] R.M. Warren, E.M. Streicher, S.L. Sampson, G.D. van der Spuy, M. Richardson, et al, Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data, J. Clin. Microbiol. 40 (2002) 4457–4465.
[35] S. Godreuil, G. Torrea, D. Terru, F. Chevenet, S. Diagbouga, et al, First molecular epidemiology study of *Mycobacterium tuberculosis* in Burkina Faso, J. Clin. Microbiol. 45 (2007) 921–927.
[36] T. Weniger, J. Krawczyk, P. Supply, S. Niemann, D. Harmsen, MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria, Nucleic Acids Res. 38 (2010) W326–331.
[37] D. Almeida, C. Rodrigues, T.F. Ashavaid, A. Lalvani, Z.F. Udwadia, et al, High incidence of the Beijing genotype among multidrug-resistant isolates of *Mycobacterium tuberculosis* in a tertiary care center in Mumbai, India, Clin. Infect. Dis. 40 (2005) 881–886.
[38] Y. Pang, Y. Zhou, B. Zhao, G. Liu, G. Jiang, et al, Spoligotyping and drug resistance analysis of *Mycobacterium tuberculosis* strains from national survey in China, PLoS One 7 (2012) e32976.
[39] V. Elisseeff, The Silk Road; Highway of culture and Commerce, UNESCO, Paris, 2000.
[40] C.W. Pannell, China Gazes West: Xinjiangs growing Rendezvous with Central Asia, Eur. Geogr. Econ. 52 (2011) 105–118.
[41] P. Farnia, F. Mohammadi, M.R. Masjedi, A. Varnerot, A.Z. Zarifi, et al, Evaluation of tuberculosis transmission in Tehran: using RFLP and spoligotyping methods, J. Infect. 49 (2004) 94–101.
[42] Y. Pang, Y. Song, H. Xia, Y. Zhou, B. Zhao, et al, Risk factors and clinical phenotypes of Beijing genotype strains in tuberculosis patients in China, BMC Infect. Dis. 12 (2012) 354.
[43] R. Durmaz, T. Zozio, S. Gunal, A. Yaman, C. Cavusoglu, et al, Genetic diversity and major spoligotype families of drug-resistant *Mycobacterium tuberculosis* clinical isolates from different regions of Turkey, Infect. Genet. Evol. 7 (2007) 513–519.
[44] E. Aktas, T. Zozio, F.B. Comert, C. Kulah, O. Aydin, et al, A first insight into the genetic diversity and population structure of *Mycobacterium tuberculosis* in Zonguldak, Turkey, Clin. Microbiol. Infect. 14 (2008) 55–59.
[45] R. Durmaz, T. Zozio, S. Gunal, C. Allix, M. Fauville-Dufaux, et al, Population-based molecular epidemiological study of tuberculosis in Malatya, Turkey, J. Clin. Microbiol. 45 (2007) 4027–4035.
[46] H.M. Gomes, A.R. Elias, M.A. Oelemann, M.A. Pereira, F.F. Montes, et al, Spoligotypes of *Mycobacterium tuberculosis* complex isolates from patients residents of 11 states of Brazil, Infect. Genet. Evol. 12 (2012) 649–656.
[47] F. Akinlade, N.J. Cowans, M.C. Kisanga, K. Spencer, Maternal serum CA 19–9 and CA 15–3 levels in pregnancies affected by trisomy 21, Prenat. Diagn. 32 (2012) 644–648.
[48] L.V. Adams, B.N. Kreiswirth, R.D. Arbeit, H. Soini, L. Mtei, et al, Molecular epidemiology of HIV-associated tuberculosis in Dar es Salaam, Tanzania: strain predominance, clustering, and polyclonal disease, J. Clin. Microbiol. 50 (2012) 2645–2650.
[49] D. Alland, G.E. Kalkut, A.R. Moss, R.A. McAdam, J.A. Hahn, et al, Transmission of tuberculosis in New York City. An analysis by DNA fingerprinting and conventional epidemiologic methods, N. Engl. J. Med. 330 (1994) 1710–1716.
[50] J.F. Reyes, A.R. Francis, M.M. Tanaka, Models of deletion for visualizing bacterial variation: an application to tuberculosis spoligotypes, BMC Bioinformatics 9 (2008) 496.
[51] M.M. Tanaka, A.R. Francis, Detecting emerging strains of tuberculosis by using spoligotypes, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 15266–15271.
