Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 7 (2007) 609–617
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Contribution of spoligotyping to the characterization of the population structure of *Mycobacterium tuberculosis* isolates in Portugal

Suzana David [a,*], Diana Raposo Ribeiro [a], Abílio Antunes [b], Clara Portugal [c], Luísa Sancho [c], José Germano de Sousa [c]

[a] *Unidade de Ensino e Investigação (UEI) de Micobactérias, Instituto de Higiene e Medicina Tropical (IHMT), Universidade Nova de Lisboa (UNL), Lisbon, Portugal*
[b] *UEI de Clínica e Doenças Tropicais/IHMT/UNL, Lisbon, Portugal*
[c] *Serviço de Patologia Clínica, Hospital Fernando Fonseca, Amadora, Portugal*

Received 14 February 2007; received in revised form 12 May 2007; accepted 17 May 2007
Available online 21 May 2007

Abstract

Tuberculosis is a major health problem in Portugal. To begin characterizing the population structure of *Mycobacterium tuberculosis*, spoligotyping was used for the systematic typing, through consecutive sampling, of patient isolates from the Amadora-Sintra area of Greater Lisbon. Distribution amongst major spoligotype families, including the Latin American Mediterranean (LAM), T, Haarlem and Beijing, was compared to that of the international spoligotype database SpolDB4 and to the European countries of traditional Portuguese immigration represented in SpolDB4. Spoligotypes from 665 isolates were analyzed and 97 shared international types (SITs) identified. In SpolDB4 Portugal is represented by part of the spoligotypes from this study explaining the reduced number of unidentified patterns. The importance of the LAM family, and especially of LAM1 and LAM9 sub-families that alone represented 38% of all the isolates in this study as compared to 8% relative to the European sub group, led us to believe that at least in this respect the population structure was closer to that of Africa and South America than to Europe. Spoligotypes characteristic of Portugal or Portuguese related settings were identified. These included SIT244 a T1 sub-family predominant in Portugal and Bangladesh, SIT64 a LAM 6 sub-family common to Portugal and Brazil, and SIT1106 a LAM 9 sub-family. These studies were the first in Portugal stressing the importance of monitoring the population structure of *M. tuberculosis* isolates, an important step towards gaining an understanding of tuberculosis and the dynamics of this disease.

Keywords: *Mycobacterium tuberculosis*; Spoligotyping; Portugal

1. Introduction

Tuberculosis is a major health problem in Portugal. Recent estimates from the World Health Organization (WHO) (http://www.who.int/tb/country/datacollection/eur_country_list/en/index.html) show a high incidence of tuberculosis in Portugal, 42/100,000, as well as in countries of Portuguese language such as Brazil, 60/100,000, and former Portuguese African territories where tuberculosis ranges from 172/100,000 in Cab Verde to 460/100,000 in Mozambique. In the past, important aspects and weak points in the control of the disease's transmission have been revealed through the characterization of a genotypic cluster and an epidemiological study that followed (Portugal et al., 1999). However, although it is the basis for epidemiological analysis on which to establish measures for tuberculosis control, genotyping of *Mycobacterium tuberculosis* has not been the current procedure.

Technical advantages have led PCR-based techniques to be proposed as first and second line genotyping methods independently or in association with the standard DNA fingerprinting methodology based on IS6110 probe analysis of restriction fragment length polymorphisms (RFLP) of genomic DNA (van Soolingen et al., 1991; Filliol et al., 2000; Sola et al., 2003; Sun et al., 2004; Cowan et al., 2005).

One such a method, Spoligotyping is based on the analysis of the presence or absence of non-repetitive spacers in the *M.*

* Corresponding author. Present address: Centro de Tuberculose e Micobactérias, Instituto Nacional de Saúde Dr. Ricardo Jorge Delegação do Porto, Porto, Portugal. Tel.: +351 226074900; fax: +351 226053512.
E-mail address: sdavid@ibmc.up.pt (S. David).

1567-1348/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.05.007

tuberculosis DR locus of the chromosome (Kamerbeek et al., 1997). Spoligotyping is highly reproducible and useful to broadly study population structure and attempt to identify main lineages, however it must be combined with other typing methods to be useful for identification of transmission chains and discriminate isolates up to the clonal level (Cowan et al., 2005; Oelemann et al., 2007).

The DR locus has been shown to be highly polymorphic, revealing the extent of genetic variability in *M. tuberculosis*. Also, the organization of an international spoligotype database has made it possible to have a global view of the geographical distribution of genotypes (Sola et al., 2001; Filliol et al., 2002, 2003). A larger number of studies on *M. tuberculosis* population structure have permitted distinction between spoligotype families of high global distribution, such as the Beijing/W genotype, as compared to those with regional specificity (Bifani et al., 2002; Douglas et al., 2003; Niobe-Eyangoh et al., 2004; Garcia de Viedma et al., 2005; McHugh et al., 2005; Brudey et al., 2006).

