

First Insight into the Population Structure of *Mycobacterium tuberculosis* in Saudi Arabia^{∇†}

Sahal A. M. Al-Hajoj,^{1*} Thierry Zozio,² Fahad Al-Rabiah,³ Viqar Mohammad,¹ Maryam Al-Nasser,¹ Christophe Sola,^{2‡} and Nalin Rastogi^{2*}

Departments of Comparative Medicine¹ and Medicine,³ King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia, and Unité de la Tuberculose et des Mycobactéries, Institut Pasteur de Guadeloupe, Abymes, Guadeloupe²

Received 12 November 2006/Returned for modification 1 January 2007/Accepted 7 May 2007

This study constitutes a first attempt to describe the genetic population structure and drug resistance of the tubercle bacilli circulating in Saudi Arabia. A total of 1,505 clinical isolates of *M. tuberculosis*, isolated between 2002 and 2005 from seven regions of Saudi Arabia, were studied. The sample studied showed a male-to-female sex ratio of 1.27, with half of the cases among foreign-born individuals and 47% within the 21- to 40-year-old age group; a total resistance rate of 19.7%; and multiple drug resistance of 4.5%. Upon spoligotyping, a total of 387 individual patterns were obtained (clustering rate, 86.4%; 182 clusters containing between 2 and 130 isolates per cluster). A total of 94% of the strains matched the spoligotype patterns in an international database. Nearly 81% of the isolates in this study belonged to established phylogeographic clades: Central Asian (CAS), 22.5%; ill-defined T clade, 19.5%; East African-Indian (EAI), 13.5%; Haarlem, 7.5%; Latin American-Mediterranean, 7.2%; Beijing, 4.4%; Manu, 2.7%; X, 0.9%; and Bovis, 0.9%. Two clonal complexes with unique spoligotyping signatures (octal codes 703777707770371 and 467777377413771) specific to Saudi Arabia were identified. These belonged to the CAS and EAI clades, respectively, as confirmed upon secondary typing using mycobacterial interspersed repetitive units (MIRUs). The results obtained underline the predominance of historic clones of principal genetic group 1, which are responsible for roughly 45% of all tuberculosis cases in Saudi Arabia. The high rate of clustering observed might be an indication of rapid ongoing transmission within certain communities and/or subpopulations in Saudi Arabia; nonetheless, spoligotyping is known to overestimate clustering, and only a systematic second-line typing, such as MIRUs, coupled with a better tuberculosis registry and epidemiological investigations would allow us to know the exact rate of ongoing transmission and associated risk factors in Saudi Arabia.

Molecular typing of *Mycobacterium tuberculosis* greatly assists our understanding of the epidemiology of this pathogen and may lead to improved control of the disease (23). Restriction fragment length polymorphism analysis using IS6110 and PCR-based methods, such as spoligotyping and mycobacterial interspersed repetitive units–variable-number tandem repeats (MIRU-VNTRs), are considered standard methodologies for elucidating the worldwide dissemination of *M. tuberculosis* (13, 14, 22, 24). Spoligotyping is particularly useful for population-based studies to determine the phylogeographic specificity of circulating clades of tubercle bacilli (7, 10). It permits assessment of the complexity of global tuberculosis (TB) transmission and the temporal evolution of the TB genetic landscape (7). MIRU-VNTR is a useful second-line method for high-throughput analysis to detect epilinked strains in a given pop-

ulation (12, 22). Considered a gold standard for molecular epidemiology of TB in the 1990s (24), IS6110-restriction fragment length polymorphism is a cumbersome method requiring large quantities of bacterial DNA and is not optimal for typing *M. tuberculosis* strains either devoid of or possessing less than five copies of the IS6110 insertion element (16). Lastly, phylogenetically relevant long-sequence and single-nucleotide polymorphisms have recently been used to study the phylogeny of the *M. tuberculosis* complex (10, 11).

The present study attempts to describe the genetic population structure and drug resistance of the tubercle bacilli circulating in Saudi Arabia. The country has a moderate incidence of active TB (30 cases per 100,000 inhabitants); however, the infection rate varies between cities. As an example, in 1990, the incidence in Jeddah (Western Province) reached 64 cases per 100,000 compared with 32 per 100,000 in Riyadh. The annual influx of pilgrims, as well as a high proportion of migrant workers in the country, the majority from countries where TB is endemic, provides an opportunity for the transmission of TB (3, 6, 18). The skin test conversion rate in unvaccinated Saudi Arabian children is about 0.5% per year, lower than in sub-Saharan countries (2%) but higher than in Europe (estimated at 0.1%). A high proportion of multidrug-resistant (MDR) TB was reported, with an estimated cure rate of 65% (1, 4, 19). In order to gain insight into the dynamics of the disease and to identify the prevalent strains and the possible routes of TB transmission, this investigation reports on the first nationwide

* Corresponding author. Mailing address for Sahal Al-Hajoj: Department of Comparative Medicine, King Faisal Specialist Hospital and Research Center (MBC 03), PO Box 3354, Riyadh 11211, Saudi Arabia. Phone: 966 (1) 4424992. Fax: 966 (1) 4427872. E-mail: hajoj@kfshrc.edu.sa. Mailing address for Nalin Rastogi: Institut Pasteur de Guadeloupe, BP484, 97183 Abymes, Guadeloupe. Phone: 590-590-897661. Fax: 590-590-893880. E-mail: nrastogi@pasteur-guadeloupe.fr.

