

Mycobacterium tuberculosis spoligotypes circulating in the Lebanese population: a retrospective study

N. Bedrossian,¹ M. Hamze,^{1,2} A.K. Rahmo,³ A. Jurjus,⁴ J. Saliba,⁴ F. Dabboussi^{1,2} and W. Karam⁵

أنماط عديدات النكليوتيد البينية للمتفطرات السلية التي تسري بين السكان اللبنانيين: دراسة استيعادية

نوره بدروسيان، منذر حمزة، عبد القادر رحمو، عبده جرجس، جيسكا صليبا، فؤاد دبوسي، وليد كرم

الخلاصة: قد يعود التنميط الجيني للمتفطرات السلية بالفوائد على المستوى الوطني في لبنان من أجل رصد الاستجابة العلاجية للمعالجة القصيرة الأمد تحت الإشراف المباشر (دوتس). وتهدف هذه الدراسة إلى تحديد خصائص المستفردات السيرية من أنماط عديدات النكليوتيد البينية للمتفطرات السلية التي جمعت من المرضى في الفترة من نيسان/أبريل 2004 وتشرين الأول/أكتوبر 2005 من جميع المناطق اللبنانية. وبلغ عدد المستفردات 60، وزرعت وتم التعرف على خصائصها الكيميائية الحيوية، وتضخيم المستخلصات من الدنا بالتفاعل السلسلي للبوليمراز، وتحديد أنماطها الجينية من خلال التنميط التفصيلي لأنماط عديدات النكليوتيد البينية. وتم التعرف على 13 نمطاً من الذراري المعقدة لفصائل المتفطرات السلية: 41.6% منها من ذراري تنتمي إلى فصيلة T1، و25% تنتمي إلى فصيلة LAM9، و10% منها تنتمي إلى فصيلة هارلم 3، و3.3% منها تنتمي إلى كل من الفصائل CAS، و LAM و BCG، والفصيلة 36، في حين كانت 1.7% منها تنتمي إلى كل من الفصائل هارلم 1، و LAM 10، و S، والمتفطرات الأفريقية، و X1، و T3. ولم يكن الغياب الملحوظ لفصائل بيجين والهندية الشرق أفريقية متماشياً مع النماذج المسجلة في البلدان المجاورة؛ مما يدل على ضرورة القيام بدراسة أكثر إحاطة وشمولاً للسكان اللبنانيين من أجل التعرف بدقة على الفصائل الأكثر انتشاراً في لبنان.

ABSTRACT Genotyping *Mycobacterium tuberculosis* in Lebanon on the national level may be beneficial for assessing patients and monitoring the therapeutic response to DOTS. This study aimed to characterize the spoligotypes of clinical isolates of *M. tuberculosis* patients collected between April 2004 and October 2005 from all Lebanese provinces. Isolates ($n = 60$) were cultured and identified by their biochemical characteristics. DNA extracts of these samples were amplified by PCR and genotyped by spoligotyping. Thirteen (13) patterns of *M. tuberculosis* complex family strains were identified: 41.6% of the strains belonged to the T 1 family, 25.0% to LAM 9, 10.0% to Haarlem 3, 3.3% to each of CAS, LAM 8, BCG and Family 36 and 1.7% to each of Haarlem 1, LAM 10, S, *M. africanum*, X 1 and T 3 families. The noticeable absence of Beijing and East African Indian families was not consistent with the patterns reported in neighbouring countries. A more inclusive study of the Lebanese population is necessary to accurately identify most of the prevailing families in the country.