Our work has been the first on the national level to begin characterizing the *M. tuberculosis* population structure through spoligotyping (David et al., 2004, 2006). It has led to the integration of Portuguese genotypes in the recently updated version of the international spoligotype database SpolDB4.

In this study spoligotyping was used for the systematic typing of isolates from tuberculosis patients of the Fernando Fonseca Hospital located in the Amadora-Sintra area of Greater Lisbon and the *M. tuberculosis* population structure was characterized with the help of SpolDB4. The hospital receives patients from communities distributed over a densely populated geographical area where tuberculosis is highly prevalent, namely amongst immigrant populations from previous Portuguese African territories, Asia, Brazil and Eastern European countries. As immigration is a relatively recent phenomenon in Portugal, this study may also contribute to the understanding of its impact on the genotypic diversity of *M. tuberculosis*.

2. Materials and methods

2.1. Strains

M. tuberculosis isolates from tuberculosis patients with positive culture hospitalized at the Fernando Fonseca Hospital located in the Amadora-Sintra area of Greater Lisbon were collected over a 6-year period from 1999 to 2005. The study collection consisted of 665 isolates corresponding to consecutive sampling of one isolate per patient. These were identified as *M. tuberculosis* using Accuprobe specific probes (GenProbe, San Diego, CA, USA). Cultures were stored on Lowenstein Jensen slants (BBL™, Becton Dickinson, Quilaban, Lisbon, Portugal).

2.2. Spoligotyping and spoligotype analysis

Spoligotyping was carried out according to the manufacturer's instructions (Isogen Bioscience B.V., Maarsen, Netherlands). The resulting spoligotypes were documented under a binary code representing either a positive or negative hybridization result (n and o) and analyzed using the Excel program for grouping and ordering of the patterns. Spoligotypes common to more than one strain were designated as "Shared Types" (ST) and attributed a "Shared International Type" number (SIT) according to the updated version of the international spoligotype database SpolDB4 available at http://www.pasteur-guadeloupe.fr:8081/SITVITdemo. They were further analyzed with the help of SpolDB4, allowing classification into spoligotype families and sub-families. Since SpolDB4 is a mixed database (Brudey et al., 2006), only *M. tuberculosis stricto sensu* and *M. africanum* families and sub-families were analyzed. *M. bovis*, *M. microti*, *M. caprae* and *M. pinipedii* strains were not detected in our sample, thus data in SpolDB4 regarding these members of the *M. tuberculosis* complex were not contemplated in percentile calculations. Likewise, as only one *M. bovis* BCG isolate was detected in our sample, *M. bovis* BCG was also removed from percentile calculations.

Distribution amongst spoligotype families and sub-families was compared to that of SpolDB4 as a whole or to that of European countries of traditional Portuguese immigration (EURp): Belgium, France, Germany, Great Britain, Ireland, Luxembourg, Netherlands, Spain and Switzerland. Since in SpolDB4 Portugal is represented exclusively by an entry of 336 spoligotypes from this study, this dataset was excluded from SpolDB4 so that the total number of isolates from our study sample could be considered as an external dataset to SpolDB4. Therefore, in the database EURp represented a total of 3759 isolates, independent from our study.

Prevalence was evaluated within each sample as the numbers of isolates with a common SIT relative to the total number of isolates from the same dataset (600 from the Fernando Fonseca Hospital or 3759 from the EURp) (expressed in percentile). However, in order to quantify the relative contribution of the Portuguese isolates to a specific SIT, an index of representativeness was defined as the number of isolates with a given SIT from our dataset to the number of isolates with the same SIT from the EURp or SpolDB4 datasets in addition to our own dataset (expressed in percentile).

3. Results

Of the 665 isolates analyzed, spoligotypes from 600 isolates (90%) were classified according to SpolDB4 into 97 Shared International Types (SITs) (Table 1). The remaining 62 patterns were not identified in the database, 59 were orphan (9%) and the remaining represented three clusters of two isolates each (1%) (Table 2). Two of these non-identified clustered STs were very similar to SIT1752 and SIT389, differing from these only by the absence of spacer 15. The third did not however resemble as closely any in the database. Eight SITs accommodated 50% of all the isolates (SIT20, SIT42, SIT53, SIT244, SIT47, SIT150, SIT64, SIT1).

