Hindawi Publishing Corporation
BioMed Research International
Volume 2014, Article ID 580981, 15 pages
http://dx.doi.org/10.1155/2014/580981

Research Article

Molecular Epidemiology and Genotyping of *Mycobacterium tuberculosis* Isolated in Baghdad

Ruqaya Mustafa Ali,[1] Alberto Trovato,[2] David Couvin,[3] Amina N. Al-Thwani,[1] Emanuele Borroni,[2] Fahim H. Dhaer,[4] Nalin Rastogi,[3] and Daniela M. Cirillo[2]

[1] *Genetic Engineering and Biotechnology Institute for Post Graduate Studies, University of Baghdad, 10070 Jadriyah, Baghdad, Iraq*
[2] *Emerging Bacterial Pathogens Unit, San Raffaele Scientific Institute (HSR), via Olgettina, 60 20132 Milano, Italy*
[3] *Institut Pasteur de la Guadeloupe, Abymes, 97183 Guadeloupe, France*
[4] *Ministry of Agriculture, Al Nidhal Street, Baghdad 5157, Iraq*

Correspondence should be addressed to Daniela M. Cirillo; cirillo.daniela@hsr.it

Received 7 November 2013; Revised 17 January 2014; Accepted 18 January 2014; Published 26 February 2014

Academic Editor: Tomasz Jagielski

Tuberculosis (TB) remains a major health problem in Iraq but the strains responsible for the epidemic have been poorly characterized. Our aim was to characterize the TB strains circulating in Bagdad (Iraq). A total of 270 *Mycobacterium tuberculosis* complex (MTBC) strains isolated between 2010 and 2011 from TB patients attending the Center of Chest and Respiratory diseases in Baghdad were analyzed by Spoligotyping. The analysis indicated that 94.1% of the isolates belong to known genotype clades: CAS 39.6%, ill-defined T clade 29.6%, Manu 7.4%, Haarlem 7%, Ural 4.1%, LAM 3.3%, X 0.7%, LAM7-TUR 0.7%, EAI 0.7%, S 0.7%, and unknown 5.9%. Comparison with the international multimarker database SITVIT2 showed that SIT 309 (CAS1-Delhi) and SIT1144 (T1) were the most common types. In addition, 44 strains were included in SITVIT2 database under 16 new Spoligotype International Types (SITs); of these, 6 SITs (SIT3346, SIT3497, SIT3708, SIT3790, SIT3791, and SIT3800) ($n = 32$ strains) were created within the present study and 10 were created after a match with an orphan in the database. By using 24-loci MIRU-VNTR-typing on a subset of 110 samples we found a high recent transmission index (RTI) of 33.6%. In conclusion, we present the first unifying framework for both epidemiology and evolutionary analysis of *M. tuberculosis* in Iraq.

1. Introduction

Tuberculosis (TB) is an ancient disease that currently represents an immense global health challenge. In 2011, WHO estimated that globally there were 8.7 million new cases of active TB leading to 1.4 million reported deaths [1].

According to the 2012 report of the Ministry of Health (MOH), the incidence rate of TB in Iraq was 45/100,000, with 13,860 new TB cases and 1140 of previously treated cases. The Iraqi laboratory network includes 124 district smear microscopy laboratories and one national reference laboratory located in Baghdad performing cultures and drug susceptibility testing of *M. tuberculosis* [2].

In the last decades, a large number of different molecular methods based on DNA fingerprints have been developed. The usefulness of these methods has been demonstrated primarily as epidemiological markers to discriminate the pathogen at the genus, species, and subspecies level. The level of strain differentiation is of crucial importance for the study of transmission dynamics, determining whether the infection is caused by single strain or by multiple strain and if recurrence of the disease is due to treatment failure or infection with new strain of *M. tuberculosis* [3, 4]. Spoligotyping, targeting the Direct Repeat locus and Mycobacterial Interspersed Repetitive Unit-Variable Number Tandem Repeats (MIRU-VNTRs) typing, has been shown to be a valuable alternative to IS*6110* [5, 6]. An optimized 24-loci MIRU-VNTR typing scheme has been proposed as international standard [7, 8]. In addition to their use for tracing TB transmission at the strain level, MIRU-VNTR markers are

also phylogenetically more informative, especially in the 24-locus format and can therefore be used to predict grouping into strain lineage [9, 10].

Although TB is still a public health problem in Iraq, there is little information about the genetic characteristics of the isolates driving the epidemic. A better knowledge of the molecular characteristics of *M. tuberculosis* complex isolates could contribute to better understanding of the transmission dynamics of the disease within the country and can guide interventions to control the epidemic. The aim of this study is to determine molecular epidemiology features of *M. tuberculosis* isolates in Baghdad, as well as in other Iraqi governorates, to assess their transmission dynamics.

2. Materials and Methods

2.1. Study Population. This study was conducted in Baghdad at the Center of Chest and Respiratory diseases. A total of 270 isolates were collected between 2010 and 2011 representing approximately 40% of new and previously treated TB patients. This study was approved by the local ethical committee.

2.2. Culture and Drug Susceptibility Test. Diagnostic specimens were cultured and isolated on Lowenstein-Jensen (LJ) media after decontamination. Drug Susceptibility Tests (DST) against first-line anti-TB drugs rifampicin (RFP), isoniazid (INH), streptomycin (SM), and ethambutol (EMB) were combined with LJ medium at the following concentrations: RFP 40 μg/mL, INH 0.2 μg/mL, SM 4 μg/mL, and EMB 2 μg/mL. DST was used to detect the drug resistance of *M. tuberculosis* by the proportion method [11].

2.3. Identification of Mycobacterium tuberculosis. *Mycobacterium tuberculosis* isolates were identified on the basis of colony morphology, growth rate, pigmentation properties, niacin accumulation, nitrate reduction, thiophene-2-carboxylic acid hydrazide (TCH), and para-nitrobenzoic acid (PNB) test [11].

2.4. DNA Extraction. DNA was extracted from cultures by the standard Cetyltrimethyl ammonium-bromide (CTAB) method. The DNA was stored in TE buffer (10 mM Tris, 1 mM EDTA) at −20°C until use [12].

2.5. Spoligotyping. All genotyping methods were performed at the Emerging Bacterial Pathogens Unit, WHO/IUATLD Supra-National Reference TB Laboratory, San Raffaele Scientific Institute (FCSR). Spoligotyping analysis was performed by using commercial kit (Ocimun Biosolutions) as described by Kamerbeek et al. [13]. The 43 spacers between the direct repeats in the target region were amplified by using DRa biotinylated at 5′ end and DRb primers. The PCR products were hybridized to a membrane containing 43 oligonucleotides by reverse line blotting. *M. tuberculosis* H37Rv and *M. bovis* BCG were used as positive controls in each run. Spoligotyping results were converted into octal code and analyzed by using the SITVIT2 proprietary database of the Pasteur Institute of Guadeloupe, which is an updated version of the previously released SpolDB4 and SITVITWEB databases [14, 15].

2.6. MIRU-VNTR. Standardized 24-loci MIRU-VNTR typing [7] was performed using the MIRU-VNTR typing kit (Genoscreen, Lille, France). PCR products were run with Genescan 1200LIZ size standard (Applied Biosystems, California, and USA) on ABI3730 sequencer. Sizing of the PCR fragments and assignment of MIRU-VNTR alleles were done by Gene Mapper software version 3.7 (Applied Biosystems, California, USA). In order to deine clusters and to build an UPMGA tree, we used the MIRU-VNTRplus web application available at http://www.miru-vntrplus.org/MIRU/index.faces. The allelic diversity of the strains was determined by using the Hunter Gaston Discriminatory Index (HGDI).

2.7. Interpretation of Typing Results. Spoligotypes in binary format and MIRU patterns in 24-digit codes were entered in the SITVIT2 database. At the time of the comparison, it contained genotyping data on more than 100,000 MTBC strains isolated from 160 countries of patient origin. In this database, "SIT" (Spoligotype International Type) designates spoligotyping shared by two or more patient isolates, "MIT" (MIRU International Type) designates 24-loci MIRU patterns shared by two or more patient isolates, and "orphan" designates patterns reported for a single isolate. Major spoligotyping-based phylogenetic clades were assigned according to revised signatures provided in SITVITWEB [15]. These clades include specific signatures for *M. tuberculosis* complex members and rules defining major lineages/sublineages for *M. tuberculosis* sensu stricto, that is, the Central Asian (CAS) clade and 2 sublineages, the East-African-Indian (EAI) clade and 9 sublineages, the Haarlem (H) clade and 3 sublineages, the Latin-American-Mediterranean (LAM) clade and 12 sublineages note that two LAM sublineages were recently raised to lineage level: LAM10-CAM as the Cameroon lineage [16] and LAM7-TUR as the Turkey lineage [17], the "Manu" family and 3 sublineages, the IS*6110*-low-banding X clade and 4 sublineages, and an ill-defined T clade with 5 sublineages. The recently described "Ural" family, subdivided into 2 sublineages (Ural-1 and Ural-2), replaced some spoligotype patterns previously classified as H3 and H4 [18].