[52] P. Supply, E. Mazars, S. Lesjean, V. Vincent, B. Gicquel, et al, Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome, Mol. Microbiol. 36 (2000) 762–771.
[53] E. Savine, R.M. Warren, G.D. van der Spuy, N. Beyers, P.D. van Helden, et al, Stability of variable-number tandem repeats of mycobacterial interspersed repetitive units from 12 loci in serial isolates of *Mycobacterium tuberculosis*, J. Clin. Microbiol. 40 (2002) 4561–4566.
[54] H. van Deutekom, P. Supply, P.E. de Haas, E. Willery, S.P. Hoijng, et al, Molecular typing of *Mycobacterium tuberculosis* by mycobacterial interspersed repetitive unit-variable-number tandem repeat analysis, a more accurate method for identifying epidemiological links between patients with tuberculosis, J. Clin. Microbiol. 43 (2005) 4473–4479.
[55] R. Stavrum, M. Mphahlele, K. Ovreas, T. Muthivhi, P.B. Fourie, et al, High diversity of *Mycobacterium tuberculosis* genotypes in South Africa and preponderance of mixed infections among ST53 isolates, J. Clin. Microbiol. 47 (2009) 1848–1856.
[56] R. Stavrum, V.P. Myneedu, V.K. Arora, N. Ahmed, H.M. Grewal, In-depth molecular characterization of *Mycobacterium tuberculosis* from New Delhi–predominance of drug resistant isolates of the 'modern' (TbD1) type, PLoS One 4 (2009) e4540.

Supplemental Table S1. Detailed results obtained including demographic, epidemiologic, drug-resistance, and genotyping information on a total of 134 *M. tuberculosis* strains isolated in Iraq.

Year	Strain	Sex/ Age	Isolation city	Spoligotype Description	Octal code	SIT*	Lineage**	MIRU-15	15-loci MIT***	Drug-R code	Specimen	Smear micros-copy	History of TB treatment
2011	2818	F/28	Baghdad-Sadr City 17		703777400001771	22	CAS1-Delhi	274653422455524	Orphan	2	Sputum	+	Retreatment
2011	3030	M/25	Diwaniya Province		703777400001771	22	CAS1-Delhi	25475342245-424	Orphan	2	Sputum	–	Unkown
2011	3540	F	Baghdad		703777400001771	22	CAS1-Delhi	27475-422455524	Orphan	2	Sputum	+	Retreatment
2011	3766	F/31	Baghdad- SE		703777400001771	22	CAS1-Delhi	27475342245-424	Orphan	2	Sputum	+	Retreatment
2011	4136	M/63	Baghdad- SE		703777400001771	22	CAS1-Delhi	264753-22455425	Orphan	3 (INH, SM)	Sputum	+	Retreatment
2011	4870	F/24	Salah Aldin Province		703777400001771	22	CAS1-Delhi	27475342246-424	Orphan	2	Sputum	+	Retreatment
2011	4877	M/40	Baghdad-Sadr City		703777400001771	22	CAS1-Delhi	274753422455424	Orphan	2	Sputum	+	Unkown
2011	5100	M	Baghdad- Central		703777400001771	22	CAS1-Delhi	27475-42245-424	Orphan	2	Sputum	+	Retreatment
2011	5242	F/50	Baghdad- NE		703777400001771	22	CAS1-Delhi	26475-42245-424	Orphan	0	Sputum	+	Unkown
2011	5452	M/34	Babil Province		703777400001771	22	CAS1-Delhi	284753422455424	Orphan	1	Sputum	+	Retreatment
2011	5536	M/40	Diwaniya Province		703777400001771	22	CAS1-Delhi	26475-422455424	Orphan	2	Sputum	+	new
2011	5932	F/55	Baghdad- SE		703777400001771	22	CAS1-Delhi	25273-222344424	Orphan	0	Bronchial wash	–	Unkown
2011	2447	F/1	Baghdad- Suburb-east		703777740003171	25	CAS1-Delhi	254553222485424	Orphan	0	Sputum	–	Unkown
2011	3240	M/8	Baghdad- Suburb-SW		703777740003171	25	CAS1-Delhi	23455322-473322	Orphan	0	Sputum	unknown	Unkown