The general features of the *M. tuberculosis* population structure in this study were close to those of the EURp (Table 3) in that the major spoligotype families were the Latin American and Mediterranean (LAM), Haarlem and T; however, differences in the proportions of these major families existed between the two settings. In the Amadora-Sintra area of Greater

Table 1
Spoligotypes identified in SpolDB4 by a SIT number[a]

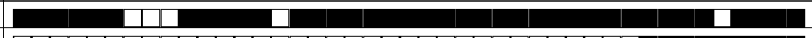
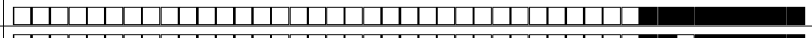
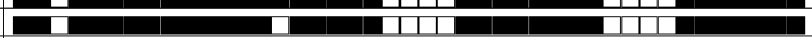
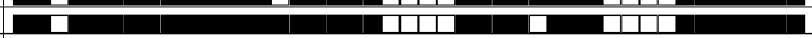
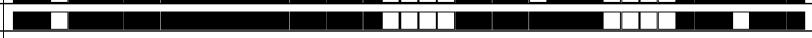
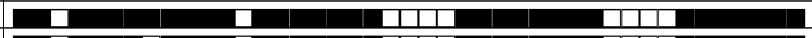
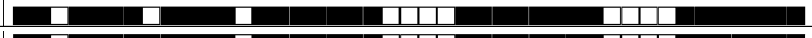
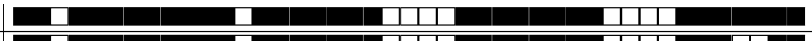
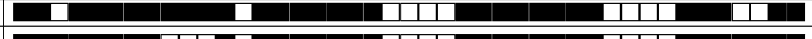
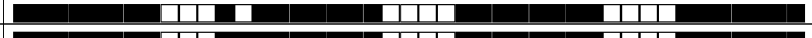
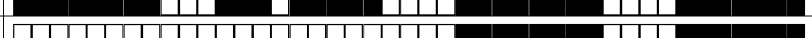
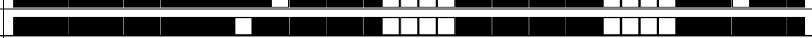
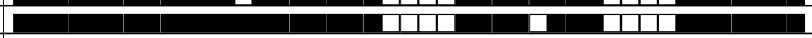
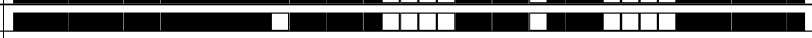
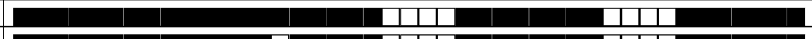
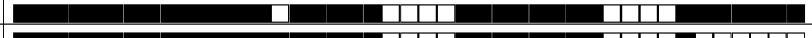
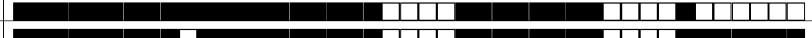
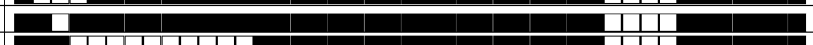
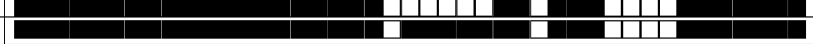
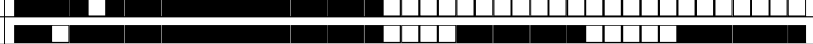
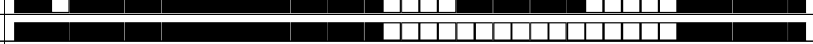
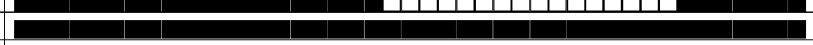
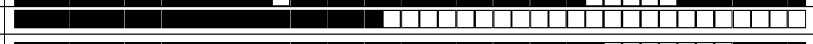
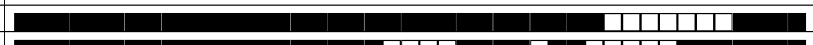
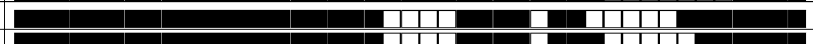
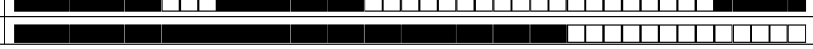
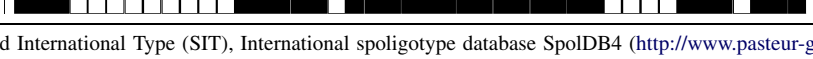
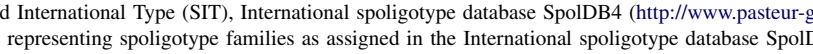
SIT[a]	Spoligotype	Label[b]	Nº of Isolates[c]	Prevalence[d] (%)
181		AFRI1	3	0,50
324		AFRI1	1	0,17
1		BEIJING	17	2,83
265		BEIJING	1	0,17
26		CAS1	2	0,33
1251		EAI1_SOM	1	0,17
129		EAI5	1	0,17
47		H1	27	4,50
151		H1	2	0,33
283		H1	2	0,33
62		H1	2	0,33
45		H1	1	0,17
50		H3	16	2,67
49		H3	4	0,67
448		H3	1	0,17
75		H3	1	0,17
740		H3	1	0,17
746		H3	1	0,17
293		H3	1	0,17
451		H37Rv	1	0,17
262		H4	1	0,17
20		LAM1	97	16,17
389		LAM1	15	2,50
1755		LAM1	2	0,33
1321		LAM1+LAM4	3	0,50
17		LAM2	10	1,67
545		LAM2	1	0,17
193		LAM2	2	0,33
179		LAM2	1	0,17
211		LAM3	9	1,50
33		LAM3	7	1,17
1760		LAM3	1	0,17
111		LAM3	1	0,17
4		LAM3+S convergent	4	0,67
60		LAM4	9	1,50
1750		LAM4	3	0,50
93		LAM5	2	0,33
64		LAM6	18	3,00
1066		LAM6	5	0,83
1759		LAM8	4	0,67
42		LAM9	69	11,50
150		LAM9	22	3,67
1106		LAM9	10	1,67
81		LAM9	5	0,83
1064		LAM9	2	0,33
1249		LAM9	1	0,17
177		LAM9	1	0,17
388		LAM9	1	0,17
61		LAM10	1	0,17
115		LAM10_CAM	1	0,17
59		LAM11_ZWE	1	0,17
34		S	15	2,50
53		T1	52	8,67
244		T1	32	5,33
51		T1	5	0,83
344		T1	4	0,67
154		T1	2	0,33
732		T1	2	0,33
173		T1	2	0,33
635		T1	1	0,17