† Supplemental material for this article may be found at <http://jcm.asm.org>.

‡ Present address: Unité de Recherche de Génétique Mycobactérienne, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris, Cedex 15, France.

[∇] Published ahead of print on 16 May 2007.

study to characterize clinical isolates of *M. tuberculosis* from different regions of Saudi Arabia.

MATERIALS AND METHODS

Mycobacterial strains, culture, and drug susceptibility testing. A total of 1,505 *M. tuberculosis* strains from seven different regions of Saudi Arabia were used in this study. The collection constituted all available isolates from the period 2002 to 2005. All isolates and clinical data were collected in accordance with the guidelines of the Ethics Committee of King Faisal Specialist Hospital and Research Center. The strains were isolated and cultured using Löwenstein-Jensen medium, and the drug susceptibility testing was performed using the proportion method. To avoid cross-contamination during isolation and culturing, all the manipulations were done individually in biological safety cabinets equipped with a negative-pressure system.

DNA extraction and spoligotyping. The bacterial DNAs were prepared in a separate negative-pressure cabinet and subjected to spoligotyping at King Faisal Specialist Hospital and Research Center using previously described methodology (13). To avoid cross-contamination, PCR was carried out in an independent UV PCR workstation dedicated to this effort; extra measures taken included UV and manual decontamination after each step. Briefly, the direct-repeat (DR) region was amplified with primers DRa (biotinylated at the 5' end) and DRb, and the amplified DNA was hybridized to inter-DR spacer oligonucleotides covalently bound to a membrane (Isogen Life Science, Lagedijk, The Netherlands). DNA from *Mycobacterium bovis* BCG and *M. tuberculosis* H37Rv were used as positive controls, whereas autoclaved ultrapure water was used as a negative control. The presence of spacers was visualized on film as black squares after incubation with streptavidin-peroxidase and detected with the enhanced chemoluminescence system detection liquid (Amersham, Little Chalfont, United Kingdom). All spoligotypes were converted into octal format within Excel spreadsheets (8) and entered into Bionumerics software (Applied Maths, Sint Maarten Latem, Belgium). Further phylogenetic analysis using parsimony was performed using the PAUP software (version 4.0b10) to draw dendrograms. The pairwise distance between patterns was computed using the unweighted pair group method using arithmetic averages and the Jaccard index. This methodology has proved useful to define major phylogeographical clades within the *M. tuberculosis* complex (7, 9, 10, 20). The individual spoligotyping patterns were compared with an updated SpolDB4 database named SITVIT2 (the initial version is available at <http://www.pasteur-guadeloupe.fr:8081/SITVIT2Demo>). The updated in-house proprietary version contained at the time of this comparison a total of 2,808 spoligotype international types (SITs) corresponding to 63,473 clinical isolates from 122 isolation countries and 160 countries of origin. Major phylogenetic clades were assigned according to signatures provided in SpolDB4 (7), which defines 62 genetic lineages/sublineages (see the original paper for a detailed description). These include specific signatures for various *M. tuberculosis* complex species, such as *M. bovis*, *Mycobacterium microti*, *Mycobacterium caprae*, *Mycobacterium pinipediti*, and *Mycobacterium africanum*, as well as rules defining major lineages/sublineages for *M. tuberculosis* sensu stricto. These include the Central Asian (CAS) clade and 2 sublineages, the East African-Indian (EAI) clade and 9 sublineages, the Haarlem clade and 34 sublineages, the Latin American-Mediterranean (LAM) clade and 12 sublineages, the "Manu" family and 3 sublineages, the S clade, the IS6110-low-binding X clade and 3 sublineages, and an ill-defined T clade with 5 sublineages.

Secondary typing using MIRU-VNTRs. Two specific clones from this study, thought to be specific to Saudi Arabia, were characterized using 12-locus MIRU-VNTRs as described previously (22). Briefly, PCR mixtures were prepared as follows: 20 ng of DNA was added to a final volume of 30 μ l containing 0.6 U of recombinant *Taq* polymerase (Amersham, Uppsala, Sweden); 3 μ l of 10 \times recombinant *Taq* buffer mix (Amersham); and 0.6 μ l of a mixture containing 25 mM of each of the deoxynucleoside triphosphates (Obiogene Inc., Carlsbad, CA), 2.4 μ l of 25 mM MgCl₂, 3 μ l of dimethyl sulfoxide, and 2 μ l of each primer diluted to 4 μ M. PCR was performed in a Perkin-Elmer GeneAmp PCR system 9600 (PE Biosystems, Foster City, CA). An initial denaturation of 10 min at 94°C was followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 1 min, and extension at 72°C for 2 min, with a final extension step at 72°C for 10 min. DNA fragments were separated by electrophoresis on an agarose gel (1.5% [wt/vol]; Invitrogen Corp., Carlsbad, CA) in 1 \times Tris-borate-EDTA containing three rows of 51 slots (slot width, 2 mm) in a subcell model Fisherbrand HU25 (Fisher-Bioblock, Illkirch, France). A 100-bp molecular size marker (Amersham) was loaded every five or six lanes, and the fragments were separated for 3 h at 150 V. The exact MIRU copy number corresponding to the respective band sizes was calculated according to information provided by Philip Supply (Institut Pasteur, Lille, France) and entered into an Excel spreadsheet. The MIRU data obtained

were entered into the SITVIT2 database (see above), which at the time of this comparison contained 12-locus MIRU patterns for a total of 8,573 isolates and 882 shared types, referred as MITs (MIRU international types).