2011	3275	M/30	Salah Aldin Province		703777740003171	25	CAS1-Delhi	254753222485424	Orphan	0	Sputum	–	Unkown
2011	3425	M/30	Musol Province		703777740003171	25	CAS1-Delhi	26355422-665324	Orphan	2	Sputum	–	Retreatment
2011	3543	M/12	Baghdad-North		703777740003171	25	CAS1-Delhi	23455322-482424	Orphan	3 (SM)	Sputum	+	Retreatment
2011	3862	M/40	Kerbala Province		703777740003171	25	CAS1-Delhi	26355422266-324	Orphan	2	Sputum	+	Retreatment
2011	4068	F/30	Diyala- Province		703777740003171	25	CAS1-Delhi	234553222475325	Orphan	2	Sputum	+	new
2011	4610	M/30	Nasiriya Province		703777740003171	25	CAS1-Delhi	244853222495624	Orphan	2	Sputum	–	Retreatment
2011	5057	F/20	Baghdad-Sadr City		703777740003171	25	CAS1-Delhi	23455322-472324	Orphan	0	Pleural Fluid	–	Unkown
2011	5068	F/30	Baghdad- North		703777740003171	25	CAS1-Delhi	234553122485324	Orphan	2	Sputum	+	Retreatment
2011	5194	M/1	Baghdad-Sadr City 4		703777740003171	25	CAS1-Delhi	234563222472324	232	0	Gastric wash	–	Unkown
2011	5291	M/60	Wasit Province		703777740003171	25	CAS1-Delhi			0	Sputum	+	Unkown
2011	5305	M/15	Baghdad- North		703777740003171	25	CAS1-Delhi	23455322-483424	Orphan	0	Sputum	+	new
2011	5508	F/27	Baghdad-Sadr City		703777740003171	25	CAS1-Delhi	234563222472324	232	3 (INH, SM)	Sputum	+	new
2011	5849	M/35	Baghdad- Central		703777740003171	25	CAS1-Delhi	23456322-472324	Orphan	0	Bronchial wash	–	Retreatment
2011	5964	M/27	Baghdad- Suburbs-SE		703777740003171	25	CAS1-Delhi	224531242362223	Orphan	0	Sputum	+	Unkown
2011	6080	F/18	Baghdad-Suburb- NE		703777740003171	25	CAS1-Delhi	263554-2266-324	Orphan	0	CSF	–	Unkown
2011	6139	M/25	Baghdad central		703777740003171	25	CAS1-Delhi	254553222475325	Orphan	2	Sputum	+	Retreatment
2011	6159	M/25	Baghdad- Central		703777740003171	25	CAS1-Delhi	26435332-565324	Orphan	0	Sputum	+	Unkown
2011	3987	M/60	Baghdad-North		703777740003771	26	CAS1-Delhi	253753222495525	Orphan	3 (INH)	Sputum	+	Retreatment
2011	4038	M/37	Mosul Province		703777740003771	26	CAS1-Delhi	264853222455544	Orphan	2	Sputum	+	Retreatment
2011	4527	M	Baghdad-Sadr City		703777740003771	26	CAS1-Delhi	253753222485624	231	1	Sputum	+	new
2011	4874	F/20	Baghdad- Suburb-South		703777740003771	26	CAS1-Delhi	253753222485624	231	2	Sputum	+	Retreatment
2011	4882	F/25	Baghdad-West		703777740003771	26	CAS1-Delhi	253753522485624	Orphan	0	Bronchial wash	–	Unkown
2011	4885	M/38	Baghdad- NE		703777740003771	26	CAS1-Delhi	253753322485624	Orphan	0	Pleural Fluid	–	Unkown
2011	5886	M/38	Baghdad- Central		703777740003771	26	CAS1-Delhi	253753222485624	231	1	Sputum	+	new
2011	6210	M/45	Muthana Province		703777740003771	26	CAS1-Delhi	253753222485624	231	2	Sputum	+	Retreatment
2011	5625	F/45	Baghdad- South		760001400000171	29	LAM	25333224426-321	Orphan	2	Sputum	+	Retreatment
2011	5732	F/25	Baghdad- East		760001400000171	29	LAM	25333223426-321	Orphan	2	Sputum	+	Unkown
2011	5575	F/59	Anbar Province		776377777760771	34	S	434534442152423	Orphan	2	Sputum	–	Retreatment
2011	5492	M/35	Baghdad- SE		777737777720771	36	H3	254533336292243	Orphan	0	Sputum	+	Unkown
2011	4588	M/30	Nasiriya Province		777737777760771	37	T3	234324443372225	Orphan	2	Sputum	–	Retreatment
2011	6179	M/54	Arbil Province		777737777760771	37	T3	234433542373226	Orphan	0	Sputum	+	new