Table 1 (*Continued*)

798		T1	1	0,17
196		T1	1	0,17
1753		T1	1	0,17
144		T1	1	0,17
156		T1	1	0,17
766		T1	1	0,17
1105		T1	1	0,17
306		T1	1	0,17
453		T1	1	0,17
373		T1	1	0,17
462		T1	1	0,17
521		T1	1	0,17
118		T2	7	1,17
784		T2+S	1	0,17
73		T2+T3	7	1,17
157		T3	4	0,67
37		T3	1	0,17
44		T5	4	0,67
58		T5_MAD2	4	0,67
1757		U	8	1,20
1752		U	6	1,00
397		U	4	0,67
523		U	3	0,50
834		U	2	0,33
402		U	2	0,33
787		U	2	0,33
1758		U	1	0,17
396		U	1	0,17
161		U	1	0,17
519		U	1	0,17
1196		U	1	0,17
106		U (LAM3?)	1	0,17
237		U (likely H3)	1	0,17
29		U likely LAM	2	0,33
92		X3	11	1,83
1756		X3	2	0,33
1273		X3	1	0,17

[a]Shared International Type (SIT), International spoligotype database SpolDB4 (http://www.pasteur-guadeloupe.fr:8081/SITVITdemo/).
[b]Lable representing spoligotype families as assigned in the International spoligotype database SpolDB4.
[c]Number of isolates in this study.
[d]Prevalence, representing the number of isolates with a common SIT relative to the total number of isolates from the same dataset (600 isolates classified by SIT from the Fernando Fonseca Hospital) (expressed in percentile).

Lisbon, the LAM family was by far more prevalent (51%) than in the EURp (16%), with a difference of 35%. On the other hand and to a lesser extent, the T and Haarlem families were more prevalent in EURp (30 and 18%, respectively) than in the Portuguese sample under study (23 and 10%, respectively). Finally, the Beijing family was in a similar proportion in either setting (3%).

The prevalence of the most relevant sub-families, accommodating approximately 77% of all the isolates from this study, in comparison to those in the EURp, representing roughly 57% of these isolates in SpolDB4, are shown in Fig. 1. The importance of the LAM family in our sample may be explained by the high prevalence of the LAM1 and LAM9 sub-families. These two sub-families represented 38% of all the isolates in the sample group studied as compared to 8% in the EURp. The presence of LAM6 may also contribute to the disparity, as it was encountered in our sample where it represents nearly 4% of the isolates, but rarely so in the EURp sample (less than 1%). Relative to the T family, the most prevalent sub-families in both contexts were T1 and T2, with 20 and 19% for T1 in the European and the Lisbon samples, respectively, and likewise 5 and 3% for T2. Haarlem1 had the same representation in EURp as in Lisbon (6%), whereas on the contrary, the Haarlem3 sub-family presented higher prevalence in EURp than in the sample under study (11 and 4%, respectively). The X family was rare in our study. X1 and X2 were not detected, yet showed some importance in the EURp population structure (2–3%).