RESULTS AND DISCUSSION

Demographic information and drug resistance. To our knowledge, this is the first description of the *M. tuberculosis* clinical isolates prevailing in Saudi Arabia ($n = 1,505$ isolates from as many patients from seven different regions of Saudi Arabia) (Table 1). Despite being as exhaustive as possible and including all the *M. tuberculosis* cultures available, a lack of proper registry did not allow us to conclude that the recruitment of strains was exhaustive. Hence, our selection of strains may be considered to reflect a "descriptive" rather than a "systematic" study; nonetheless, our sampling remains more representative of the TB situation in Saudi Arabia than what would have been obtained through a purely "convenience" sample. The related demographic information and drug resistance patterns obtained are summarized in Table 1. As can be seen from the table, lack of a proper registry system did not allow us to have homogeneous data for all settings. Nonetheless, the following observations could be made: the male-to-female sex ratio varied from 0.8 to 1.3, with a mean value of 1.27 (data were available for 82.9% of the cases); age was reported for 70.2% of cases, with disease being most frequent among the 21- to 40-year-old age group (47.3% of all cases); the nationality was reported for 58.7% of cases, with half of the cases diagnosed among foreign born individuals; and smear-microscopy data were available for 44.3% of cases, with 75% being smear positive. Regarding drug resistance determination, data were available for 1,124 out of 1,505 (75%) strains as follows: pansusceptible, 80.3%; total resistance, 19.7%; total MDR, 4.5% (i.e., strains resistant to isoniazid [I] and rifampin [R], with or without additional resistance to streptomycin [S] and ethambutol [E]). The distribution of major resistance patterns of the MDR strains ($n = 50$) among the total drug susceptibility test results known ($n = 1,124$) was as follows: I plus R, 10 (0.9% of strains tested and 20% of MDR strains); I plus R plus S, 18 (1.6% and 36%, respectively); I plus R plus S plus E, 21 (1.87% and 42%, respectively). Unfortunately, the TB registry did not allow discrimination of new from recurrent cases or classification by treatment outcomes.

Distribution of major phylogenetic clades of *M. tuberculosis*. A total of 387 individual patterns were obtained upon spoligotyping, and the patterns of 1,415/1,505 (94%) of the strains matched those present in the database. A total of 1,300/1,505 (86.4%) strains were clustered within our study (in 182 clusters containing between 2 and 130 isolates per cluster) (Fig. 1), and 205 (13.6%) were unclustered, i.e., had a genetic profile with a single representative in Saudi Arabia. In a second step, we attempted to see how many of these 205 unclustered isolates were really unique (i.e., not reported elsewhere in the world); 90 were found to be true orphans (6% of the total strains), found preferentially in Taif, Riyadh, and Dammam. In conclusion, a total of 1,415/1,505 (94%) strains from this study had a counterpart in the SITVIT2 database. The distribution of phylogenetic clades in Saudi Arabia is summarized in Table 2 (for further information, see the supplemental material). Nearly 81% of the isolates belonged to known genotype clades or

TABLE 1. Geographic origins of 1,505 *M. tuberculosis* clinical isolates and related demographic information and drug susceptibility patterns

Parameter	Value							Total
	Riyadh	Damman	Taif	Medina	Tabuk	Jizan	Al-Bahah	
Geographic location	Central	East	West-South	West	Northwest	South	South	
Total no. of isolates	428	419	443	117	60	25	13	1,505
Sex								
Male	195	220	252	NA ^a	25	6	NA	698
Female	165	164	191	NA	22	7	NA	549
Unknown	68	35	0	NA	13	12	NA	258
Sex ratio	1.2	1.3	1.3	NA	1.1	0.8	NA	1.27
Age group (yr) (no.)								
0–20	40	33	46	NA	5	1	NA	125
21–40	112	226	138	NA	17	6	NA	499
41–60	106	86	76	NA	9	5	NA	282
>60	83	26	24	NA	16	1	NA	150
Unknown	87	48	159	NA	13	12	NA	449
Nationality (no.)								
Saudi Arabian	NA	160	224	NA	47	9	NA	440
Foreign born	NA	217	211	NA	13	2	NA	443
Unknown	NA	42	8	NA	0	14	NA	622
Smear microscopy (no.)								
Positive	88	137	258	NA	11	4	NA	498
Negative	39	43	42	NA	44	1	NA	169
Unknown	301	239	143	NA	5	20	NA	838
Drug resistance (no.)								
Any resistance ^b	47	51	117	6	1	NA	NA	222
I + R	4	2	4	0	0	NA	NA	10
I + R + S	4	1	12	1	0	NA	NA	18
I + R + E	0	0	1	0	0	NA	NA	1
I + R + S + E	5	3	11	2	0	NA	NA	21
Pansusceptible	284	262	325	17	14	NA	NA	902
Unknown	97	106	1	94	45	NA	NA	381

^a NA, information not available.

^b Any resistance includes all drug-resistant cases (including MDR TB).

subclades. These include, in decreasing order, CAS, 22.5%; ill-defined T clade, 19.5%; EAI, 13.5%; Haarlem, 7.5%; LAM, 7.2%; Beijing, 4.4%; Manu, 2.7%; X, 0.9%; and Bovis, 0.9%. Regarding the distribution of major *M. tuberculosis* clades, CAS strains were predominant in Riyadh (37% of all strains) and EAI were predominant in the coastal areas of Al-Bahah (92%), Jizan (44%), and Damman (22%).