2011	4257	M/31	Mosul Province		777777404760771	41	Turkey	1511333-2343443	Orphan	3 (INH, SM, EMB)	Sputum	+	new
2011	5439	M/44	Baghdad- SE		777777404760771	41	Turkey	151133342343443	Orphan	2	Sputum	+	Retreatment
2011	5562	M/50	Mosul Province		777777404760771	41	Turkey	151132342363443	Orphan	2	Sputum	+	Retreatment

2011	5699	F/19	Misan Province		777777404760771	41	Turkey	151132342363445	Orphan	2	Sputum	–	Retreatment
2011	5282	M/35	Baghdad- Suburb-SE		777777770000000	46	Unknown	222638343222225	Orphan	2	Sputum	+	Retreatment
2011	2352	M/45	Baghdad- Suburb-South		777777774020771	47	H1	253533337433244	Orphan	0	Urine	–	Unkown
2011	2485	F/45	Babil Province		777777774020771	47	H1	253533337333244	Orphan	0	Bronchial wash	unknown	Unkown
2011	2849	F/31	Baghdad-Sadr City 79		777777774020771	47	H1	253532334363244	Orphan	0	Sputum	+	Unkown
2011	3413	F/45	Baghdad-Sadr City		777777774020771	47	H1	24353333437324-	Orphan	0	Sputum	+	Unkown
2011	5917	M/60	Wasit Province		777777777720731	49	H3	243533333372244	Orphan	2	Sputum	–	Retreatment
2011	3234	M/40	Nasiriya Province		777777777720771	50	H3	163534326365243	Orphan	2	Sputum	+	Retreatment
2011	4305	M/35	Najaf Province		777777777720771	50	H3	273534325263243	Orphan	2	Sputum	+	Retreatment
2011	4846	M/49	Kerbala-Province		777777777720771	50	H3	243534337363244	Orphan	2	Sputum	+	Retreatment
2011	5204	M/32	Baghdad-Sadr City 47		777777777720771	50	H3	26353432736-244	Orphan	2	Sputum	–	Retreatment
2011	5511	M/50	Mosul Province		777777777720771	50	H3	263534327363244	Orphan	2	Sputum	+	Retreatment
2011	6191	M/33	Mosul Province		777777777720771	50	H3	24353-337363244	Orphan	2	Sputum	+	Retreatment
2011	2674	F/24	Baghdad- West		777777777760771	53	T1	232634344253224	Orphan	0	Sputum	+	Unkown
2011	4469	F/9	Diyala Province		777777777760771	53	T1	25212224235-245	Orphan	0	pus	+	Unkown
2011	4478	M/55	Nasiriya Province		777777777760771	53	T1	252122242354246	Orphan	2	Sputum	+	Retreatment
2011	4603	M/4	Baghdad- Suburbs-east		777777777760771	53	T1	232434244253223	Orphan	0	Gastric wash	–	Unkown
2011	2259	F/75	Baghdad- NE		777777777763771	54	MANU2	----34---------	Orphan	0	Bronchial wash	–	Unkown
2011	2615	M/37	Baghdad- NE		777777777763771	54	MANU2	2----4------22-	Orphan	0	Bronchial wash	–	Unkown
2011	3895	M/19	Diyala Province		777777777763771	54	MANU2	----34------2--	Orphan	0	Bronchial wash	–	Unkown
2011	3977	M/62	Kerbala Province		777777777763771	54	MANU2	--5---------2--	Orphan	0	Bronchial wash	–	Retreatment
2011	5022	M/70	Baghdad- NE		777777777763771	54	MANU2	---------1-----	Orphan	2	Sputum	+	Retreatment
2011	6423	M/65	Baghdad- Central		777777777763771	54	MANU2	----------	Orphan	0	Bronchial wash	+	Unkown
2011	4292	F/25	Baghdad-Sadr City		777777774020731	62	H1	254533336384435	Orphan	0	Bronchial wash	–	Unkown
2011	6468	M/35	–		777777607560771	64	LAM6	24252514237-313	Orphan	0	Sputum	+	Unkown
2011	2389	F/17	Baghdad- Suburb-South		777777737760771	86	T1	22363-3-5292323	Orphan	2	Sputum	–	Retreatment
2011	5618	M/40	Baghdad-Sadr City 3		777703777760771	102	T	223422343263225	Orphan	2	Sputum	–	Retreatment
2011	3271	M/25	Wasit Province		577777777420771	127	H4 (Ural-2)	224631342562223	Orphan	0	Sputum	–	Unkown