Spoligotypes de *Mycobacterium tuberculosis* circulant dans la population libanaise : une étude rétrospective

RÉSUMÉ Le génotypage de *Mycobacterium tuberculosis* au Liban au niveau national peut être utile pour dépister les patients et surveiller la réponse au traitement de brève durée sous surveillance directe ou DOTS. L'objectif de l'étude visait à caractériser les spoligotypes des isolats cliniques de *M. tuberculosis* prélevés entre avril 2004 et octobre 2005 chez des patients de toutes les provinces libanaises. Les isolats ($n = 60$) ont été mis en culture puis identifiés en fonction de leurs caractéristiques biochimiques. Les extraits d'ADN de ces échantillons ont été amplifiés par PCR puis génotypés par spoligotypage. Treize (13) souches différentes de la famille du complexe *M. tuberculosis* ont été identifiées : 41,6 % des souches appartenaient à la famille T 1 ; 25,0 % à la famille LAM 9 ; 10,0 % à la famille Haarlem 3 ; 3,3 % chacune aux familles CAS, LAM 8, BCG et à la famille 36 et 1,7 % chacune aux familles Haarlem 1, LAM 10, S, *M. africanum*, X 1 et T 3. L'absence remarquable des familles Beijing, indiennes et d'Afrique de l'Est ne coïncidait pas avec les tendances rapportées dans les pays voisins. Une étude plus globale de la population libanaise est nécessaire pour identifier avec précision la plupart des familles dominantes dans notre pays.

¹Faculty of Medicine; ²Department of Biology, Lebanese University, Hadath Campus, Hadath, Lebanon (Correspondence to M. Hamze: mhamze@monzerhamze.com). ³AZM Center for Biotechnology, Lebanese University, Tripoli, Lebanon. ⁴National Commission of Biotechnology, Damascus, Syrian Arab Republic. ⁵Department of Human Morphology, American University of Beirut, Beirut, Lebanon.

Received: 12/08/11; accepted: 05/02/12

Introduction

After HIV/AIDS, tuberculosis (TB) is the second most common cause of death due to an infectious disease, and current trends suggest that TB will still be among the 10 leading causes of global disease burden in the year 2020 [1].

Molecular tools have enhanced our understanding of the epidemiology of TB by providing new insight into the transmission, dynamics, source and spread of *Mycobacterium tuberculosis* clones [2,3], which are often difficult to identify by traditional epidemiological investigations alone [4]. International databases, such as the World Spoligotyping Database, *SpolDB3.0*, have revealed the clonal structure of *M. tuberculosis* isolates in different geographical settings. The *SpolDB4.0* database further defines super-families specific to certain locations [5]. The abundance of polymorphism indicates that transposition and homologous recombination are the major events contributing to the diversity of *M. tuberculosis* strains [6].

Due to the importance and gravity of this disease in Lebanon, the Ministry of Public Health set up the national TB programme in 1992, which started the DOTS implementation/expansion in 1998. DOTS refers to a broad TB control strategy outlined by the World Health Organization (WHO) which aims to halt the spread of the infection and of multidrug resistant TB strains [7]. The total number of cases in Lebanon fell from 993 in 1993 to 195 in 2005, with pulmonary TB representing 60% to 65% of all TB cases. At present, according to the WHO report, the incidence of TB per 100 000 people in Lebanon was reported as 14 in 2008, 15 in 2009 and 17 in 2010 (report released in 2011) [8].

Genotyping *M. tuberculosis* in Lebanon on the national level in parallel with the DOTS programme may prove beneficial for assessing patients and monitoring the therapeutic response

to DOTS. In a previous study, characterization of *M. tuberculosis* in Lebanese patients by double-repetitive-element polymerase chain reaction (DRE-PCR) indicated the presence of several genotypes with evidence of certain groups specific to geographic areas, implicating separate evolution of *M. tuberculosis* strains [9]. This investigation reports on the first nationwide study to characterize by spoligotyping clinical isolates of *M. tuberculosis* from different regions of Lebanon. It will further provide insight into the future dynamics of the disease in our country and even among neighbouring countries.

Methods

Sampling of patients

Following an agreement with the national TB programme of the Lebanese Ministry of Health, we obtained sputum samples from newly detected pulmonary TB cases (60 samples). Samples were collected between April 2004 and October 2005 from all Lebanese provinces (*muhafazat*) and were stored in the laboratory of the hospital at the Middle East Health Centre in Bsalim/Metn in Lebanon. Relevant information (e.g. sex, age, location, new/old case) was also obtained for each of the samples; however no information about the HIV status was available.