2.8. Phylogenetic Analysis. The evolutionary relationships among all the observed spoligotypes were studied by drawing minimum spanning trees (MSTs) with BioNumerics Software version 6.6 (Applied Maths, Sint-Martens-Latem, Belgium).

3. Results

3.1. Studied Population. This study was performed over a period of 13 months on 270 *M. tuberculosis* complex strains from Iraqi patients. The related demographic information and drug resistance patterns obtained are summarized in Table 1. One hundred and fifty-seven (58.14%) of the cases included in the study were from Baghdad, 113 (41.85%) were from other governorates in Iraq but diagnosed in Baghdad.

Table 1: Demographic characteristics of the population studied and drug resistance profiles.

Parameters	Origin of patients		
	Baghdad (%)	Other governorates (%)	Total
Total no. of strain	157 (58.1)	113 (41.9)	270
Sex			
Male	94 (59.9)	85 (75.2)	179
Female	63 (40.1)	28 (24.8)	91
Sex ratio M : F	1.49	3.04	1.97
Age group (yr) (no.)			
0–20	17 (10.8)	4 (3.5)	21
21–40	82 (52.2)	71 (62.8)	153
41–60	48 (30.6)	25 (22.1)	73
>60	10 (6.4)	13 (11.5)	23
Smear microscopy (no.)			
Positive	98 (62.4)	69 (61.1)	167 (61.8)
Negative	59 (37.6)	44 (38.9)	103 (38.2)
Treatment			
New cases	119 (75.8)	68 (60.2)	187 (69.3)
P.T.C	38 (24.2)	45 (39.8)	83 (30.7)
Drug susceptibility			
Susceptible to all drugs	102 (65)	69 (61.1)	171 (63.3)
MDR (INH – RIF)	31 (19.8)	33 (29.2)	64 (23.7)
Unknown	24 (15.3)	11 (9.7)	35 (13.0)

As shown in the table, male patients are predominants (male to female sex ratios varied from 1.49 to 3.04 depending on the origin of the patients, from Baghdad or other governorates, resp.) and 56.7% of the TB patients included in the study are between 21 and 40 years old with the mean age 36.3. As for smear microscopy, 167 tested positive and 103 were negative. 187 patients had newly TB diagnosed, while 83 patients were previously treated cases (PTC). Lastly, 171 cases were susceptible to all antituberculosis drugs, whereas 64/270 (23.7%) patients had MDR. 35 patients had an unknown resistance profile.

3.2. Distribution of Phylogenetic Clades. Spoligotyping results are summarized in Tables 2 and 3. A total of 117 different patterns were observed among 270 strains studied; a total of 53 patterns corresponding to 53 strains belonged to orphan patterns not yet reported to the SITVIT2 database (Table 2), whereas the remaining 64 patterns ($n = 217$) corresponded to SITs, of which 48 patterns ($n = 173$ strains) matched a preexisting shared type in the SITVIT2 database, while 16 SITs ($n = 44$ strains) were newly created. Of these, 6 SITs (SIT3346, SIT3497, SIT3708, SIT3790, SIT3791, and SIT3800) ($n = 32$ strains) were created within the present study and 10 were created after a match with an orphan in the database (SIT3275 from France metropolitan area, SIT3789 from Iraq, SIT3792, SIT3793, and SIT3794 from India, SIT3795 and SIT3799 from Brazil, SIT3796 from Pakistan, SIT3797 from Mexico, and SIT3798 from Iraq).

Nearly 94.1% of the isolates belonged to known genotype clades. These include, in decreasing order: CAS 39.6%, ill-defined T clade 29.6%, Manu 7.4%, Haarlem 7%, Ural 4.1%, LAM 3.3%, X 0.7%, LAM7-TUR 0.7%, EAI 0.7%, S 0.7%, and unknown 5.9%. This observation on the complex diversity of *M. tuberculosis* in Iraq was further supported by the minimum spanning tree (MST) illustrated in Figure 1, which was constructed on all isolates ($n = 270$, including 53 orphan patterns).

As shown in Figure 1, the CAS lineage strains constitute the biggest group of strains infecting Iraqi patients; the major shared types being SIT309, 26, 141, 25, 3497, and 22. At a considerable distance is seen a separate group of strains made up of smaller nodes: T lineage (SIT1144, 3346, 284, 118, 53), Haarlem (SIT75, 50), LAM (SIT42, 1470), Ural (SIT127), and Manu (SIT54). Even smaller nodes concerned other lineages such as Turkey, S, X, and the ancestral EAI lineages. Indeed, as many as 129/270 (47.7%) strains from this study belonged to Principal Genetic Group 1 (PGG1), according to the *KatG-gyrA* polymorphism-based classification of Sreevatsan et al. [19], and these strains represent all of known lineages associated with PGG1 strains (CAS, Manu, and EAI with the exception of Beijing). However, evolutionary modern PGG2/3 strains constitute a high proportion in this study such as ill-defined T, Haarlem, Ural, LAM, LAM7-TUR, S,

Table 2: Orphan strains ($n = 53$) and corresponding spoligotyping defined lineages/sublineages recorded among a total of 270 *M. tuberculosis* strains (one isolate per patient) from Iraqi patients.