2011	4752	M/26	Wasit- Province		577777777420771	127	H4 (Ural-2)	22463134256-223	Orphan	2	Sputum	–	Retreatment
2011	5895	M/30	Baghdad- SE		577777777420771	127	H4 (Ural-2)	224533342532323	Orphan	0	Sputum	+	Unkown
2011	5981	F/45	Baghdad-Suburb- east		577777777420771	127	H4 (Ural-2)	224533342562223	Orphan	2	Sputum	–	Retreatment
2011	2806	M/59	Diwaniya Province		703777740003471	247	CAS1-Delhi	244754322492624	Orphan	2	Sputum	+	Retreatment
2011	3259	F/56	Baghdad- NE		703777740003471	247	CAS1-Delhi	24475422249-624	Orphan	2	Sputum	–	Retreatment
2011	3429	M/53	Baghdad-Sadr City 25		703777740003471	247	CAS1-Delhi	244754222492624	229	2	Sputum	+	Retreatment
2011	4748	M/40	Baghdad- Suburb-east		703777740003471	247	CAS1-Delhi	244753322472624	Orphan	0	Sputum	–	Retreatment
2011	6014	M/53	Baghdad-Sadr City 546		703777740003471	247	CAS1-Delhi	244754222492624	229	2	Sputum	+	Retreatment
2011	2260	M/30	Diwaniya Province		037637777760771	284	T1	244632432232224	Orphan	2	Sputum	–	Retreatment
2011	2360	F/27	Babil Province		037637777760771	284	T1	333532432333125	Orphan	2	Sputum	–	Retreatment
2011	2824	F/51	Baghdad- West		037637777760771	284	T1	243632432333125	Orphan	0	Sputum	+	Unkown
2011	3313	M/35	Sulaimaniya Province		037637777760771	284	T1	243542432333125	228	2	Sputum	+	Retreatment
2011	4081	F/57	Baghdad-Sadr City		037637777760771	284	T1	243532432333125	Orphan	2	Sputum	+	new
2011	4689	M	Sulaimaniya- Province		037637777760771	284	T1	243542432333125	228	0	Sputum	+	Unkown
2011	5448	M/63	Diyala Province		037637777760771	284	T1	233532452363125	Orphan	0	Bronchial wash	–	Unkown
2011	6255	F/21	Basra Province		777777777360771	516	T1			0	Sputum	–	Unkown
2011	5263	M/43	Baghdad-east		777777777760571	520	T1	252122442354228	Orphan	2	Sputum	+	Retreatment
2011	6158	F/27	Baghdad central		703777740003111	704	CAS1-Delhi	264653322665324	Orphan	0	Bronchial wash	+	Unkown
2011	5979	M/31	Diwaniya Province		777777777420771	777	H3/H4 (Ural-1)	223533342462224	Orphan	2	Sputum	+	new
2011	3705	F/18	Baghdad-Suburb- NE		777776777760571	878	X1	254533332341543	Orphan	0	pus	+	Unkown
2011	5159	F/35	Babil Province		703677740003771	954	CAS1-Delhi	264753422485424	Orphan	2	Sputum	+	Retreatment
2011	2478	M/38	Basra Province		777777600760771	1144	T1			2	Sputum	+	Retreatment
2011	2834	M/60	Baghdad- West		777777600760771	1144	T1	252734244333424	222	2	Sputum	–	Retreatment
2011	3178	M/45	Baghdad- SE		777777600760771	1144	T1	252734144333424	Orphan	0	Sputum	+	Unkown
2011	3233	M/27	Nasiriya Province		777777600760771	1144	T1	243734244333424	223	2	Sputum	–	Retreatment
2011	3980	M/25	Najaf Province		777777600760771	1144	T1	242734244333424	230	2	Sputum	+	Retreatment
2011	4728	M/55	Baghdad- West		777777600760771	1144	T1	252734244333424	222	0	Sputum	+	Unkown
2011	4826	M/34	Kerbala-Province		777777600760771	1144	T1			2	Sputum	+	Retreatment
2011	4863	F/12	Baghdad- Suburb- east		777777600760771	1144	T1	242734244333424	230	2	Sputum	+	Retreatment

2011	5198	M/24	Baghdad-Sadr City		777777600760771	1144	T1	242734244333424	230	2	Sputum	+	Retreatment
2011	5220	M/45	Baghdad- NE		777777600760771	1144	T1	243734244333424	223	2	Sputum	+	Retreatment
2011	5473	M/34	Kerbala Province		777777600760771	1144	T1	252734244333424	222	2	Sputum	+	Retreatment