Further distinction was possible through analysis of the most representative spoligotype families and sub-families by SIT, as shown in Tables 4 and 5. As regards the LAM family, in the samples from this study LAM1 was represented by a tight structure of only three genotypes: SIT20, SIT389 and SIT1755. In the database, Portugal is highly representative of these SITs within the EURp sample and SpolDB4 as a whole. Contrary to LAM1, LAM9 displayed a looser structure with a larger number of genotypes. Of these two were also important in the

Table 2
Spoligotypes of orphan strains and clusters of spoligotypes not identified in SpolDB4 by a SIT number[a]

Nº of Isolates[b]	Spoligotype
1	□□□□□□■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	□□□□□□■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■□□■■
1	□□■■■■■■□□□■□■■■■■□■□□□□■■■■■■■■□□□□■■■■■■■
1	□□■■■■■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	□□■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■□□□□■
1	■□□■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■□□□
1	■■□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■■■
1	■■□□■■■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	■■□■□□■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	■■□■■■□■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	■■□■■■■□■■□■■■□■■■■■□□□□■■■■■■■■□□□□■■□■■■■
1	■■□■■■■■■□□■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	■■□■■■■■■■□■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	■■□■■■■■■■■□■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■
2	■■□■■■■■■■■■■■□■■■■■□□□□■■■■■■■□□□□□■■■■■■■
1	■■□■■■■■■■■■■■■■■■■■□□□□■□□■■■■■□□□□■■■■■■■
1	■■□■■■■■■■■■■■■■■■■■□□□□■■□□□□■■□□□□■■■■■■■
1	■■□■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■□□□□■
1	■■■□□□□□□□□□■■□■■□■■■■■■■■■■■■■■□□□□■■■□■■■
1	■■■□□□□□□□■■■■■■■■■■■■□□□□□□□□□□□□□□■■■■■■□■■
1	■■■□□□□■■■■■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□□□
1	■■■□□□■■■■■■■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□□
1	■■■■■□□□□□□□□□□□□■■■□□□□■■■■■■□■□□□□□□■■■■■
1	■■■■■□□□□■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■□□■■
1	■■■■■□□■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■□■■■
1	■■■■■■□□□■■■■■□■■■■□□■■■■■■■■■■■■■■■■■■■■■■
1	■■■■■■□□□■■■■■■■■■■■■■■■■■■□□□□□□□□□□□□□□□□
1	■■■■■■□□■■■□■■■■■■■■□□□□■■■■□■■■□□□□■■■■■■■
1	■■■■■■□□■■■■■■■■■■■■□□□□■■■■□■■■□□□□■■■■■■■
1	■■■■■■□■□□□■■■■■■■■■□□□□■■□■□■■■□□□□■■■■■■■
1	■■■■■■□■■■□■■■□■■■■■□■■■■□■■□■■■□□□□■■□■■■■
1	■■■■■■■□■■■■■■□□□■□□□□□□□□□□□□□□□□□□□□□□■■■
1	■■■■■■■■□□□□■■□■■■■■□□□□□□□□□□□□□□□□■■■■■■■
1	■■■■■■■■□□□■□■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■
1	■■■■■■■■□□□■■■□■■■■■■■■■■■■■■■■■□□□□■■■■■■■
1	■■■■■■■■□□□■■■■■■■■■□□□□□□□□□□□□□□□□□■■■■■■
1	■■■■■■■■□□□■■■■■■■■■□□□□■■■■■■■■□□■■■■■■■■■
1	■■■■■■■■□□■■■■■■■■■□□□□□□□□□□□□□□□□□□□■■■■■
1	■■■■■■■■■□■■■■■■■■■■■■■■■■□□■■■■□□□□■■□□□□■
1	■■■■■■■■■■□■■■□■■■■■□□□□■■■■■■■■□□□□■■■■■■■
1	■■■■■■■■■■■□□□□□■□□□□□□□■■■■□□□□■□■■□■■■■■■
1	■■■■■■■■■■■■□■□□■■■■■■■■■■■□■■■■□□□□■■■■■■■
1	■■■■■■■■■■■■□■□■■■■■■■■■■■■■■■■■□□□□□□□■■■■
1	■■■■■■■■■■■■□■■■■■■■□□□□■■■■■■■■□□□□■■■■■□□
1	■■■■■■■■■■■■□■■■■■■■■■■□■□□□□■□■□□□□■■■■■■■
1	■■■■■■■■■■■■■■□■■■■□■■□■■■□□■■■■□□□□■■■■■■■
1	■■■■■■■■■■■■■■□■■■■■□□□□□□□□□□□□□□□□□□□□□□□
1	■■■■■■■■■■■■■■□■■■■■□□□□■■□□□□■□□□□□■■■■■■■
1	■■■■■■■■■■■■■■□■■■■■□□□□■■■■■□■■□□□□■■■■■■■
1	■■■■■■■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■□□□□□□
1	■■■■■■■■■■■■■■□■■■■■□□□□■■■■■■■■□□□□■□□□□■■
1	■■■■■■■■■■■■■■□■■■■■■■■■■■□□□■■■□□□□■■□□□□■
1	■■■■■■■■■■■■■■□■■■■■■■■■■■■■■■■■□□□□□□■■■■■
1	■■■■■■■■■■■■■■□■■■■■■■■■■■■■■■■■□□□□■■□□□■■
2	■■■■■■■■■■■■■■□■■■■■■■■■■■■■■■■■□□□□■■■□□□□
1	■■■■■■■■■■■■■■■■■■■□□□□□□□□■■■■■□□□□■■■□■■■
1	■■■■■■■■■■■■■■■■■■■■□□□□■□□■■■■■□□□□■■■■■■■
1	■■■■■■■■■■■■■■■■■■■■□□□□■■□□□□□■□□□□■■■■■■■
1	■■■■■■■■■■■■■■■■■■■■□□■■■■■■■■□■□□□□■■■■□■■
1	■■■■■■■■■■■■■■■■■■■■□□■■■■■■■■■■□□□□■■■■□■■
1	■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■□□□□■■■■■■■
2	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□■■■□□□□■■□□□□■