Iso number	Year	Strain	Spoligotype description	Octal code	Lineage*	Label	Drug R**	Isolation city	Sex/age
IRQ02201004351	2010	4351	■■□■■■■■■■■■■■■■□■■■■■■■■□□□□■□■■■■■■■■■	677777377413771	EAI5	O01	0	Wasit	F/60
IRQ02201023876	2010	3876	■■■□□□□■■■■■■■□■■■■■■■□□□□□□□□□□□□■■□□□■■■■	703767740003071	CASI-Delhi	O02	2	Baghdad	M/29
IRQ02201015527	2010	5527	□□□□■■■■■■■■□■□■■■■■■■■■■■■■■■■□□□□■■■■■■■	037727777760771	T3	O03	1	Babil	F/29
IRQ02201003954	2010	3954	□□□□■■■■■■■■□■□■■■■■■■■■■■■■■■■□□□□■■■■■■■□	037657777741770	Unknown	O04	0	Baghdad	M/22
IRQ02201113313	2011	3313	■□□□■■■■■■■■□■□■■■■■■■■■■■■■■■■□□□□■■■■■■■	437727777760771	T3	O05	1	Sulaimaniya	M/35
IRQ02201114243	2011	4243	■□■□■■■■■■■■□□■■■■■■■■■■■■■■■■□□□□■■■■■■■	537637777760771	T1	O06	1	Baghdad	F/35
IRQ02201014888	2010	4888	□□■■■■■■■■■■■■□■■■□■■■■■■■■□□□□□□□□■■□■■■■■	177767377400671	Unknown	O07	1	Diyala	M/33
IRQ02201016952	2010	6952	■□■■■■■■■■■■■■■■■■■■■■■■■■□■□■□□□□■■■■■■■	577777777520771	H3	O08	1	Baghdad	M/35
IRQ02201017197	2010	7197	■□■■■□■■■■■■■□■□■■■■■■■■■■■■□□□□□□□□■■■■■■■	557727777400771	Unknown	O09	1	Baghdad	M/39
IRQ02201115282	2011	5282	■■■■□■■■■■■■■□■■■■■■■□□□□□□□□□□□■■□□■■■■■	757767740003171	Unknown	O10	1	Baghdad	M/35
IRQ02201115448	2011	5448	■■■■□■■■■■■■■■■■■■■■■□□□□□■■■■■□□■■■■■■■■■	757777740763771	Manu2	O11	1	Diyala	M/63
IRQ02201013167	2010	3167	■■■□□□□■■■■□□■□■■■■□□□□□□□□□□□□□□■■■■■■■■■	703627400003771	CAS	O12	1	Baghdad	F/16
IRQ02201003924	2010	3924	■■■□□□□■■■■□□■■■■■■□□□□□□□□□□□□□□■■■■■■■■■	703637400003771	CAS	O13	0	Baghdad	M/48
IRQ02201005145	2010	5145	■□□■■■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■	477767607760771	LAM1	O14	0	Najaf	F/35
IRQ02201004981	2010	4981	■□■■■□■■■■■■■□□■■■■■□□□□■■■■■■■□□□□□■■□■■□■	557747607740661	LAM9	O15	0	Dhi-qar	M/37
IRQ02201025861	2010	5861	■■□□□■■■■■■■□■■■■■■■■■■■■■■■■■□□□□■■■■■■■	637737777760771	T3	O16	2	Al-Qadisiya	M/29
IRQ02201026647	2010	6647	■■■■■■■■■■■■■■■□■■■■■□□■■■■□■■■■■□□□□■■■■■■■	777767636760771	T1	O17	2	Basra	F/29
IRQ02201015057	2010	5057	■■■■■■■■■■■■■■□■□□□□□■■■■■■■■■■■□□□□■■■■■■■	777764077760771	T1	O18	1	Baghdad	M/47
IRQ02201024897b	2010	4897b	■■□■■■■■■■■■■■■□■■■□■■■■■■■■□□□□■□■■■■■■■■■	677767377413771	EAI5	O19	2	Basra	M/21
IRQ02201024258	2010	4258	■□■□□□□■■■■■■■□■■■■■■■□□□□□□□□□□□□■■■■■■■□■	503767740003761	CASI-Delhi	O20	2	Baghdad	M/54
IRQ02201014287	2010	4287	■■■■■■■■■■■■□■□■■■■■■■■■■■■■□□□■□□□□■■■■■■■	777727777420771	H3 (Ural-1)	O21	1	Dhi-qar	F/51
IRQ02201016432	2010	6432	■■■■■■■□■■■■■■□■■■■■□□□□□□□■■■■■□□□□■■■■■■■	775767600760771	T1	O22	1	Baghdad	M/60
IRQ02201014433	2010	4433	■□■□□□□■■■■□■■□■■■■■■□□□□□□□□□□□□□□■■■■■■■■■	503667700003771	CAS	O23	1	Babil	M/23
IRQ02201016038	2010	6038	■■■■□■■■■■■■■□■■■■■■□□■■■■■■■■■□□■■■■■■■■■	757767637763771	Manu2	O24	1	Baghdad	M/28
IRQ02201013391	2010	3391	■■■■■■■■■■■■□■□■■■■■■■■■■■■□□□□□□□□■■■■■■■	777727777000771	Unknown	O25	1	Baghdad	M/31
IRQ02201113592	2011	3592	■□■□□□□■■■■■■■□■■■■□□□□□□□□□□□□□□□□■■■■■■■■	503767400001771	CAS	O26	1	Baghdad	M/21
IRQ02201113234	2011	3234	■■■■■■■■■■■■■■□■■■■■■■■■■■■■■■□■□□■□■■■■■■■	777767777722771	Manu2	O27	1	Dhi-qar	M/40
IRQ02201004179	2010	4179	■■□■■■■■■■■■■□■■■■■□□■■■■■■□■■■□□■■■■■■■■■	677767637563771	Manu2	O28	0	Baghdad	M/25
IRQ02201016082	2010	6082	■■■■■■■■■■■■■■□■■□■■□□□□■■■■■■■■□□□□■■■■■■■	777766607760771	LAM9	O29	1	Baghdad	F/24
IRQ02201103164	2011	3164	■■■■■■■■■■■■■■□■■□■■□□■■■■■■■■■■□□□□■■■■■■■	777766637760771	X1	O30	0	Baghdad	M/24
IRQ02201012729	2010	2729	■■■□□□□□■■■■■■□■■■■■■□□□□□□□□□□□□□□■■□□□■■■■	701767700003071	CAS	O31	1	Dhi-qar	M/65
IRQ02201015095	2010	5095	■□□□□□□■■■■■■■□■■■■□□□□□□□□□□□□□□□□■■■■■■■■	403767400001771	Unknown	O32	1	Najaf	M/70
IRQ02201014584	2010	4584	■□■■■■■■■■■■■□■■■■■■■■■■■■□■□■□□□□■■■■■■■	577767777520771	H3	O33	1	Baghdad	F/18
IRQ02201004326	2010	4326	■■■■■■■■■■■■■□■■■■■■■□□□□□□□□□□□□■■■■■■■■	777767770000771	Unknown	O34	0	Baghdad	M/25
IRQ02201113543	2011	3543	■□■□□□□■■■■■■□■■■■■■■□□□□□□□□□□□□■■□□□■■■■	501767740003071	CAS	O35	1	Baghdad	M/22
IRQ02201026360b	2010	6360b	■□■■■■■■■■■■■□■■■■■□□□□□□■■■■■□□□□■■■■■■■	577767600760771	T1	O36	2	Basra	M/35
IRQ02201014885	2010	4885	■□■■■■■■■■■□■■■■■■■■■■■■□■□■□□□□■■■■■■■	577737777520771	H3	O37	1	Baghdad	M/29
IRQ02201026915	2010	6915	■■■■□□□■■■■■■□■■■■■■■□□□□□□□□□□□■■□□■■■■■	743767740003171	Unknown	O38	2	Salah	M/37
IRQ02201015996	2010	5996	■■■■□■■■■■■■□■■■■□□□□□□□■■■■□□□□■■■■■■■■	757767400741771	Unknown	O39	1	Baghdad	M/20

Table 2: Continued.

Iso number	Year	Strain	Spoligotype description	Octal code	Lineage*	Label	Drug R**	Isolation city	Sex/age
IRQ02201014761	2010	4761	■□■□□	500000000000000	CAS	O40	1	Baghdad	F/56
IRQ02201113275	2011	3275	■■■□□□□■■■■■■■□■■■■□□□□□□□□□□□□□□□■■□□□■■■■	703767400003071	CAS	O41	1	Salah	M/30
IRQ02201004309	2010	4309	■□■□□□□□■■■■□■□■■■■■■■□□□□□□□□□□□□■■□□□■■■■	503767740003000	CAS1-Delhi	O42	0	Salah	M/37
IRQ02201016005	2010	6005	■■■□□□□■■■■■■■□■■■■■■■□□□□□□□□□□□□□□□□■■■■■	757767777720771	H3	O43	1	Baghdad	F/54
IRQ02201123429	2011	3429	■□■□□□□■■■■■■■□■■■■■■■□□□□□□□□□□□□■■□□□□□□□	501727740003071	CAS	O44	2	Baghdad	M/53
IRQ02201113672	2011	3672	■■■■□■■■■■■■■■□■■■■■■■■■■■■■■■□■□□□□■■■■■■■	703767740000171	CAS	O45	1	Dhi-qar	M/58
IRQ02201017297	2010	7297	■■■□□□□■■■■■□■□■■■■■■■□□□□□□□□□□□□■■□□■□□■■	703727740003111	CAS1-Delhi	O46	1	Kerbala	M/24
IRQ02201013852	2010	3852	■■■□□□□■■■■■■■□■■■■■■■□□□□□□□□□□□□■■■■□■■□■	703767740003661	CAS1-Delhi	O47	1	Baghdad	M/20
IRQ02201015137	2010	5137	■□■■■■■□■■■■■■□■■■■■■□■■■■■■■■■■□□□□■■□■■■■	575767737760671	T1	O48	1	Baghdad	M/25
IRQ02201012104	2010	2104	■■■■□■■■■■■■■■■■■■■■■■□□■□□■■■■□□□■■■■■■■■■	757777744743771	Unknown	O49	1	Dhi-qar	M/65
IRQ02201013811	2010	3811	■■■■□■■■■■■■■■□■■■■■■■■■■■■■■■■□□□■■■■■■■■■	757767777743771	Unknown	O50	1	Baghdad	F/35
IRQ02201113484	2011	3484	■□■■■■■■■■■■□■□■■■■■■■■■■■■■□□□■□□□□■■■■■■■	577727777420771	H4 (Ural-2)	O51	1	Baghdad	F/53
IRQ02201112326	2011	2326	■□■■□■■■■■■■■■□■■■■■■■■■■■■■■■■□□□■■■■■■■■■	557767777743771	Unknown	O52	1	Dhi-qar	M/56
IRQ02201016511	2010	6511	■□■■□■■□■■■■□■□■■■■■□□□□■□□■■□■□□□■■□□□■■■■	555727604643071	Unknown	O53	1	Wasit	M/30

*Lineage designations for orphan patterns were done manually as expert-based interpretations using revised SpolDB4 rules.

**Drug resistance (drug R) information is shown as 0, unknown; 1, non-MDR; 2, MDRTB, that is, combined resistance to INH − RIF (with or without resistance to other drugs); 3, any other resistance(s); 4, proven XDRTB, that is, resistance to INH + RIF + a fluoroquinolone + any 1 of 3 injectable 2nd-line drugs (capreomycin, kanamycin, amikacin).