2011	5489	F/20	Baghdad- east		777777600760771	1144	T1	242734244333424	230	2	Sputum	+	Retreatment
2011	5546	M/30	Baghdad- east		777777600760771	1144	T1	242734244333424	230	2	Sputum	–	Retreatment
2011	5700	F/45	Misan Province		777777600760771	1144	T1	242734244333424	230	2	Sputum	+	Retreatment
2011	5769	F/25	Mosul Province		777777600760771	1144	T1	243734244333424	223	2	Sputum	+	Retreatment
2011	5844	M/20	Baghdad- Central		777777600760771	1144	T1	252734244333424	222	2	Sputum	+	Retreatment
2011	5934	M/34	Basra Province		777777600760771	1144	T1	242734244333424	230	2	Sputum	+	new
2011	6010	M/57	Baghdad-Sadr City		777777600760771	1144	T1	243734244332523	Orphan	2	Sputum	–	Retreatment
2011	6166	M/32	Basra Province		777777600760771	1144	T1	242734244333424	230	2	Sputum	–	new
2011	6223	F/50	Basra Province		777777600760771	1144	T1			2	Sputum	+	Retreatment
2011	2397	M/40	Baghdad- Suburb-South		703737740003171	1198	CAS1-Delhi	264353622475324	221	0	Sputum	+	Unkown
2011	4183	F/16	Baghdad-Sadr City 41		703737740003171	1198	CAS1-Delhi	264353622475324	221	0	Bronchial wash	–	Unkown
2011	4686	M/35	Baghdad-Centre		703737740003171	1198	CAS1-Delhi	264353622475324	221	3 (INH, SM)	Sputum	+	Retreatment
2011	5388	M/40	Baghdad- NE		703737740003171	1198	CAS1-Delhi	264353622475324	221	3 (INH)	Sputum	+	Retreatment
2011	5514	F/33	Diyala Province		703737740003171	1198	CAS1-Delhi	264353622475324	221	0	Pleural Fluid	–	Unkown
2011	5928	M/62	Baghdad- West		703737740003171	1198	CAS1-Delhi	264353622474324	Orphan	1	Sputum	+	Retreatment
2011	3672	M/58	Nasiriya Province		703777740000171	1789	CAS	26574342258-424	Orphan	0	Sputum	+	Retreatment
2011	5790	M/42	Baghdad- NE		703777640003771	2145	CAS1-Delhi	266833432465544	Orphan	2	Sputum	+	Retreatment
2011	2483	F/9	Baghdad-Sadr City 191		777700077760771	2230	T1	23253234524232-	Orphan	0	Sputum	+	Unkown
2011	3814	M/1	Baghdad-Sadr City 17		703777740003761	3371	CAS1-Delhi	24595442258-623	Orphan	0	Gastric wash	–	new
2011	6216	M	Baghdad-Sadr City		703777740003761	3371	CAS1-Delhi	266953432475623	Orphan	2	Sputum	+	Retreatment
2011	4039	M/26	Misan Province		037637777760731	3776*	T2	234532453372224	Orphan	1	Sputum	+	new
2011	3284	M/60	Salah Aldin Province		777717777760371	Orphan	T1	233532252252224	Orphan	0	Sputum	–	Unkown
2011	5760	M/60	Diyala Province		713777777763171	Orphan	MANU2	2----------4	Orphan	0	Bronchial wash	–	Unkown
2011	5953	M/60	Baghdad- West		777777740763771	Orphan	MANU2	25675343249-623	Orphan	0	Bronchial wash	+	Unkown
2011	6005	M/25	Anbar Province		177777777420671	Orphan	H4 (Ural-2)	224633342572323	Orphan	2	Sputum	+	Retreatment
2011	6320	M/58	Baghdad- West		703777740003161	Orphan	CAS1-Delhi	26655343247-423	Orphan	0	Sputum	–	Unkown

* SIT marked by an asterix (SIT3776*) was newly created after a match with an orphan from Iran, and is highlighted in yellow. Note that spoligotype orphans are highlighted in blue.
** Lineage designation is based on signatures provided in SpolDB4 (Brudey et al., 2006). Note that new lineage designations since SpolDB4 are highlighted in mauve.
*** 15-loci MITs highlighted in ochre color were newly created after match with another orphan in the database, or due to 2 or more strains belonging to a new pattern within this study.