As this was a retrospective study, and the samples tested consisted of stock bacterial cultures and did not directly involve any human subjects, ethical approval was not required.

Specimen preparation and culture

Specimens were decontaminated by the 2% N-acetyl-L-cysteine NaOH method. After neutralization and centrifugation, 0.2 mL of the concentrated specimen was inoculated onto one slant of Lowenstein-Jensen (LJ) medium (Becton Dickinson), 0.5 mL was also inoculated into modified Middlebrook

7H9 broth (BD BBL MGIT, Becton Dickinson) supplemented with PANTA (BBL MGIT PANTA, Becton Dickinson) and oleic acid, albumin, dextrose, and catalase (OADC) enrichment. The 2 media were incubated at 35–37 °C, the LJ slant was examined for growth twice weekly. The BD BBL MGIT tube was read daily with ultraviolet light starting on the 2nd day of incubation using a positive control and a negative control. Fluorescence is detected as a bright orange colour in the bottom of the tube and also as an orange reflection on the meniscus. A positive tube was subcultured and an acid-fast smear prepared. A positive acid-fast smear result indicated the presumptive presence of viable microorganisms in the tube.

Identification of strains

Identity of strains was based on the following biochemical characteristics: production of niacin, nitrate reductase, catalase at laboratory temperature, catalase at 68 °C, hydrolysis of Tween 80 in 10 days, urease in 18 hours, arylsulfatase in 3 days [10].

DNA extraction

A loopful of each culture was suspended in animal-tissue lysis buffer (0.2 mL). DNA samples were extracted in a laboratory free of all mycobacterial products using QIAamp DNA blood mini kit (Qiagen). The accompanied procedure was adhered to, except for the incubation period which was extended to 3 hours. DNA was eluted in 100 µL PCR water (Gibco).

PCR amplification

All DNA samples were positive for the IS6110 insertion element as determined by PCR amplification followed by detection by agarose gel electrophoresis using the method of Eisenach [11]. PCR set-up was also performed in a remote area separate from all subsequent procedures. Measures to avoid cross- and carry-over PCR

contaminations were observed: use of aerosol barrier-fitted pipette tips and molecular biology-grade water and reagents and frequent decontamination of work areas with diluted bleach and short-ultraviolet surface performed regularly between experiments.

Spoligotyping

Spoligotyping was performed for each sample in duplicate and was initially repeated twice to assess the reproducibility of the method. Duplicate positive and negative controls were included in each run. DNA was replaced with water in negative control reactions and DNA of *M. tuberculosis* strains H37Rv and *M. bovis* BCG P3 supplied in the kit were used in positive control reactions. The procedure was performed using a reverse dot-blot spoligotyping kit with chemiluminescent detection (Isogen Bioscience) as follows: 5 µL of mycobacterial DNA was added to a PCR amplification reaction in which the forward direct repeat (DR) primer was biotinylated. Amplification was performed for 25 cycles as recommended by the manufacturer.

Results

The prevalence of the different *M. tuberculosis* complex family strains in the tested samples are summarized in Table 1. Table 2 shows a detailed analysis of the different *M. tuberculosis* complex family clades with the individual shared type (ST) designation, the number of isolates and the corresponding lack of spacers for each.

Of the samples tested 25 (41.6%) belonged to the T1 family. In this clade, all isolates lacked spacers 33–36. Separate STs exhibited additional absence of spacers: 17 strains were ST 53 and lacked spacers 39–41; 1 was ST 154 and lacked spacers 5; 1 was ST 751 and lacked spacers 1–3, 12–15, 21 and 39–42; 1 was ST 879 and lacked spacers 10, 21 and 39; and 3 isolates were ST

Table 1 Prevalence of the different *Mycobacterium tuberculosis* complex family strains in the tested samples from Lebanese patients (*n* = 60)