Table 3: Description of 64 shared types (SITs; $n = 217$ isolates) and corresponding spoligotyping defined lineages/sublineages starting from a total of 270 *M. tuberculosis* strains isolated from Iraqi patients.

SIT*	Spoligotype description	Octal number	Nb in study	% in study	% in study versus database	Lineage**	Clustered versus unique patterns***
4	▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪	000000007760771	1	0.37	0.28	Unknown	Unique
20	▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪	677777607760771	1	0.37	0.12	LAM1	Unique
22	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▫▪▪▪▪▪▪▪▪	703777400001771	8	2.96	10.26	CAS1-Delhi	Clustered
25	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▫▪▪▫▫▪▪▪▪▪▪	703777740003171	10	3.7	1.77	CAS1-Delhi	Clustered
26	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▫▪▪▪▪▪▪▪▪▪▪	703777740003771	12	4.44	0.87	CAS1-Delhi	Clustered
34	▪▪▪▪▪▪▪▪▫▫▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	776377777760771	1	0.37	0.12	S	Unique
41	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▪▫▫▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777404760771	1	0.37	0.21	LAM7-TUR	Unique
42	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777607760771	1	0.37	0.03	LAM9	Unique
47	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▪▫▫▫▫▪▪▪▪▪▪▪▪	777777774020771	1	0.37	0.07	H1	Unique
50	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▫▫▫▫▪▪▪▪▪▪▪▪	777777777720771	6	2.22	0.18	H3	Clustered
53	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777777760771	5	1.85	0.08	T1	Clustered
54	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▪▪▪▪▪▪▪▪▪	777777777763771	8	2.96	3.31	Manu2	Clustered
64	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▪▪▪▪▫▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777607560771	1	0.37	0.28	LAM6	Unique
75	▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▫▪▫▫▫▫▪▪▪▪▪▪▪▪▪▪	777767777720771	6	2.22	8.7	H3	Clustered
117	▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▫▪▪▪	777767777760731	2	0.74	8	T2	Clustered
118	▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777767777760771	6	2.22	3.9	T1	Clustered
124	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▪▪▪▪▪▪▪▪	777777777700771	1	0.37	1.54	Unknown	Unique
127	▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▪▫▫▫▫▪▪▪▪▪▪▪▪	577777777420771	6	2.22	3.08	H4 (Ural-2)	Clustered
141	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▪▪▪▪▪▪▪▪	703767740003771	9	3.33	28.13	CAS1-Delhi	Clustered
186	▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▫▫▫▫▪▫▫▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777767404760771	1	0.37	8.33	LAM7-TUR	Unique
247	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▪▪▫▫▪▪▪▪▪▪	703777740003471	3	1.11	8.11	CAS1-Delhi	Clustered
284	▫▫▫▫▪▪▪▪▪▪▪▫▫▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	037637777760771	9	3.33	5.11	T1	Clustered
309	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▪▪▫▫▪▪▪▪▪▪	703767740003171	17	6.3	80.95	CAS1-Delhi	Clustered
485	▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▫▪▪▪▪▪▪▪▪▪▪	703777400003771	1	0.37	4	CAS1-Delhi	Unique
516	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777777360771	1	0.37	5.56	T1	Unique
568	▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▫▫▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777767477760771	1	0.37	20	T-H37Rv	Unique
610	▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▪▫▫▫▫▪▪▪▫▪▪▪	777767774020731	1	0.37	14.29	H1	Unique
831	▪▪▪▪▪▪▪▪▫▫▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	776367777760771	1	0.37	7.69	S	Unique
878	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▫▪▪▪▪▪	777776777760571	1	0.37	4	X1	Unique
954	▪▪▪▫▫▫▫▪▪▪▪▫▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▪▪▪▪▪▪▪▪▪	703677740003771	1	0.37	6.67	CAS1-Delhi	Unique
1088	▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▪▪▪▪▪▪▪▪▪▪▪▪	777767777763771	4	1.48	33.33	Manu2	Clustered
1144	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▫▫▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777600760771	17	6.3	41.46	T1	Clustered
1198	▪▪▪▫▫▫▫▪▪▪▪▪▫▪▪▪▪▪▪▪▫▫▫▫▫▫▫▫▫▫▫▫▪▪▫▫▪▪▪▪▪▪	703737740003171	4	1.48	10.53	CAS1-Delhi	Clustered
1318	▪▫▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	577767777760771	2	0.74	25	T1	Clustered
1470	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▪▪▫▫▫▫▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777776607760771	2	0.74	10	LAM9	Clustered
1547	▪▪▪▪▪▪▪▪▪▪▪▫▪▫▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777727777760771	3	1.11	27.27	T3	Clustered
1630	▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▫▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777777601760771	1	0.37	16.67	T1	Unique
1638	▪▫▪▪▪▪▪▪▪▪▪▪▫▪▪▪▪▪▪▪▪▪▪▪▪▪▫▫▪▪▪▪▪▪▪▪▪▪▪	577767777763771	1	0.37	33.33	Manu2	Unique
1655	▪▪▪▪▪▪▪▪▪▪▪▪▫▪▫▫▪▪▪▪▪▪▪▪▪▪▪▪▫▫▫▫▪▪▪▪▪▪▪▪	777723777760771	1	0.37	2.7	T3	Unique

Table 3: Continued.

SIT*	Spoligotype description	Octal number	Nb in study	% in study	% in study versus database	Lineage**	Clustered versus unique patterns***
1679		777767737760771	2	0.74	20	T1	Clustered
1913		777767757760771	1	0.37	25	T5	Unique
2032		037627777760771	1	0.37	20	T1	Unique
2230		777700077760771	1	0.37	16.67	Unknown	Unique
2359		703677740003171	1	0.37	4.55	CAS1-Delhi	Unique
2686		703767740003471	4	1.48	66.67	CAS1-Delhi	Clustered
2691		037737777760771	1	0.37	20	T3	Unique
2707		701767740003671	1	0.37	33.33	CAS	Unique
2728		577767777420771	3	1.11	50	H4 (Ural-2)	Clustered
3275*		700037777760731	1	0.37	50	T2	Unique
3346*		777767600760771	11	4.07	100	T1	Clustered
3497*		703767400001771	8	2.96	100	CAS	Clustered
3708*		703767740003761	3	1.11	100	CAS1-Delhi	Clustered
3789*		777777740763771	2	0.74	66.67	Manu2	Clustered
3790*		777767600360771	4	1.48	100	T1	Clustered
3791*		703727740003171	4	1.48	100	CAS1-Delhi	Clustered
3792*		703767640003771	2	0.74	66.67	CAS1-Delhi	Clustered
3793*		557767777763771	1	0.37	50	Manu2	Unique
3794*		703767740000771	1	0.37	50	CAS	Unique
3795*		277767607760771	1	0.37	50	LAM1	Unique
3796*		703767740000371	1	0.37	50	CAS	Unique
3797*		777777774120771	1	0.37	50	H3	Unique
3798*		777717777760371	1	0.37	50	T1	Unique
3799*		777767614760771	1	0.37	50	T1	Unique
3800*		703767400003771	2	0.74	100	CAS	Clustered

A total of 48/64 SITs containing 173 isolates matched a preexisting shared type in the database, whereas 16/64 SITs ($n = 44$ isolates) were newly created. A total of 31/64 SITs containing 184 isolates were clustered within this study (2 to 17 isolates per cluster) while 33/64 SITs containing 33 strains were unique (for total unique strains, one should add to this number the 53 orphan strains, which brings the number of unclustered isolates in this study to 86/270 or 31.85% and clustered isolates to 184/270 or 68.15%). Note that SITs followed by an asterisk indicates "newly created" SITs due to 2 or more strains belonging to an identical new pattern within this study or after a match with an orphan in the database; SIT designations followed by number of strains: 3275 this study $n = 1$, FXX $n = 1$; 3346* this study $n = 11$; 3497* this study $n = 8$; 3708* this study $n = 3$; 3789* this study $n = 2$, IRQ $n = 1$; 3790* this study $n = 4$; 3791* this study $n = 4$; 3792* this study $n = 2$, IND $n = 1$; 3793* this study $n = 1$, IND $n = 1$; 3794* this study $n = 1$, IND $n = 1$; 3795* this study $n = 1$, BRA $n = 1$; 3796* this study $n = 1$, PAK $n = 1$; 3797* this study $n = 1$, MEX $n = 1$; 3798* this study $n = 1$, IRQ $n = 1$; 3799* this study $n = 1$, BRA $n = 1$; 3800* this study $n = 2$.