Supplemental Table S2. Description of major *M. tuberculosis* clusters in Iraq (≥ 4 isolates) as compared to neighboring countries, and their worldwide distribution in the database.

SIT (Clade) Octal Number Spoligotype Description	n/t (%) in Iraq	% in Iraq vs. database	Distribution in neighboring countries *			Worldwide distribution in SITVIT2 database *	
			Iran n/t (%)	Saudi Arabia n/t (%)	Turkey n/t (%)	Distribution in Regions with ≥ 3% of a given SIT *	Distribution in countries with ≥ 3% of a given SIT **
1144 (T1) 777777600760771 ■■■■■■■■■■■■■■■■■■■■□□□□□□□■■■■■□□□■■■■■■■	**20/134 (14.93) §**	83.33	0/565 (0)	0/1818 (0)	0/1187 (0)	ASIA-W 83.33, AMER-N 8.33, EURO-S 4.17, AMER-S 4.17	IRQ 83.33, USA 8.33, VEN 4.17, GRC 4.17
25 (CAS1-Delhi) 703777740003171 ■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□■■□□■■■■■	**19/134 (14.18) †**	3.49	17/565 (3)	94/1818 (5.2)	2/1187 (0.2)	AFRI-E 27.71, ASIA-W 21.28, AFRI-N 15.41, ASIA-S 12.11, AMER-N 9.54, EURO-N 5.14, EURO-W 4.95	ETH 26.42, SAU 17.25, SDN 12.66, USA 9.54, IND 6.97, IRQ 3.49, IRN 3.12
22 (CAS1-Delhi) 703777400001771 ■■■□□□□■■■■■■■■■■■■□□□□□□□□□□□□□□□■■■■■■■■	**12/134 (8.96) †**	17.14	6/565 (1.1)	30/1818 (1.7)	0/1187 (0)	ASIA-W 61.43, ASIA-S 14.29, AFRI-E 8.57, EURO-N 5.71, EURO-W 4.29	SAU 42.86, IRQ 17.14, IRN 8.57, TZA 4.29
26 (CAS1-Delhi) 703777740003771 ■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□■■■■■■■■■	8/134 (5.97)	0.6	**62/565 (11) †**	78/1818 (4.3)	0/1187 (0)	ASIA-S 53.15, AMER-N 16.44, ASIA-W 6.91, EURO-W 6.38, AFRI-E 5.33, EURO-N 4.35, EURO-S 3.75	IND 31.23, USA 16.44, PAK 10.81, SAU 5.86, BGD 5.63, IRN 4.66, ITA 3.6, NLD 3.23, ETH 3.23
284 (T1) 037637777760771 □□□□■■■■■■■□□■■■■■■■■■■■■■■■■■□□□□■■■■■■■	**7/134 (5.22) †**	4.19	4/565 (0.7)	23/1818 (1.3)	35/1187 (3)	ASIA-W 39.52, EURO-E 23.95, EURO-W 16.77, EURO-N 7.78, EURO-S 3.59, ASIA-S 2.99	BGR 23.95, TUR 20.96, SAU 13.77, SWE 6.59, AUT 5.99, IRQ 4.19, FXX 3.59
50 (H3) 777777777720771 ■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■	6/134 (4.48)	0.18	0/565 (0)	25/1818 (1.4)	**65/1187 (5.5) †**	AMER-N 18.83, EURO-W 18.04, AMER-S 17.56, EURO-S 11.86, EURO-E 5.67, EURO-N 4.4, AFRI-N 4.37, AFRI-S 4.15, CARI 3.58, ASIA-W 3.03	USA 18.01, BRA 7.25, FXX 7.07, AUT 6.25, ITA 5.58, ESP 5.58, PER 4.22, ZAF 4.15, CZE 3.76
54 (Manu2) 777777777763771 ■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□■■■■■■■■■	**6/134 (4.48) †**	2.64	1/565 (0.2)	20/1818 (1.1)	2/1187 (0.2)	ASIA-E 16.3, AFRI-N 15.86, ASIA-W 14.1, ASIA-S 13.22, AMER-N 10.57, AFRI-S 7.05, ASIA-N 4.41, AMER-S 3.97	CHN 16.3, EGY 14.98, IND 12.34, USA 8.81, SAU 8.81, ZAF 7.05, RUS 4.41
1198 (CAS1-Delhi) 703737740003171 ■■■□□□□■■■■■□■■■■■■■■□□□□□□□□□□□□□■■□□■■■■■	**6/134 (4.48) †**	17.65	6/565 (1.1)	3/1818 (0.2)	0/1187 (0)	ASIA-W 26.47, AFRI-N 23.53, ASIA-S 17.65, EURO-N 8.82, EURO-W 5.88, AMER-N 5.88, AFRI-S 5.88, AFRI-E 5.88,	SDN 23.53, IRQ 17.65, IRN 17.65, SAU 8.82, ZAF 5.88, USA 5.88, GBR 5.88, ETH 5.88