[a]Shared International Type (SIT), International spoligotype database SpolDB4 (http://www.pasteur-guadeloupe.fr:8081/SITVITdemo/).
[b]Number of isolates in this study.

Table 3
Proportion of major families of isolates of *M. tuberculosis* in the European countries of traditional Portuguese immigration (EURp) and from the present study in the Amadora-Sintra area of Greater Lisbon

Genotypic family[a]	Prevalence (%)[b]	
	EURp	This study
LAM	16	51
T	30	23
Haarlem	18	10
Beijing	3	3
Other	33	13

[a] Spoligotype families Latin American Mediterranean (LAM), T, Haarlem and Beijing as assigned in the International spoligotype database SpolDB4 (http://www.pasteur-guadeloupe.fr:8081/SITVITdemo/).

[b] All isolates belonged to families and sub-families of *M. tuberculosis stricto sensu* and *M. africanum*. Although *M. bovis*, *M. microti*; *M. caprae* and *M. pinipedii* integrate SpolDB4 these were not found in our sample and were thus not contemplated in percentile calculations. Prevalence, represented the number of isolates with a common SIT relative to the total number of isolates from the same dataset which represented 3759 isolates from SpolDB4 for EURp and 600 isolates classified by SIT from the present study in the Lisbon area (expressed in percentiles).

EURp (SIT42 and SIT150). On the other hand, Portugal was the only representative for SIT1106 within the EURp. Some of the genotypes with a limited number of isolates were found seemingly common to South America or Africa according to SpolDB4 (SIT81, SIT177, SIT388). In the case of SIT1249, confirmation of origin from Eastern Europe was obtained. As for the LAM6 sub-family, SIT64 was found to be important in Brazil as well as in our sample from Lisbon. The Brazilian origin was indeed established for one isolate. SIT 1066 found in our study also appeared in South America and Africa. In the Lisbon sample, the T1 sub-family SIT244 was distinguishable due the high index of representativeness of this genotype, 84 and 53% in the EURp and SpolDB4, respectively. Amongst the genotypes in reduced numbers, two were found with possible Brazilian and African origins, SIT306 and SIT144, respectively. In Table 4, the Haarlem families are represented by two major SITs, SIT47 for Haarlem1 and SIT50 for Haarlem3. From analysis of SpolDB4, these did not appear to be particularly representative of our sample. Also, foreign origin was not confirmed for any of these isolates.

Analysis of origin was simplified when specific designation was available from other studies as was the case with a number of genotypes labeled accordingly in SpolDB4 (Brudey et al., 2006), such as AFRI_1, in the case of SIT 181 and SIT 324, CAS_DEHLI for SIT 26, EAI1_SOM for SIT 1251, LAM10_CAM for SIT 61 and SIT 115, LAM11_ZWE for SIT 59, T5_MAD2 for SIT 58 and Beijing for SIT 1 and SIT 265. In our study, these genotypes represented 32 isolates or 5% of the total number of isolates. Of these 18 isolates, 3% were Beijing, with SIT1 as the major Beijing genotype. As of yet only 3 isolates were attributed a probable Eastern European or Asian origin. The first Beijing isolate resistant to all five first line antibiotics was detected in 2005 through this study, presumably of Eastern European origin. The other genotypes including from the AFRI, CAS, S and EAI (East Asian Indian) families were very poorly represented in the Lisbon sample. The four isolates belonging to the AFRI1 genotypes, were presumed original of Guinea Bissau. The only CAS genotype identified was SIT 26 a CAS_DEHLI genotype and there were no data on the origin of the strains.