**Lineage designations according to SITVIT2 using revised SpolDB4 rules; "unknown" designates patterns with signatures that do not belong to any of the major lineages 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. Unique strains matching a preexisting pattern in the SITVIT2 database are classified as SITs, whereas in case of no match, they are designated as "orphan" (see Table 2).

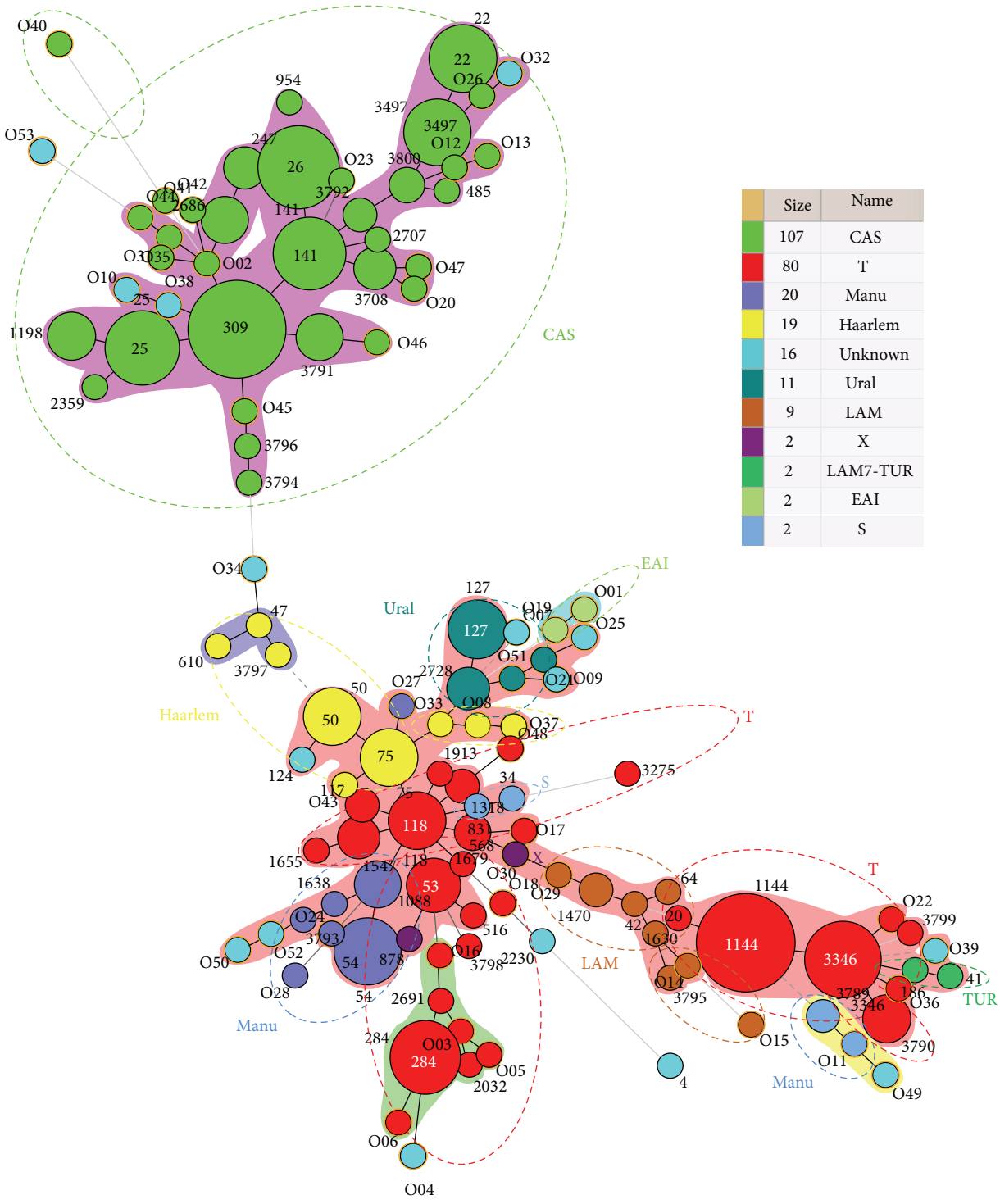

FIGURE 1: A minimum spanning tree (MST) illustrating evolutionary relationships between the *Mycobacterium tuberculosis* spoligotypes. This kind of tree connects each genotype based on degree of changes required to go from one allele to another. The structure of the tree is represented by branches (continuous versus dashed and 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, gray dashed, and gray dotted) denotes the number of allele/spacer changes between two patterns: solid lines represent 1, 2, or 3 changes (thicker ones indicate a single change, while the thinner ones indicate 2 or 3 changes); gray dashed lines represent 4 changes; and gray dotted lines represent 5 or more changes. The size of the circle is proportional to the total number of isolates in our study, illustrating unique isolates (smaller nodes) versus clustered isolates (bigger nodes). The color of the circles indicates the phylogenetic lineage to which the specific pattern belongs. Note that orphan patterns are circled in orange. Patterns colored in cyan-blue indicate a strain with an unknown signature (unclassified).

and X. These strains accounted for 125/270 (46.3%) of the studied strains.

A total of 31/64 SITs containing 184 strains were clustered within this study (2 to 17 strains per cluster) while 33/64 SITs containing 33 strains were unique. For total unique strains, one should add to this number the 53 orphan strains, which brings the number of unclustered strains in this study to 86/270 or 31.85% and clustered strains to 184/270 or 68.15%. 23 clusters are shown in Table 3, along with 8 additional clusters (2 to 11 strains per cluster) represented by the newly created shared types.

The two largest clusters of 17 strains were composed of SIT309 (CAS1-Delhi lineage) and SIT1144 (T1 lineage).

Description of clusters containing 5 or more isolates in this study, and their worldwide distribution in the SITVIT2 database, is detailed in Table 4 and will be commented further under the discussion section.

The MIRU-VNTR analysis detected a total of 73 MIRU patterns from 110 strains using the full 24 MIRU-VNTR locus set, including 17 clusters and 56 unique (Figure 2). Allelic diversity for each locus was calculated in order to determine the discriminatory power of these loci in a combined group for the *Mycobacterium tuberculosis* population studied. Based on their discriminatory index (HGDI), 7 loci (MIRU02, MIRU04, MIRU20, Mtub29, ETR-B, MIRU24, and MIRU27) showed poor discriminatory power (HGDI < 0.3). Seven loci (MIRU 23, Qub11b, Mtub30, Mtub34, Mtub39, MIRU 39, and QUB 4156) discriminated the isolated moderately ($0.3 \leq$ HGDI ≤ 0.6). Lastly, 10 loci (Mtub 04, ETR-C, MIRU 40, MIRU 10, MIRU 16, Mtub21, ETR-A, MIRU 26, MIRU 31, and QUB-26) were highly discriminative (HGDI > 0.6). In this study the locus QUB-26 was found to be the most discriminatory allele in order to distinguish between strains (HGDI of 0.83). Conversely, locus MIRU-24 was found to be the least discriminatory with an HGDI of 0.03. The Recent Transmission Index (RTI_{n-1}) was found at 33.6% showing evidence of ongoing transmission.

3.3. Drug Resistance Patterns. Both the drug resistance patterns and the treatment status of the patients (new versus retreated cases) were studied in detail on all the 270 strains included in this study in function of their spoligotyping-based genotypic lineages, and the results were concomitantly exploited to draw a minimum spanning tree (MST) shown in Figure 3. Of the 270 strains studied, 171 (63.3%) were sensitive to all five of the first-line drugs tested, 64 (23.70%) were MDR, while the drug-susceptibility information was not available for 35 (12.96%) of strains. Regarding the treatment status of the patients, 187 (69.26%) were new cases while 83 (30.74%) were previously treated. It is noteworthy that all the 64 MDR cases were exclusively found among the retreated patients, bringing the proportion of MDR isolates in this group to 77.1% the difference of drug resistance between new versus retreated cases being highly significant (P value < 0.0001). As shown in Figure 3(a), a high frequency of MDR-TB was associated with SIT53/T1 which contained 60% of MDR strains. Interestingly, this same shared type was also exclusively associated with retreated cases (Figure 3(b)). Although the clustering rate between drug susceptible and drug resistant isolates did not vary significantly (P value > 0.4), differences were noted when comparing predominant SITs. Indeed, one may notice a relatively high frequency of MDR-TB among patterns related to SIT53/T1 (60% of MDR) in Figure 3(a), which was significantly higher than other SITs such as SIT26/CAS1-Delhi (16.67% of MDR), P value $= 0.0372$, and SIT141/CAS1-Delhi (11.11% of MDR), P value $= 0.0030$. Lastly, although the rate of MDR-TB was slightly higher among patients from cities other than Baghdad (Table 1), the difference was not statistically significant (P value $= 0.1224$).