The CAS genotype family prevails in Pakistan, northern India, and neighboring countries of Central Asia, Tanzania, and Madagascar, whereas the EAI family prevails in Southeast Asia (mainly the Philippines, Vietnam, Bangladesh, and Thailand) and East Africa (references 7 and 9 and the updated SITVIT2 database). The prototype signature representing the CAS1-Delhi sublineage (SIT 26) was predominant in Riyadh (7.25% of all isolates) and Damman (6.9%). The prevalence of CAS and EAI families in Saudi Arabia is not surprising, since the country is host to millions of expatriates from the Indian subcontinent and Southeast Asian countries. Furthermore, the country receives more than 2 million pilgrims throughout the year. However, as the genetic diversity of *M. tuberculosis* has not been studied in neighboring countries, it is difficult to ascertain if most of these strains constitute imported or endemic cases. Nonetheless, some of the immigrants came very

early, in the 1960s and 1970s, and settled in different parts of Saudi Arabia. These immigrants primarily came from the Indian subcontinent, Africa, and Central Asia (e.g., Kazakhstan) and form sizeable communities in the country today. In this connection, the link between a specific Saudi clonal complex 1 (SIT 1887) and the CAS genotype suggests that the CAS genotype may have been undergoing evolution in Saudi Arabia (see below).

Regarding the “T” group of families, which represent modern TB strains, a total of five sublineages (T1 to T5) based on single-spacer differences were described in SpolDB4 (7). Among these, the T3-ETH sublineage (prototype SIT 149) was prevalent in our study (Table 2). This low-banding IS6110 clone, identified as early as 1995, has been shown to be frequent in Ethiopia and among Ethiopian immigrants in Denmark and may suggest imported cases of TB (7). The Haarlem clade (7.5% of all isolates) accounted for 11.3% of isolates in Taif; however, a single genotype (SIT 50; H3) accounted for about 7% of the cases in Tabuk, and another (SIT 47; H1 subclade) accounted for 3.5% of the cases in Medina (Table 2; see the supplemental material). Strains belonging to sublineages H2 and H4 (recent results suggest that the H4 sublineage belongs to a new “Ural” lineage that is distinct from Haarlem