Family strains	No.	%
T1	25	41.6
LAM 9	15	25.0
Haarlem 3	6	10.0
CAS	2	3.3
LAM 8	2	3.3
Family 36	2	3.3
BCG	2	3.3
Haarlem 1	1	1.7
LAM 10	1	1.7
S family	1	1.7
<i>M. africanum</i>	1	1.7
X1	1	1.7
T3	1	1.7
Total	60	100.0

272 and lacked spacers 1–4 and 12–13. Two isolates could not be classified under any of the shared types: 1 having an octal designation 017711727760760 and lacking spacers 1–5, 14–15, 21–22, 32–36 and 42, and 1 with an octal designation 777737777760371 and lacking spacers 2, 5, 13–15, 19, 32–35, 37–39 and 42.

The second predominant family in our results was the LAM 9 family in 15 samples (25.0%). Most of these strains (14/15) were prototyped as ST 41, 13 simultaneously lacking 20–24, 26–27 and 33–36 spacers and 1 in addition lacking 13 and 15. The other LAM 9 strain belonged to ST 42, lacking spacers 21–24 and 33–36. Another 3.3% of samples (*n* = 2) belonged to the LAM 8 family: 1 was prototyped as ST 511, lacking simultaneously 13–17, 19–31, 33–36 and 39–42 spacers; the other with an octal format 777777400000371 (lacking 14–17, 19–31, 33–36 and 39–42 spacers) and was unclassified. One sample belonged to LAM 10 (1.7%), lacking 23–25 and 33–36, and was prototyped as ST 61.

In the Haarlem (H) lineage, 6 samples (10.0%) belonged to Haarlem 3. One was prototyped as ST 35 (lacking 13, 29–31, 33–36 and 40 spacers),

another as ST 418 (lacking 3–5, 31 and 33–36), and 4 others as ST 775, 2 of which lacked 28–37 and 2 others additionally lacking spacer 13.

There were 10 other family strains identified. One sample (1.7%) belonged to the T3 family as ST 149, lacking spacers 33–36 and 10–19. One specimen (1.7%) belonged to Haarlem 1 as ST 602 and lacked spacers 25–36. Two samples (3.3%) were Asian (CAS) lineage; 1 belonged to ST 25 (lacking 4–7, 21–34 and 37–38 spacers) and 1 to ST 1199 (lacking 2, 4–7, 10, 13–15, 21–34, 37–39 and 42 spacers). Two of the samples (3.3%) belonged to Family 36; 1 was ST 4 and lacked 1–24 and 33–36 and the other was ST 125 and additionally lacked spacers 40–41. One sample (1.7%) belonged to the X1 family. With an octal format 017100777760760 and the absence of spacers 1–5, 10, 12–15, 21–22, 32–36 and 42, it could not be classified under any of the shared types. One strain (1.7%) belonged to *M. africanum* ST 536, with the absence of 4–19 and 39 spacers. Two samples (3.3%) belonged to *M. bovis* BCG were ST 482, lacking spacers 3, 9, 16 and 39–43. One specimen (1.7%) was characterized as the S family (ST 784) and lacked spacers 9–10, 33–36 and 40.

Table 2 Summary of the different clades of *Mycobacterium tuberculosis* complex family strains identified, their shared type designations, the number of isolates and the absence of spacers for the specimen tested from Lebanese patients