4. Discussion

In this study we characterized, by spoligotyping, 270 *M. tuberculosis* isolates collected from patients diagnosed in Baghdad/Reference TB laboratory in a 13-month period. One hundred and ten samples were also characterized by 24-loci MIRU-VNTR.

All TB cases reported in this study were caused by *M. tuberculosis*. The molecular investigation of strain by spoligotyping did not show the specific indicator for other members of *M. tuberculosis* complex such as *M. bovis*. This situation has been described in other settings and countries such as by Godreuil et al. in west African countries [20], Nakajima et al. in Bangladesh [21], and Viegas in Mozambique [22]. Spoligotyping of the 270 *M. tuberculosis* strains revealed 31 clusters consisting of 184 strains, with clustering rate 57%, whereas 86 strains were unique.

One hundred and seven (39.6%) of the 270 studied strains were CAS belonging to different SITs. CAS has also been identified as predominant family in Saudi Arabia (22.5%) and also the predominance of Delhi genogroup in Iran [23, 24] was reported. In Pakistan this lineage represents 61% of the total [25]. And CAS has been also identified by recent study as a predominant strain in north India [26]. The data from Turkey suggested that there is no dominant *M. tuberculosis* clade such as what has been observed in Asia and former USSR republics [17]. CAS clade is followed, in decreasing order, by ill-defined T clade 29.6%, Manu 7.4%, Haarlem 7%, Ural 4.1%, LAM 3.3%, X 0.7%, LAM7-TUR 0.7%, EAI 0.7%, S 0.7%, and unknown 5.9% (Table 4). These observations emphasize the complex diversity of circulating *M. tuberculosis* strains in Iraq that could reflect the different transmission pathways occurring within the country. Besides, it has been suggested that particular lineages of *M. tuberculosis* might be adapted to specific human population [27, 28]. Indeed, many strains of this study belonged to Principal Genetic Group (PGG1), according to the *KatG-grA* polymorphism-based classification of Sreevatsan et al. [19], and these strains represent all of the known lineages associated with PGG1 with the exception of the Beijing clade. However, evolutionary modern PGG 2/3 strains were also found with different distribution.

In this study 110 strains were classified into 73 MIRU patterns of which 17 were clusters and 56 unique strains. The high level of Recent Transmission Index (RTI), at 33.6%,

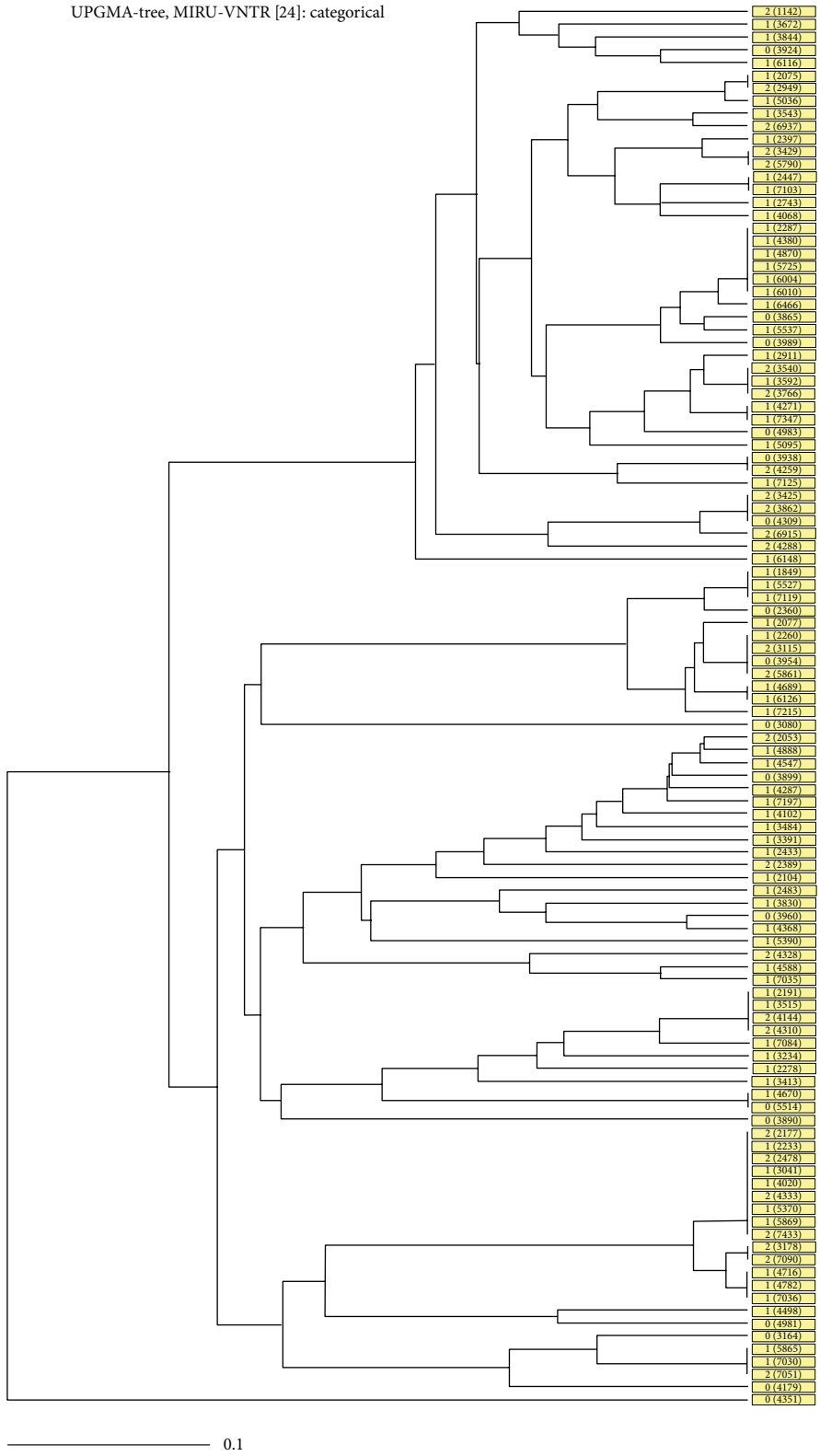

Figure 2: UPGMA tree based on MIRU 24 pattern of the subset of 110 samples. Drug resistance (Drug R) information is shown as 0, unknown; 1, non-MDR; 2, MDR-TB, that is, combined resistance to INH-RIF (with or without resistance to other drugs).

Table 4: Description of clusters containing 5 or more isolates in this study and their worldwide distribution in the SITVIT2 database.