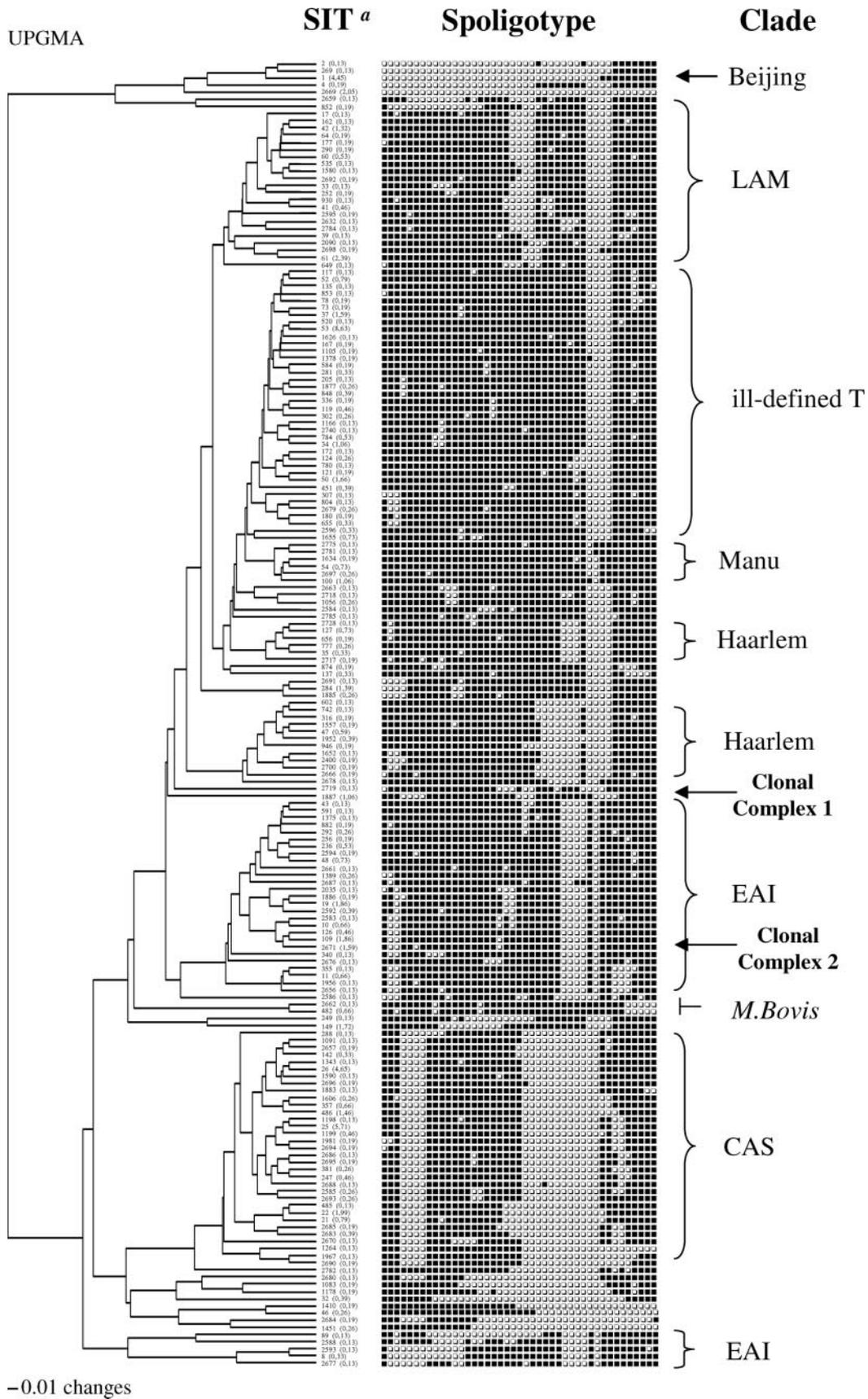


FIG. 1. Tree of spoligotype patterns of clustered *M. tuberculosis* isolates ($n = 1,300$ isolates in 182 clusters, containing 2 to 130 isolates per cluster), created by using PAUP 4.0b software and the unweighted pair group method using arithmetic averages (UPGMA). Unique patterns are not represented (SIT^a, SIT as defined in the SITVIT2 international database). The column below "SIT" shows the individual SIT numbers followed by the percentage of each pattern in this study).

TABLE 2. Distribution of major clades and subclades of *M. tuberculosis* found in Saudi Arabia

Clade	Subclade	SIT ^a	No. (%)							
			Total for this study	Damman (n = 419)	Medina (n = 117)	Riyadh (n = 428)	Taif (n = 443)	Jizan (n = 25)	Tabuk (n = 60)	Al-Baha (n = 13)
Beijing		1	67 (4.4)	29 (6.9)	5 (4.3)	19 (4.4)	14 (3.2)	0	0	0
EAI	Clone 2 ^b	2671		0	0	2	0	11	0	11
	EAI8-MDG	109		1	0	9	8	0	10	0
	EAI3-IND	11 + VP ^c		10	1	3	4	0	0	0
	EAI2-MAN	19 + VP		34	1	4	2	0	0	0
	EAI1-SOM	48 + VP		4	6	6	3	0	0	0
	EAI	VP		42	15	14	10	0	3	1
Total EAI			203 (13.5)	91 (21.7)	23 (19.7)	38 (8.9)	27 (6.1)	11 (44.0)	13 (21.7)	12 (92.3)
CAS	CAS1-KILI	21		0	1	6	5	0	0	0
	Ancestor CAS1-KILI	22		8	2	12	5	2	2	0
	CAS-1DELHI	26		26	2	29	13	0	0	0
	Variant CAS-1DELHI	25 + VP		33	14	65	29	0	8	0
	CAS	VP		8	7	46	14	2	0	0
Total CAS			339 (22.5)	75 (17.9)	26 (22.2)	158 (36.9)	66 (14.9)	4 (16.0)	10 (16.7)	0
T	T	VP		47	15	80	102	9	14	1
	T3-ETH	149		4	0	7	15	0	0	0
Total T			294 (19.5)	51 (12.2)	15 (12.8)	87 (20.3)	117 (26.4)	9 (36.0)	14 (23.3)	1 (7.7)
LAM		VP	109 (7.2)	31 (7.4)	11 (9.4)	13 (3.0)	54 (12.2)	0	0	0
Haarlem		VP	113 (7.5)	28 (6.7)	9 (7.7)	21 (4.9)	50 (11.3)	0	5 (8.3)	0
X		VP	14 (0.9)	0	0	5 (1.17)	9 (2.0)	0	0	0
Manu		VP	41 (2.7)	5 (1.2)	5 (4.3)	4 (0.9)	24 (5.4)	0	3 (5.0)	0
Bovis		VP	13 (0.9)	2 (0.5)	0	11 (2.6)	0	0	0	0
Clone 1 ^d		1887	16 (1.1)	1 (0.24)	0	11 (2.6)	4 (0.9)	0	0	0
UNK ^e			284 (18.9)	106 (25.3)	23 (19.7)	61 (14.2)	78 (17.6)	1 (4.0)	15 (25.0)	0

^a A shared type number is given when the clade is represented in this study by a unique prototype as defined in SpolDB4.

^b SIT 2671 (octal code 467777377413771) had a signature close to the EAI8-MDG subclade (an additional spacer missing at position 6).

^c VP, various patterns of SIT represented this clade and/or subclade (a single major SIT is given when applicable; details are not provided for other minor SITs).

See SpolDB4 and the text for additional information.

^d SIT 1887 (octal profile 703777707770371) defines a clonal complex based on its distinct signature (see the text for further details).

^e UNK, unknown patterns that did not define any particular clade based on available information.

[15; K. Kremer, unpublished results]) were limited to 2 strains of H2 (SIT 2) in Damman and 15 strains of the “Ural” lineage (SIT 127, 11 strains; SIT 777, 4 strains) in Riyadh. In the updated SITVIT2 database, more than 60% of SIT 127 isolates are localized in Armenia, Austria, Finland, Georgia, Iran, and Russia. The related pattern SIT 777 (with a single-spacer difference from SIT 127) found in Saudi Arabia has also been found in Kazakhstan, Russia, and Georgia (n = 26). Our findings on large variations in the distribution of various Haarlem genotypes in the country suggest that these strains were imported (probably through European business communities, as well as Central Asians) at different times and remain localized within the these areas.