Clade	Shared type	No. of isolates	Absent spacers	No. of isolates
T 1	53	17	33-36	13
			33-36, 39-41	2
			33-36, 40-41	1
			33-36, 39-40	1
T 1	154	1	33-36, 5	1
T 1	751	1	33-36, 1-3, 12-15, 21, 39, 42	1
T 1	879	1	33-36, 10, 21, 39	1
T 1	272	3	33-36, 1-4, 12-13	3
T 1	Unclassified	2	1-5, 14-15, 21-22, 32-36, 42	2
T 3	149	1	33-36, 10-19	1
LAM 9	41	14	33-36, 20-24, 26-27	13
			33-36, 20-24, 26-27, 13, 15	1
LAM 9	42	1	33-36, 21-24	1
LAM 8	511	1	33-36, 13-17, 19-31, 39-42	1
LAM 8	Unclassified	1	14-17, 19-31, 33-36, 39-42	1
LAM 10	61	1	33-36, 23-25	1
Haarlem 1	602	1	25-36	1
Haarlem 3	35	1	33-36, 13, 29-31, 40	1
Haarlem 3	418	1	33-36, 3-5, 31	1
Haarlem 3	775	4	28-37	2
			28-37, 13	2
CAS	25	1	4-7, 21-34, 37-38	1
CAS	1199	1	4-7, 13-15, 21-34, 37-39, 2, 10, 42	1
BCG	482	2	39-43, 3, 9, 16	2
BCG (control)	683	1	39-43, 8-12, 3, 16	1
M. africanum	536	1	4-19, 39	1
Family 36	4	1	33-36, 1-24	1
Family 36	125	1	33-36, 1-24, 40-41	1
S	784	1	33-36, 9-10, 40	1
X 1	Unclassified	1	1-5, 10, 12-15, 21-22, 32-36, 42	1
H37Rv (control)	451	1	33-36, 20-21	1

Discussion

Many genetic loci within the *M. tuberculosis* complex genomes are polymorphic and may be used for molecular evolutionary studies [12]. Among these, the DR locus, which consists of alternating identical DRs and variable spacers, can be genotyped using spoligotyping methods [13].

Several studies have shown that there is an apparent stable association of TB bacilli populations with their human hosts in various environments

[14], hence strong phylogeographical clustering of TB bacilli populations has been reported.

Our results found that the majority of the samples tested (41.6%) belonged to the T 1 family. This family is widely present in populations in all continents and corresponds to about 30% of all entries in the international database [15]. One sample (1.7%) belonged to the T 3 family, which has recently been sub-classified as T 3 Ethiopia [15]. The presence of this type with an Ethiopian origin can be explained by the migration of thousands

of Ethiopian women in the past few years to work as housemaids in Lebanon.

The second predominant family in our results was the LAM family: 25.0% were LAM 9, 3.3% LAM 8 and 1.7% LAM 10. In general, the LAM family has been found to be most common in Venezuela, in the Mediterranean basin and in the Caribbean region [16]. Lebanon's geographical location in the Mediterranean basin and the tendency of the Lebanese, since ancient times, to migrate to Venezuela, Central America and to some other African countries,

with intermittent visits to their mother country could explain our findings.

In the Haarlem (H) lineage, 10.0% of our samples belonged to Haarlem 3 and 1.7% to Haarlem 1. In Europe, this lineage represents about 25.0% of the isolates [15]. Outside Europe, the Haarlem strains were mainly found in Central America and the Caribbean (about 25%), suggesting a link of Haarlem to the post-Columbus European colonization [17]. Today its widespread distribution in different geographical regions of the world such as Asia (except for India [18]), Europe and Africa has been documented [19]. The presence of Haarlem may always be related to European settlers [14] and in our areas, to the historical contacts that existed between Lebanon and many of the European communities.

The Central Asian (CAS) lineage covered 3.3% of the samples tested. Strains of this family have been reported in different countries of the Middle East (Islamic Republic of Iran, Pakistan, India and Afghanistan) [20], and to a lesser extent in several other regions (Africa 5.3%, Central America 0.1%, Europe 3.3%, Far East Asia 0.4%, North America 3.3% and Oceania 4.8%).

Another 3.3% of the samples belonged to Family 36. Until now, this family has been identified as solely of USA origin [21].

There were 3.3% of samples belonging to *M. bovis*. This strain is characterized by the presence of spacers 33 and 34 [22] and shares 99.95% identity with *M. tuberculosis* [23]. Humans are rarely affected, but people in some occupations such as veterinarians, farmers and abattoir workers may be more at risk [23].