Presumably, Portuguese related *M. tuberculosis* genotypes are shown in Table 5. Relative to the number of strains and the high degree of contribution of the isolates from this study to SpolDB4 genotype, and based on the analysis of global distribution using the database, a number of these genotypes appeared to be Portuguese specific genotypes, as with SIT 244 and SIT 1106, but also SIT 1752, SIT 1757 and SIT 1759. Other genotypes appeared Portuguese and Brazilian related, as with SIT 64, or more generally speaking were highly representative genotypes with possible linkage to the natural history of tuberculosis in Portugal (Table 5). These included LAM1 genotypes (SIT 20 and SIT 389), LAM6 (SIT 64), LAM9 (SIT 1106) and T1 (SIT 244) genotypes as mentioned, but also LAM8 (SIT 1759), T3 (SIT 157) and Undefined (SIT 1752 and SIT 1757) genotypes.

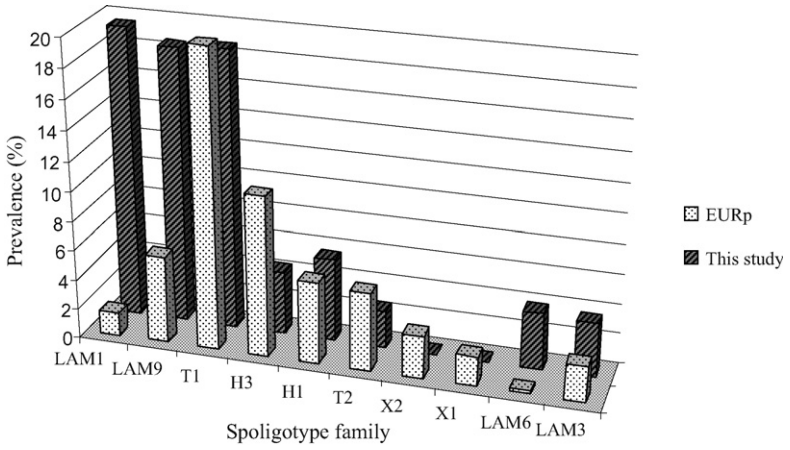


Fig. 1. Prevalence of the major spoligotype families and sub-families of *M. tuberculosis* isolates from the Amadora-Sintra area of Greater Lisbon, accommodating approximately 77% of all the isolates in the present study, in comparison to those in the European countries of traditional Portuguese immigration, representing roughly 57% of these isolates in SpolDB4 (http://www.pasteur-guadeloupe.fr:8081/SITVITdemo/).

Table 4
Analysis by SIT of the most representative spoligotype families and sub-families of *M. tuberculosis* isolates from this study in the Amadora-Sintra area of Greater Lisbon[a]

Spoligotype family	SIT	No. isolates in this study	Prevalence[b] (%)
LAM1	20	97	16.17
	389	15	2.50
	1755	2	0.33
LAM6	64	18	3.00
	1066	5	0.83
LAM9	42	69	11.50
	150	22	3.67
	1106	10	1.67
	81	5	0.83
	1064	2	0.33
	177	1	0.17
	388	1	0.17
	1249	1	0.17
T1	53	52	8.67
	244	32	5.33
	51	5	0.83
	344	4	0.67
	154	2	0.33
	173	2	0.33
	732	2	0.33
	144	1	0.17
	156	1	0.17
	196	1	0.17
	306	1	0.17
	373	1	0.17
	462	1	0.17
	521	1	0.17
	635	1	0.17
	766	1	0.17
	798	1	0.17
	1105	1	0.17
	1753	1	0.17
H1	47	27	4.50
	62	2	0.33
	151	2	0.33
	283	2	0.33
	45	1	0.17
H3	50	16	2.67
	49	4	0.67
	75	1	0.17
	293	1	0.17
	448	1	0.17
	740	1	0.17
	746	1	0.17

[a] Spoligotype families Latin American Mediterranean (LAM), T and Haarlem (H), were labeled according to the International spoligotype database SpolDB4 (http://www.pasteur-guadeloupe.fr:8081/SITVITdemo/).

[b] Prevalence, representing the number of isolates with a common SIT relative to the total number of isolates from the same dataset (600 isolates classified by SIT from the Fernando Fonseca Hospital) (expressed in percentile).

4. Discussion

Portugal is closely related to many countries by the traditional emigration of Portuguese populations. However, more recently, it has also received immigrants mostly from previous Portuguese African territories, Asia, Brazil and Eastern European countries. The Fernando Fonseca Hospital, where this study was undertaken, is situated in the Amadora-Sintra area of Greater Lisbon. This hospital receives patients from a densely populated geographical area where tuberculosis is highly prevalent and which also harbors important nuclei of foreign communities.