This study also showed for the first time an unexpected but moderately high frequency of Beijing genotype (SIT 1, characterized by the absence of spacers 1 to 34) in Saudi Arabia (67/1,505, or 4.4% of all strains). In this study, almost a quarter of the Beijing strains showed drug resistance, and among these, more than half (64%) were MDR (results not shown). The Beijing isolates were found equitably among Saudi and foreign

nationals in Taif (3.2% of strains), Riyadh (4.4%), Medina (4.3%), and Damman (6.9%). Nonetheless, as mentioned above, because our data are not exhaustive and due to a lack of proper registry, these observations for Beijing, as well as for other clades of *M. tuberculosis*, should be considered only descriptive.

Descriptions of specific clonal complexes. Two clonal complexes that may be specific to Saudi Arabia were identified. They are defined by two newly created shared types numbered SIT 1887 (octal code, 703777707770371) and SIT 2671 (octal code, 467777377413771). Both had distinct genetic profiles; the signature of SIT 1887 did not match any of the defined clades in SpolDB4, whereas SIT 2671 had a signature close to that of the EAI8-MDG clade (prototype SIT 109, characterized by the absence of spacers 2, 3, 19, 29 to 32, and 34, with an additional spacer missing at position 6) (7, 9). These two SITs are tentatively referred to as clonal complexes 1 (SIT 1887) and 2 (SIT 2671). Clonal complex 1 (16 cases in total) was essentially limited to Riyadh and Taif and is characterized by the absence of spacers 4 to 7, 22 to 24, and 34 to 37. Clonal complex 2 (SIT

2671, close to the EAI8-MDG subclade found in Madagascar), characterized by the absence of spacers 2, 3, 6, 19, 29 to 32, and 34, was essentially limited to the southwestern regions (Al-Bahah and Jizan), representing 84% and 44% of all cases. It appears to be responsible for most of the ongoing transmission in this region. This clonal complex may be endemic in the region, as it has not yet been reported elsewhere, and the patients for whom nationalities were known (13/24 cases) were all Saudi nationals. Considering that spoligotypes may preferentially evolve by loss of spacers, it may have evolved from SIT 109 by loss of spacer 6. In this context, it is worth mentioning that SIT 109 coexists with SIT 2671 in Saudi Arabia (Table 2). The fact that it is essentially limited to Jizan and Al-Bahah suggests that it either has evolved very recently or was introduced a short time ago. Indeed, Jizan is the largest port of Saudi Arabia, situated very close to Al-Bahah, at a crossroads of population movements due to trade and religious tourism (e.g., it is a privileged port of entry for pilgrims from Madagascar).

To determine if these clonal complexes really reflect a newly emerging genotype within Saudi Arabia, their phylogeographical specificities were confirmed by MIRU-VNTRs. Out of the total of 40 strains characterized as two clones described above, 8/16 clone 1 strains (SIT 1887) and 10/24 clone 2 strains (SIT 2671) were further analyzed by MIRU-VNTRs. Clone 1 strains were split into six patterns, and a comparison with the SITVIT2 database showed that they belonged to MITs 68, 75, 264, 649, and 872 and an orphan pattern. One of these patterns (MIT 75) was shared by three isolates. For clone 1, which did not show the spoligotyping signature of a well-defined lineage, the MIRU pattern obtained (MIT 75; pattern 225525183533) suggested a link to the CAS family. Indeed, MIT 75 has a VNTR (ETR-A to -E) pattern, 42235 which in the SITVIT database belongs to the CAS family. This result, as well as its spoligotyping pattern (SIT 1887; octal code 703777707770371), suggests that clone 1 may be ancestral to the CAS lineage, which constitutes a predominant clade in Saudi Arabia. Clone 2 (SIT 2671; octal code 46777377413771) was split into seven MIRU patterns: MITs 636, 848, and 885 and four orphan patterns). One of the clusters (MIT 885; pattern 225125213322) was shared by four isolates, showing the clonality of this particular strain. Nonetheless, for all 10/24 clone 2 strains studied, MIRU locus 24 was equal to 2 (results not shown), which undoubtedly characterized this clone as the EAI ancestral lineage (22), a fact that further corroborates the EAI-specific spoligotyping signatures of these strains.

Conclusions. This study constitutes a first attempt to describe the genetic population structure and drug resistance of the tubercle bacilli circulating in Saudi Arabia. It corroborates the predominance and historical presence of historic clones of principal genetic group 1, according to the classification of Sreevatsan et al. (21), which are responsible for roughly 45% of all TB cases in Saudi Arabia (CAS, EAI, and Manu, as well as Beijing) (Table 2). Whether these clones are deeply rooted or represent recently expanded epilinked clusters remains to be investigated in future studies. Nonetheless, these values are quite similar to those previously reported from Mumbai, India (17). This observation, together with the fact that organisms presumably of European descent were only rarely found in Saudi Arabia (7.5% Haarlem strains in our study), underlines the fact that TB in this country is both caused by historical

clones of tubercle bacilli undergoing active circulation and due to high migration from countries where TB is endemic. A rapid population growth in high-rise cities and surrounding ghettos where low-skilled construction workers live and a paucity of human resources to tackle health problems caused by a large influx of migrants constitute favorable conditions for TB to disseminate. In addition, the influx of visitors for religious purposes may be considered a potential and permanent risk factor for TB transmission. This influx peaks during pilgrimage to Mecca, which started 1,426 years ago and continues to attract large numbers of visitors. TB was reported among the diseases causing pneumonia during Hajj days (6). The high incidence of clustering after spoligotyping (86%; 182 clusters containing between 2 and 130 isolates per cluster) indicates both a high rate of diversity and probable active transmission of TB in the present study. However, spoligotyping is known to overestimate clustering, and only a systematic second-line, typing, such as MIRU-VNTRs, coupled with a better TB registry and epidemiological investigations would allow us to know the exact rate of ongoing transmission and associated risk factors.