SIT (lineage) octal number Spoligotype description	Number (%) in study	% in study versus database	Distribution in regions with ≥3% of a given SIT*	Distribution in countries with ≥3% of a given SIT**
22 (CAS1-Delhi) 703777400001771	8 (2.96)	10.26	ASIA-W 65.39, ASIA-S 12.82, AFRI-E 7.69, EURO-N 5.13, EURO-W 3.85	SAU 38.46, IRQ 25.64, IRN 7.69, TZA 3.85
25 (CAS1-Delhi) 703777740003171	10 (3.7)	1.77	AFRI-E 26.68, ASIA-W 22.26, AFRI-N 14.84, ASIA-S 13.6, AMER-N 9.19, EURO-N 4.95, EURO-W 4.77	ETH 25.44, SAU 16.61, SDN 12.19, USA 9.19, IND 8.66, IRQ 5.12, IRN 3.0
26 (CAS1-Delhi) 703777740003771	12 (4.44)	0.87	ASIA-S 54.08, AMER-N 15.81, ASIA-W 7.51, EURO-W 6.14, AFRI-E 5.13, EURO-N 4.19, EURO-S 3.61	IND 33.0, USA 15.81, PAK 10.4, SAU 5.63, BGD 5.42, IRN 4.48, ITA 3.47, NLD 3.11, ETH 3.11
50 (H3) 777777777720771	6 (2.22)	0.18	AMER-S 18.03, EURO-W 17.88, AMER-N 17.88, EURO-S 11.75, EURO-E 5.62, EURO-N 4.36, AFRI-N 4.33, AFRI-S 4.12, CARI 3.55, ASIA-W 3.19	USA 17.85, BRA 7.18, FXX 7.0, AUT 6.19, ITA 5.53, ESP 5.53, PER 4.81, ZAF 4.12, CZE 3.73
53 (T1) 777777777760771	5 (1.85)	0.08	EURO-W 16.17, AMER-N 13.93, AMER-S 12.43, EURO-S 9.72, ASIA-W 7.19, EURO-N 5.51, AFRI-S 5.13, AFRI-E 4.66, ASIA-E 4.38, AFRI-N 3.64, EURO-E 3.37, AMER-C 3.34	USA 13.64, FXX 8.14, ITA 5.51, BRA 5.3, ZAF 5.02, TUR 3.59, AUT 3.54, CHN 3.19
54 (Manu2) 777777777763771	8 (2.96)	3.31	ASIA-W 16.53, ASIA-S 15.29, ASIA-E 15.29, AFRI-N 14.88, AMER-N 8.26, AFRI-S 6.61, ASIA-N 4.13, AMER-S 3.72	CHN 15.29, IND 14.46, EGY 14.05, USA 8.26, SAU 8.26, ZAF 6.61, IRQ 5.79, RUS 4.13
75 (H3) 777767777720771	6 (2.22)	8.7	AMER-S 18.84, ASIA-W 15.94, EURO-W 11.59, ASIA-S 10.15, EURO-S 7.25, AMER-N 7.25, AFRI-M 7.25, AFRI-E 5.8, EURO-E 4.35, CARI 4.35, AFRI-W 4.35	IRN 10.15, FXX 10.15, BRA 10.15, IRQ 8.7, USA 7.25, CMR 7.25, TUR 4.35, MDG 4.35, GUF 4.35, ARG 4.35
118 (T1) 777767777760771	6 (2.22)	3.9	ASIA-W 18.83, EURO-W 17.53, AMER-S 13.64, EURO-S 9.74, ASIA-E 9.74, AMER-N 5.2, AFRI-E 5.2, EURO-E 4.55, AMER-C 4.55	FXX 15.58, TUR 12.34, ITA 6.49, VEN 5.84, CHN 5.84, USA 5.2, MOZ 4.55, JPN 3.9, IRQ 3.9, CZE 3.9, BRA 3.9, MEX 3.25, ARG 3.25
127 (Ural-2) 577777777420771	6 (2.22)	3.08	ASIA-S 54.36, EURO-W 16.41, ASIA-W 11.8, EURO-N 5.64, ASIA-E 5.13	IRN 49.23, NLD 7.69, SAU 6.15, IRQ 5.13, CHN 5.13, SWE 4.62, AUT 4.62, DEU 3.59
141 (CAS1-Delhi) 703767740003771	9 (3.33)	28.13	ASIA-S 37.5, ASIA-W 31.25, EURO-W 12.5, AMER-N 12.5, ASIA-SE 3.13, AFRI-E 3.13	IND 34.38, IRQ 28.13, USA 12.5, FXX 9.38, SAU 3.13, PAK 3.13, NLD 3.13, MOZ 3.13, MMR 3.13
284 (T1) 037637777760771	9 (3.33)	5.11	ASIA-W 42.61, EURO-E 22.73, EURO-W 15.91, EURO-N 7.39, EURO-S 3.41	BGR 22.73, TUR 19.89, SAU 13.07, IRQ 9.09, SWE 6.25, AUT 5.68, FXX 3.41
309 (CAS1-Delhi) 703767740003171	17 (6.3)	80.95	ASIA-W 85.71, AFRI-M 9.52, AFRI-E 4.76	IRQ 80.95, CAF 9.52, SAU 4.76, ETH 4.76
1144 (T1) 777777600760771	17 (6.3)	41.46	ASIA-W 90.24, AMER-N 4.88	IRQ 90.24, USA 4.88
3346 (T1) 777767600760771	11 (4.07)	100	ASIA-W 100.0	IRQ 100.0

Table 4: Continued.

SIT (lineage) octal number Spoligotype description	Number (%) in study	% in study versus database	Distribution in regions with ≥3% of a given SIT*	Distribution in countries with ≥3% of a given SIT**
3497 (CAS) 703767400001771 ■■■□□□■■■■■■■□■■■■□□□□□□□□□□□□□□□■■■■■■■■	8 (2.96)	100	ASIA-W 100.0	IRQ 100.0

*Worldwide distribution is reported for regions with more than 3% of a given SITs as compared to their total number in the SITVIT2 database. The definition of macrogeographical regions and subregions (http://unstats.un.org/unsd/methods/m49/m49regin.htm) is according to the United Nations; regions: AFRI (Africa), AMER (Americas), ASIA (Asia), EURO (Europe), and OCE (Oceania) are 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 subregions, AUST (Australasia), MEL (Melanesia), MIC (Micronesia), and POLY (Polynesia). Note that in our classification scheme, Russia has been attributed a new subregion by itself (northern Asia) instead of including it among the rest of 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 ≥3% of a given SITs as compared to their total number in the SITVIT2 database.

FIGURE 3: A minimum spanning tree (MST) illustrating evolutionary relationships between the *M. tuberculosis* spoligotypes in our study in function to studied parameters. (a) Drug resistance and (b) treatment status, that is, new versus retreated cases.

indicates that the most cases are due to recent transmissions rather than reactivation of *M. tuberculosis* infections. In our study, ten loci (Mtub 04, ETR-C, MURU 40, MIRU 10, MIRU 16, Mtub21, ETR-A, MIRU 26, MIRU 31, and QUB-26) were highly discriminatory (HGDI > 0.6), seven loci (MIRU23, Qub11b, Mtub30, Mtub34, Mtub39, MIRU 39, and QUB 4156) moderately discriminate ($0.3 \leq \text{HGDI} \leq 0.6$), and seven loci (MIRU02, MIRU04, MIRU20, Mtub 29, ETR-B, MIRU 24, and MIRU27) showed poor discriminatory power (HGDI < 0.3).

MIRU-VNTR allelic results have been correlated with definition of ancestral and modern MTB lineages, with the presence of single allele in locus 24 being related to a modern strain type. We found that 98% of our strains contained only single repeat at locus 24, further confirming their modern lineage. This is comparable with previous reports for CAS strain from Bangladesh and Singapore [29, 30] and also with the studies from Pakistan and Bulgaria as supported by the finding that 62% of their CAS family strains contained only one allele at the locus 24 [31, 32]. Moreover, the relative discriminatory powers of particular VNTR loci vary depending on the strain in question [33–37].

This study found that in Iraqi population, the characteristic of MDR in *M. tuberculosis* is mostly acquired as a result of treatment failure, due to irregularity in taking of drug (anti-TB), neglect, and incorrect prescriptions. Although the extremely high level of MDR among previously treated patients might indicate transmission chains within the population, molecular epidemiology revealed that except for SIT53/T1 genotype, no significant differences were found and the MIRU analysis in the subset of strains did not show clusters of exclusively MDR strains. The fact that identical MIRU patterns were shared both by MDR and non-MDR strains, and that the isoniazid and rifampicin resistance patterns were independent of their genotypes, suggests that MDR strains most probably emerged due to the selective pressure because of problems in adherence to treatment (in addition to other environmental factors), (high population, poor housing, overcrowding, and malnutrition).

The data of this study provide important baseline information on the genetic diversity of *M. tuberculosis* in Iraq. Therefore, it could be used to monitor change in the transmission pattern of tuberculosis. Spoligotyping has been proved useful for categorizing strains into different families and can be used as an initial technique to be subsequently followed by MIRU-VNTR. This study showed that 24-loci MIRU-VNTR typing offers a higher discriminatory power. Iraq needs to conduct epidemiological survey by using conventional and genotyping methods in order to provide adequate data that can be used for the formulation of control strategies of tuberculosis transmission.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

Nalin Rastogi and Daniela M. Cirillo contributed equally to this paper.

Acknowledgment

Ruqaya Mustafa Ali would like to thank Baghdad University, Iraq, for funding and support.

References

[1] WHO, "Global tuberculosis control," WHO Report, WHO, Geneva, Switzerland, 2012.

[2] Ministry of Health of Iraq, "New diagnosis of tuberculosis in Iraq," Tech. Rep., Ministry of Health of Iraq, Baghdad, Iraq, 2012.

[3] B. Mathema, N. E. Kurepina, P. J. Bifani, and B. N. Kreiswirth, "Molecular epidemiology of tuberculosis: current insights," *Clinical Microbiology Reviews*, vol. 19, no. 4, pp. 658–685, 2006.