There was 1 sample (1.7%) belonging to the X1 family. The X super-family is subdivided into at least 3 distinct families (X1 to X3) [24]. Its prevalence in North America is 21.5%, and in Central America 11.9%. It has been linked to an Anglo-Saxon ancestry, since it has been encountered in the UK and in English-colonized areas such as the

USA, Australia, South Africa and the Caribbean [25]. However, according to other investigations, this group of strains is currently correlated with African-Americans, a fact that may not represent the ancestry of this genotype [26].

Another sample (1.7%) belonged to *M. africanum*. It was first described in Senegal in 1968, in comparison to half of the smear-positive pulmonary TB cases in West Africa, but has not been found in other geographical areas except among recent West African migrants [27]. There is a specificity of the *M. africanum* lineage for the African continent, with a decreasing gradient of prevalence from west to east [27].

Finally, 1 sample (1.7%) was characterized as the S family. This lineage, which is highly prevalent in Sicily and Sardinia, is thought to be identical to the F28 in South Africa [15], although its origin remains unknown.

Noticeably none of the samples tested belonged to the Beijing family (characterized by the absence of spacers 1–34), which is quite prevalent in countries of the Middle East and Far East. Actually, they represent around 50% of the strains in Far East Asia, 16.5% in the Middle East and Central Asia, 17.2% in Oceania [15], 10% in India [18], and 13% of isolates globally [15]. This genotype, which may have been endemic in China for a long time [28], is emerging in some parts of the world, especially in countries of the former Soviet Union, and to a lesser extent in the Western hemisphere [29]. Either this genotype is actually absent in our country or additional samples should be tested in order to confirm this finding.

Another spoligotype that could not be detected was the East African Indian lineage. This super-family is highly prevalent in Far East Asia (33.8%), in the Middle East and Central Asia (24.3%) and in Oceania (22.9%).

The current study demonstrated the circulation of several *M. tuberculosis* spoligotypes in the Lebanese population. The genetic diversity of *M. tuberculosis* isolates

from different regions of Lebanon necessitates additional studies by the simple spoligotyping test in order to achieve accurate epidemiological tracing of specific clusters and shared types which, in conjunction with TB control programmes, will help in prevention and treatment of TB. Determination of the sensitivity of the identified spoligotypes to the various antituberculosis drugs is in progress. This will provide further evidence of the clinical applicability of this work and quicker recognition of the long-feared virulent multidrug resistant strains.

Conclusion

Our results fit fairly well the molecular epidemiology of *M. tuberculosis* genotypes described in the international database. Four unclassified strains were detected: 2 strains belonging to the T1 family (octal formats 017711727760760 and 777737777760371), 1 strain belonging to LAM8 (777777400000371) and 1 strain belonging to X1 (017100777760760). However, the absence of Beijing and East African Indian is not consistent with the patterns reported in neighbouring countries, despite the fact that Lebanon is open to workers from the Far and Middle East (e.g. Philippines, Bangladesh, Nepal, Sri Lanka, India, Pakistan, Islamic Republic of Iran, Russia, Iraq), from Africa (e.g. Egypt, Sudan, Ethiopia), in addition to the multinational forces lodging in south Lebanon. A more inclusive study of the Lebanese population is necessary to more accurately identify most of the prevailing families in our country and further extrapolate the benefits for prevention and possibly treatment of TB.

Acknowledgements

This work was totally funded by the Lebanese University. The *SpolDB4* database was kindly provided by Christopher Sola and was used as a basis for data analysis.