Lastly, this investigation underlined that the component dealing with TB infection control, regulation, and enforcement of public health policies within the National Tuberculosis Program in Saudi Arabia should be improved urgently. Indeed, many of the laboratories that constitute the backbone of the TB control program were recently considered suboptimal in Saudi Arabia; some missed even the most basic requirements for bacteriological diagnostics of TB, as well as a basic registry system (2, 5). Adequate professional training of the health care workers in these laboratories may be helpful to solve the issue of an efficient quality control system, as well as the updating and renovation of many TB laboratories in the country to international standards (2, 5).

ACKNOWLEDGMENTS

We thank the Sultan Al-Sedairy, Executive Director of the Research Center, without whose help this work would not have seen the light.

Many thanks go to King Abdulaziz City for Science and Technology for funding this work under project AT-21-8. The work performed in Nalin Rastogi's laboratory was partially funded by the European Regional Development Fund, European Commission (ERDF/FEDER, A34-05). Thierry Zozio received a Ph.D. fellowship awarded by the European Union and the Regional Council of Guadeloupe and the International Network of Pasteur Institutes.

REFERENCES

1. Al-Hajjaj, M. S., F. A. Al-Kassimi, A. F. Al-Mobeireek, and A. H. Alzeer. 2001. Progressive rise of *Mycobacterium tuberculosis* resistance to rifampicin and streptomycin in Riyadh, Saudi Arabia. *Respirology* **6**:317-322.
2. Al-Hajjaj, S. A., and F. A. Alrabiah. 2004. Role of tuberculosis laboratories in Saudi Arabia. A call to implement standardized procedures. *Saudi Med. J.* **25**:1545-1548.
3. Al-Kassimi, F. A., A. K. Abdullah, M. S. al-Hajjaj, I. O. al-Orainey, E. A. Bangboye, and M. N. Chowdhury. 1993. Nationwide community survey of tuberculosis epidemiology in Saudi Arabia. *Tuber. Lung. Dis.* **74**:254-260.
4. Al-Mazrou, Y. Y., T. A. Khoja, K. M. Aziz, and A. M. Salem. 1997. High proportion of multi-drug resistant *Mycobacterium tuberculosis* in Saudi Arabia. *Scand. J. Infect. Dis.* **29**:323.
5. Alrajhi, A. A., S. Abdulwahab, E. Almodovar, and H. M. Al-Abdely. 2002. Risk factors for drug-resistant *Mycobacterium tuberculosis* in Saudi Arabia. *Saudi Med. J.* **23**:305-310.
6. Alzeer, A., A. Mashlah, N. Fakim, N. Al-Sugair, M. Al-Hedaithy, S. Al-Majed, and G. Jamjoom. 1998. Tuberculosis is the commonest cause of pneumonia requiring hospitalization during Hajj (pilgrimage to Makkah). *J. Infect.* **36**:303-306.
7. Brudey, K., J. R. Driscoll, L. Rigouts, W. M. Prodinger, A. Gori, S. A. Al-Hajjaj, C. Allix, L. Aristimuno, J. Arora, V. Baumanis, L. Binder, P.