[4] I. S. Kontsevaya, V. V. Nikolayevsky, and Y. M. Balabanova, "Molecular epidemiology of tuberculosis: objectives, methods, and prospects," *Molecular Genetics, Microbiology and Virology*, vol. 26, no. 1, pp. 1–9, 2011.

[5] P. Supply, S. Lesjean, E. Savine, K. Kremer, D. van Soolingen, and C. Locht, "Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units," *Journal of Clinical Microbiology*, vol. 39, no. 10, pp. 3563–3571, 2001.

[6] L. S. Cowan, L. Diem, T. Monson et al., "Evaluation of a two-step approach for large-scale, prospective genotyping of *Mycobacterium tuberculosis* isolates in the United States," *Journal of Clinical Microbiology*, vol. 43, no. 2, pp. 688–695, 2005.

[7] P. Supply, C. Allix, S. Lesjean et al., "Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*," *Journal of Clinical Microbiology*, vol. 44, no. 12, pp. 4498–4510, 2006.

[8] A. Thong-On, N. Smittipat, T. Juthayothin et al., "Variable-number tandem repeats typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110 in Thailand," *Tuberculosis*, vol. 90, no. 1, pp. 9–15, 2010.

[9] T. Wirth, F. Hildebrand, C. Allix-Béguec et al., "Origin, spread and demography of the *Mycobacterium tuberculosis* complex," *PLoS Pathogens*, vol. 4, no. 9, Article ID e1000160, 2008.

[10] M. Cardoso Oelemann, H. M. Gomes, E. Willery et al., "The forest behind the tree: phylogenetic exploration of a dominant *Mycobacterium tuberculosis* strain lineage from a high tuberculosis burden country," *PloS ONE*, vol. 6, no. 3, Article ID e18256, 2011.

[11] H. David, V. Levy-Frébault, and M. F. Thorel, *Méthodes de laboratoire de mycobactériologie clinique*, Commission des Laboratoires de Référence et d'Expertise de l'Institut Pasteur, Paris, France, 1989.

[12] D. van Soolingen, P. W. M. Hermans, P. E. W. de Haas, D. R. Soll, and J. D. A. van Embden, "Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis," *Journal of Clinical Microbiology*, vol. 29, no. 11, pp. 2578–2586, 1991.

[13] J. Kamerbeek, L. Schouls, A. Kolk et al., "Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology," *Journal of Clinical Microbiology*, vol. 35, no. 4, pp. 907–914, 1997.

[14] K. Brudey, J. R. Driscoll, L. Rigouts et al., "*Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology," *BMC Microbiology*, vol. 6, article 23, 2006.

[15] C. Demay, B. Liens, T. Burguière et al., "SITVITWEB—a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology," *Infection, Genetics and Evolution*, vol. 12, no. 4, pp. 755–766, 2012.

[16] F. Koro Koro, Y. K. Simo, F. F. Piam et al., "Population dynamics of tuberculous Bacilli in Cameroon as assessed by spoligotyping," *Journal of Clinical Microbiology*, vol. 51, no. 1, pp. 299–302, 2013.

[17] O. Kisa, G. Tarhan, S. Gunal et al., "Distribution of spoligotyping defined genotypic lineages among drug-resistant *Mycobacterium tuberculosis* complex clinical isolates in Ankara, Turkey," *PLoS ONE*, vol. 7, no. 1, Article ID e30331, 2012.

[18] I. Mokrousov, "The quiet and controversial: ural family of *Mycobacterium tuberculosis*," *Infection, Genetics and Evolution*, vol. 12, no. 4, pp. 619–629, 2012.

[19] S. Sreevatsan, X. Pan, K. E. Stockbauer et al., "Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 18, pp. 9869–9874, 1997.

[20] S. Godreuil, G. Torrea, D. Terru et al., "First molecular epidemiology study of *Mycobacterium tuberculosis* in Burkina Faso," *Journal of Clinical Microbiology*, vol. 45, no. 3, pp. 921–927, 2007.

[21] C. Nakajima, Z. Rahim, Y. Fukushima et al., "Identification of *Mycobacterium tuberculosis* clinical isolates in Bangladesh by a species distinguishable multiplex PCR," *BMC Infectious Diseases*, vol. 10, article 118, 2010.

[22] S. O. Viegas, A. Machado, R. Groenheit et al., "Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mozambique," *BMC Microbiology*, vol. 10, article 195, 2010.

[23] S. A. Al-Hajoj, "Tuberculosis in Saudi Arabia: can we change the way we deal with the disease?" *Journal of Infection and Public Health*, vol. 3, no. 1, pp. 17–24, 2010.

[24] M. Jafarian, M. Aghali-Merza, P. Farnia, M. Ahmadi, M. R. Masjedi, and A. A. Velayati, "Synchronous comparison of *Mycobacterium tuberculosis* epidemiology strains by MIRU-VNTR and MIRU-VNTR and spoligotyping technique," *Avicenna Journal of Medical Biotechnology*, vol. 2, no. 3, pp. 145–152, 2010.

[25] M. Tanveer, Z. Hasan, A. R. Siddiqui et al., "Genotyping and drug resistance patterns of *M. tuberculosis* strains in Pakistan," *BMC Infectious Diseases*, vol. 8, article 171, pp. 2–10, 2008.

[26] M. Varma-Basil, S. Kumar, J. Arora et al., "Comparison of spoligotyping, mycobacterial interspersed repetitive units typing and IS6110-RFLP in a study of genotypic diversity of *Mycobacterium tuberculosis* in Delhi, North India," *Memórias do Instituto Oswaldo Cruz Rio de Janeiro*, vol. 106, no. 5, pp. 524–535, 2011.

[27] Z. H. Helal, M. S. E.-D. Ashour, S. A. Eissa et al., "Unexpectedly high proportion of ancestral manu genotype *Mycobacterium*

tuberculosis strains cultured from tuberculosis patients in Egypt," *Journal of Clinical Microbiology*, vol. 47, no. 9, pp. 2794–2801, 2009.

[28] S. Gagneux, K. DeRiemer, T. Van et al., "Variable host-pathogen compatibility in *Mycobacterium tuberculosis*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 8, pp. 2869–2873, 2006.

[29] S. Banu, S. V. Gordon, S. Palmer et al., "Genotypic analysis of *Mycobacterium tuberculosis* in bangladesh and prevalence of the Beijing strain," *Journal of Clinical Microbiology*, vol. 42, no. 2, pp. 674–682, 2004.

[30] Y.-J. Sun, R. Bellamy, A. S. G. Lee et al., "Use of mycobacterial interspersed repetitive unit-variable-number tandem repeat typing to examine genetic diversity of *Mycobacterium tuberculosis* in Singapore," *Journal of Clinical Microbiology*, vol. 42, no. 5, pp. 1986–1993, 2004.

[31] A. Ali, Z. Hasan, M. Tanveer et al., "Characterization of *Mycobacterium tuberculosis* central Asian Strain 1 using mycobacterial interspersed repetitive unit genotyping," *BMC Microbiology*, vol. 7, article 76, 2007.

[32] V. Valcheva, I. Mokrousov, O. Narvskaya, N. Rastogi, and N. Markova, "Utility of new 24-locus variable-number tandem-repeat typing for discriminating *Mycobacterium tuberculosis* clinical isolates collected in Bulgaria," *Journal of Clinical Microbiology*, vol. 46, no. 9, pp. 3005–3011, 2008.

[33] Y. Murase, S. Mitarai, I. Sugawara, S. Kato, and S. Maeda, "Promising loci of variable numbers of tandem repeats for typing Beijing family *Mycobacterium tuberculosis*," *Journal of Medical Microbiology*, vol. 57, no. 7, pp. 873–880, 2008.

[34] I. Comas, S. Homolka, S. Niemann, and S. Gagneux, "Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies," *PLoS ONE*, vol. 4, no. 11, Article ID e7815, 2009.

[35] S. Ferdinand, G. Valétudie, C. Sola, and N. Rastogi, "Data mining of *Mycobacterium tuberculosis* complex genotyping results using mycobacterial interspersed repetitive units validates the clonal structure of spoligotyping-defined families," *Research in Microbiology*, vol. 155, no. 8, pp. 647–654, 2004.

[36] M. Aminian, A. Shabbeer, and K. P. Bennett, "A conformal Bayesian network for classification of *Mycobacterium tuberculosis* complex lineages," *BMC Bioinformatics*, vol. 11, supplement 3, article S4, 2010.

[37] N. H. Mendes, F. A. F. Melo, A. C. B. Santos et al., "Characterization of the genetic diversity of *Mycobacterium tuberculosis* in São Paulo city, Brazil," *BMC Research Notes*, vol. 4, article 269, 2011.