247 (CAS1-Delhi) 703777740003471 ■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□■■■□□■■■■	**5/134 (3.73)**	14.71	0/565 (0)	10/1818 (0.6)	0/1187 (0)	ASIA-W 44.12, AFRI-E 35.29, ASIA-S 8.82, AMER-N 5.88	ETH 32.35, SAU 29.41, IRQ 14.71, USA 5.88, IND 5.88
41 (Turkey) 777777404760771 ■■■■■■■■■■■■■■■■■■■□□□□□■□□■■■■■□□□■■■■■■■	4/134 (2.99)	0.84	0/565 (0)	7/1818 (0.4)	**212/1187 (17.9) †**	ASIA-W 47.26, EURO-E 21.52, EURO-W 12.45, AFRI-E 5.49, EURO-N 5.06, EURO-S 4.01	TUR 44.73, BGR 20.46, NLD 5.91, ETH 5.49, FXX 3.59
47 (H1) 777777774020771 ■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□□□■■■■■■■	**4/134 (2.99)**	0.27	0/565 (0)	13/1818 (0.7)	35/1187 (2.9)	EURO-W 21.57, AMER-N 18.18, EURO-S 14.32, AMER-S 12.01, EURO-E 7.67, EURO-N 6.72, ASIA-W 3.94, AFRI-N 3.87	USA 16.15, ITA 8.75, AUT 8.55, BRA 7.46, FXX 7.06, CZE 4.0, ESP 3.8, SWE 3.6
53 (T1) 777777777760771 ■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	4/134 (2.99)	0.07	21/565 (3.7)	157/1818 (8.6)	**218/1187 (18.4) †**	AMER-N 16.94, EURO-W 16.23, AMER-S 12.37, EURO-S 9.76, ASIA-W 7.13, EURO-N 5.53, AFRI-S 5.15, AFRI-E 4.67, ASIA-E 4.39, AFRI-N 3.65, EURO-E 3.39	USA 13.69, FXX 8.17, ITA 5.53, BRA 5.32, ZAF 5.04, TUR 3.6, AUT 3.55, CHN 3.2
127 (H4 / Ural-2) 577777777420771 ■□■■■■■■■■■■■■■■■■■■■■■■■■■□□□■□□□□■■■■■■■	4/134 (2.99)	2.13	**96/565 (17) †**	12/1818 (0.7)	1/1187 (0,1)	ASIA-S 55.85, EURO-W 17.02, ASIA-W 9.04, EURO-N 5.85, ASIA-E 5.32	IRN 51.06, NLD 7.98, SAU 6.38, CHN 5.32, SWE 4.79, AUT 4.79, DEU 3.72

***** Note that distribution in Iraq versus neighboring countries is **highlighted in bold** for the highest proportion of a given SIT.

****** Worldwide distribution is reported for regions with more than 3% of a given SIT. Definition of macro-geographical regions and sub-regions according to the United Nations (http://unstats.un.org/unsd/methods/m49/m49regin.htm); Regions: AFRI (Africa), AMER (Americas), ASIA (Asia), EURO (Europe), and OCE (Oceania), subdivided in: E (Eastern), M (Middle), C (Central), N (Northern), S (Southern), SE (South-Eastern), and W (Western). Furthermore, CARIB (Caribbean) belongs to Americas, while Oceania is subdivided in 4 sub-regions, AUST (Australasia), MEL (Melanesia), MIC (Micronesia), and POLY (Polynesia). Note that Russia has been attributed a new sub-region by itself (Northern Asia) instead of including it among rest of the Eastern Europe. It reflects its geographical localization as well as due to the similarity of specific TB genotypes circulating in Russia (a majority of Beijing genotypes) with those prevalent in Central, Eastern and South-Eastern Asia. The 3 letter country codes are according to http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3; countrywide distribution is only shown for SITs with ≥5% of a given SITs as compared to their total number in the SITVIT2 database.

§ Exclusively present in Iraq in this comparison; high phylogeographical specificity for Iraq in SITVIT2 database.

† Difference statistically significant, p<0.001