- Cafrune, A. Cataldi, S. Cheong, R. Diel, C. Ellermeier, J. T. Evans, M. Fauville-Dufau, S. Ferdinand, D. Garcia de Viedma, C. Garzelli, L. Gazzola, H. M. Gomes, M. C. Gutierrez, P. M. Hawkey, P. D. van Helden, G. V. Kadival, B. N. Kreiswirth, K. Kremer, M. Kubin, S. P. Kulkarni, B. Liens, T. Lillebaek, M. L. Ho, C. Martin, C. Martin, I. Mokrousov, O. Narvskaja, Y. F. Ngeow, L. Naumann, S. Niemann, I. Parwati, Z. Rahim, V. Rasolof-Razanamparany, T. Rasolonavalona, M. L. Rossetti, S. Rusch-Gerdes, A. Sajduda, S. Samper, I. G. Shemyakin, U. B. Singh, A. Somoskovi, R. A. Skuce, D. van Soolingen, E. M. Streicher, P. N. Suffys, E. Tortoli, T. Tracevska, V. Vincent, T. C. Victor, R. M. Warren, S. F. Yap, K. Zaman, F. Portaels, N. Rastogi, and C. Sola. 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* 6:23.
8. Dale, J. W., H. Al-Ghusein, S. Al-Hashmi, P. Butcher, A. L. Dickens, F. Drobniowski, K. J. Forbes, S. H. Gillespie, D. Lamprecht, T. D. McHugh, R. Pitman, N. Rastogi, A. T. Smith, C. Sola, and H. Yesilkaya. 2003. Evolutionary relationships among strains of *Mycobacterium tuberculosis* with few copies of IS6110. *J. Bacteriol.* 185:2555–2562.
 9. Filliol, I., J. R. Driscoll, D. Van Soolingen, B. N. Kreiswirth, K. Kremer, G. Valetudie, D. D. Anh, R. Barlow, D. Banerjee, P. J. Bifani, K. Brudey, A. Cataldi, R. C. Cooksey, D. V. Cousins, J. W. Dale, O. A. Dellagostin, F. Drobniowski, G. Engelmann, S. Ferdinand, D. Gascoyne-Binzi, M. Gordon, M. C. Gutierrez, W. H. Haas, H. Heersma, G. Kallenius, E. Kassa-Kelembho, T. Koivula, H. M. Ly, A. Makristathis, C. Mammina, G. Martin, P. Mostrom, I. Mokrousov, V. Narbonne, O. Narvskaya, A. Nastasi, S. N. Niobe-Eyangoh, J. W. Pape, V. Rasolof-Razanamparany, M. Ridell, M. L. Rossetti, F. Stauffer, P. N. Suffys, H. Takiff, J. Texier-Maugein, V. Vincent, J. H. De Waard, C. Sola, and N. Rastogi. 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. *Emerg. Infect. Dis.* 8:1347–1349.
 10. Filliol, I., A. S. Motiwala, M. Cavatore, W. Qi, M. H. Hazbon, M. Bobadilla del Valle, J. Fyfe, L. Garcia-Garcia, N. Rastogi, C. Sola, T. Zozio, M. I. Guerrero, C. I. Leon, J. Crabtree, S. Angiuoli, K. D. Eisenach, R. Durmaz, M. L. Joloba, A. Rendon, J. Sifuentes-Osornio, A. Ponce de Leon, M. D. Cave, R. Fleischmann, T. S. Whittam, and D. Alland. 2006. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *J. Bacteriol.* 188:759–772 (Erratum, 188:3162–3163.).
 11. Huard, R. C., M. Fabre, P. de Haas, L. C. Lazzarini, D. van Soolingen, D. Cousins, and J. L. Ho. 2006. Novel genetic polymorphisms that further delineate the phylogeny of the *Mycobacterium tuberculosis* complex. *J. Bacteriol.* 188:4271–4287.
 12. Kam, K. M., C. W. Yip, L. W. Tse, K. L. Wong, T. K. Lam, K. Kremer, K., B. K. Au, and D. Van Soolingen. 2005. Utility of mycobacterial interspersed repetitive unit typing for differentiating multidrug-resistant *Mycobacterium tuberculosis* isolates of the Beijing family. *J. Clin. Microbiol.* 43:306–313.
 13. Kamerbeek, J., L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, D., S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, and J. Van Embden. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J. Clin. Microbiol.* 35:907–914.
 14. Kanduma, E., T. D. McHugh, and S. H. Gillespie. 2003. Molecular methods for *Mycobacterium tuberculosis* strain typing: a user's guide. *J. Appl. Microbiol.* 94:781–791.
 15. Kovalev, S. Y., E. Y. Kamaev, M. A. Kravchenko, N. E. Kurepina, and S. N. Skorniakov. 2005. Genetic analysis of *Mycobacterium tuberculosis* strains isolated in Ural region, Russian Federation, by MIRU-VNTR genotyping. *Int. J. Tuberc. Lung. Dis.* 9:746–752.
 16. Kremer, K., D. van Soolingen, R. Frothingham, W. H. Haas, P. W. Hermans, C. Martin, P. Pallitapongarnpim, B. B. Plikaytis, L. W. Riley, M. A. Yakrus, J. M. Musser, and J. D. Van Embden. 1999. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J. Clin. Microbiol.* 37:2607–2618.
 17. Kulkarni, S., C. Sola, I. Filliol, N. Rastogi, and G. Kadival. 2005. Spoligotyping of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mumbai, India. *Res. Microbiol.* 156:588–596.
 18. Milaat, W., A. Ali, H. Afif, and T. Ghabrah. 1994. Epidemiology of tuberculosis in Jeddah region, Saudi Arabia. *Saudi Med. J.* 15:133–137.
 19. Schiott, C. R., H. C. Engbaek, B. Vergmann, M. Al Motez, and I. Kassim. 1984. Resistant strains of *Mycobacterium tuberculosis* in the Gizan Area, Saudi Arabia. *Ugeskr. Laeger.* 146:4024–4026.
 20. Sola, C., I. Filliol, E. Legrand, I. Mokrousov, and N. Rastogi. 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR, and DR-based spoligotyping suggests the existence of two new phylogeographical clades. *J. Mol. Evol.* 53:680–689.
 21. Sreevatsan, S., X. Pan, K. E. Stockbauer, N. D. Connell, B. N. Kreiswirth, T. S. Whittam, and J. M. Musser. 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci. USA* 94:9869–9874.
 22. Supply, P., S. Lesjean, E. Savine, K. Kremer, D. van Soolingen, and C. Locht. 2001. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *J. Clin. Microbiol.* 39:3563–3571.
 23. van Soolingen, D. 2001. Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J. Intern. Med.* 249:1–26.
 24. van Soolingen, D., P. E. de Haas, P. W. Hermans, and J. D. van Embden. 1994. DNA fingerprinting of *Mycobacterium tuberculosis*. *Methods Enzymol.* 235:196–205.