References

1. Murray M. Determinants of cluster distribution in the molecular epidemiology of tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99:1538–1543.
2. Gopaul KK et al. Progression toward an improved DNA amplification-based typing technique in the study of *Mycobacterium tuberculosis* epidemiology. *Journal of Clinical Microbiology*, 2006, 44:2492–2498.
3. Mathema B et al. Molecular epidemiology of tuberculosis: current insights. *Clinical Microbiology Reviews*, 2006, 19:658–685.
4. Durmaz R et al. Population-based molecular epidemiological study of tuberculosis in Malatya, Turkey. *Journal of Clinical Microbiology*, 2007, 45:4027–4035.
5. Sola C et al. Spoligotype database of *Mycobacterium tuberculosis*: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives. *Emerging Infectious Diseases*, 2001, 7:390–396.
6. Sreevatsan S 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*, 1997, 94:9869–9874.
7. *The five elements of DOTS*. World Health Organization [online factsheet]. (<http://www.who.int/tb/dots/whatisdots/en/index2.html>, accessed 12 November 2012).
8. *Global tuberculosis control report 2012*. Geneva, World Health Organization, 2012.
9. Hamze M, Rahmo A, Saade M. Characterization of *Mycobacterium tuberculosis* of Lebanese patients by double-repetitive-element polymerase chain reaction. *Eastern Mediterranean Health Journal*, 2010, 16:812–819.
10. Shinnick TM, Good RC. Mycobacterial taxonomy. *European Journal of Clinical Microbiology and Infectious Diseases*, 1994, 13:884–901.
11. Eisenach KD et al. Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *Journal of Infectious Diseases*, 1990, 161:977–981.
12. Mostowy S, Behr MA. The origin and evolution of *Mycobacterium tuberculosis*. *Clinics in Chest Medicine*, 2005, 26:207–216, v–vi.
13. Groenen PM et al. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Molecular Microbiology*, 1993, 10:1057–1065.
14. Hirsh AE et al. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101:4871–4876.
15. Brudey K et al. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiology*, 2006, 6:23–50.
16. Sola C, Rastogi N. *Molecular epidemiology and population genetics of tuberculosis*. Kuala Lumpur, Malaysia, Academy of Sciences of Malaysia, 2006.
17. Duchêne V et al. Phylogenetic reconstruction of *Mycobacterium tuberculosis* within four settings of the Caribbean region: tree comparative analyse and first appraisal on their phylogeography. *Infection, Genetics and Evolution*, 2004, 4:5–14.
18. Kulkarni S et al. Spoligotyping of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mumbai, India. *Research in Microbiology*, 2005, 156:588–596.
19. Mardassi H et al. Tuberculosis due to resistant Haarlem strain, Tunisia. *Emerging Infectious Diseases*, 2005, 11:957–961.
20. Farnia P et al. Prevalence of Haarlem I and Beijing types of *Mycobacterium tuberculosis* strains in Iranian and Afghan MDR-TB patients. *Journal of Infection*, 2006, 53:331–336.
21. *TB insight. Tuberculosis tracking and control* [website] (<http://tbinsight.cs.rpi.edu>, accessed 28 November 2012).
22. Singh UB et al. Predominant tuberculosis spoligotypes, Delhi, India. *Emerging Infectious Diseases*, 2004, 10:1138–1142.
23. Taylor GM et al. Genotypic analysis of *Mycobacterium tuberculosis* from medieval human remains. *Microbiology*, 1999, 145:899–904.
24. Filliol I et al. Global distribution of *Mycobacterium tuberculosis* spoligotypes. *Emerging Infectious Diseases*, 2002, 8:1347–1349.
25. Sebban M et al. A data-mining approach to spacer oligonucleotide typing of *Mycobacterium tuberculosis*. *Bioinformatics*, 2002, 18:235–243.
26. Kempf MC et al. Long-term molecular analysis of tuberculosis strains in Alabama, a state characterized by a largely indigenous, low-risk population. *Journal of Clinical Microbiology*, 2005, 43:870–878.
27. De Jong BC et al. Use of spoligotyping and large sequence polymorphisms to study the population structure of the *Mycobacterium tuberculosis* complex in a cohort study of consecutive smear-positive tuberculosis cases in The Gambia. *Journal of Clinical Microbiology*, 2009, 47:994–1001.
28. Qian L et al. Retrospective analysis of the Beijing family of *Mycobacterium tuberculosis* in preserved lung tissues. *Journal of Clinical Microbiology*, 1999, 37:471–474.
29. Glynn JR et al. Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerging Infectious Diseases*, 2002, 8:843–849.