The metropolitan area of Lisbon receives roughly 55% of Portugal's immigrant population mostly from previous Portuguese African territories, Asia, Brazil and Eastern European countries. Immigrants from African countries of Portuguese language represent close to 44% of the total number of immigrants in Portugal, however in the communities served by the hospital they represent a high 76% of the total (Centro de Estudos para a Intervenção Social. Diagnóstico Social do Conselho da Amadora–Programa Rede Social, 2004 (http://www.cm-amadora.pt/web/_pdf/as00ca.pdf; Instituto Nacional de Estatística (INE), Portugal–Censos 2001).

Tuberculosis amongst the foreign community is high. In some areas of Greater Lisbon area it may represent close to 30% of the total number of cases, over the double reported in the rest of the country (A. Fonseca Antunes. Direcção-Geral da Saúde, Programa Nacional de Luta Contra a Tuberculose: Ponto da Situação Epidemiológica e de Desempenho, Ano 2003. Fevereiro 2004. http://www.dgsaude.pt/upload/membro.id/ficheiros/i006111.pdf; A. Fonseca Antunes. Direcção-Geral da Saúde, Tratamento de Dados do Sistema de Vigilância da Tuberculose (SVIG-TB), 2000–2005 unpublished).

To begin characterizing the population structure of the *M. tuberculosis* isolates, spoligotyping was used for the systematic typing through consecutive sampling of isolates from tuberculosis patients of the Fernando Fonseca Hospital from 1999 to 2005. These studies have recently led to the integration of Portuguese genotypes in SpolDB4. In SpolDB4 Portugal is represented by an important part of the isolates from this study (336 isolates), explaining the reduced number of patterns for which a SIT number could not be attributed (62 patterns).

In this study the distribution of the 665 isolates from patients of the Fernando Fonseca Hospital amongst major spoligotype families, including the Latin American Mediterranean (LAM), T, Haarlem and Beijing, was compared to that of the International spoligotype database SpolDB4 as a whole and more specifically to European countries of traditional Portuguese immigration (EURp): Belgium, France, Germany, Great Britain, Ireland, Luxembourg, Netherlands, Portugal, Spain and Switzerland.

Although *M. tuberculosis* presented a tight population structure with eight SITs accommodating 50% of all the isolates (SIT20, SIT42, SIT53, SIT244, SIT47, SIT150, SIT64, SIT1), our analysis also showed a high degree of diversity with 97 SITs accommodating 600 of the isolates (90%). The remaining 65 isolates were represented by 62 spoligotype patterns not identified in the database.

The major spoligotype families, Latin American and Mediterranean group (LAM), T, Haarlem and Beijing being conserved in Portugal relative to EURp indicated an overall similar population structure for *M. tuberculosis*. However,

Table 5
Portuguese related *Mycobacterium tuberculosis* genotypes

Spoligotype family[a]	SIT[a]	No. of isolates			Index of representativeness relative to	
		Present study	EURp	SpolDB4	EURp[b] (%)	SpolDB4[c] (%)
LAM1	20	97	55	377	64	21
T1	244	32	6	28	84	53
LAM6	64	18	6	143	75	11
LAM1	389	15	0	7	100	68
LAM9	1106	10	0	2	100	83
U	1757	8	0	0	100	100
U	1752	6	0	1	100	86
LAM8	1759	4	0	0	100	100
T3	157	4	1	3	80	57

[a] Spoligotype families, including Latin American Mediterranean (LAM), T and Undefined (U), and Shared International Type (SIT) numbers, were assigned as in the International spoligotype database SpolDB4 (http://www.pasteur-guadeloupe.fr:8081/SITVITdemo/).
[b] Number of isolates with a given SIT from our dataset relative to the number of isolates with the same SIT from the European countries of traditional Portuguese immigration (EURp) in addition to our own dataset.
[c] Number of isolates with a given SIT from our dataset relative to the number of isolates with the same SIT from SpolDB4 in addition to our own dataset, expressed in percentile.

important differences between the two settings were apparent from the proportions of each family. In this respect, the LAM had a much higher representation in Portugal whereas T and Haarlem were somewhat more common in EURp.

The LAM family, represented by three important sub-families LAM1 (19%), LAM6 (4%) and LAM9 (19%), was the most highly represented in the Portuguese panorama with 308 isolates. Further analysis revealed the prevalence of SITs where the number of strains from Portugal contributed in an important manner to these SITs relative the EURp sample or to SpolDB4 as a whole (SIT 20, SIT 389 and SIT 1755 from LAM1; SIT 64 from LAM6; SIT 1106 from LAM9). The LAM6 SIT 64 was found to be important in Brazil as well as in our sample from Lisbon.

The high prevalence of the LAM family in our sample is a feature also seen amongst *M. tuberculosis* isolates in South America and some African countries represented in the database. This lead us to believe that at least in this respect the population structure from our study in Portugal was closer to that of Africa and South America than to Europe.

Through the analysis of the more representative genotypes, their contribution to the genotype in number of isolates and their geographical distribution according to SpolDB4, it was also possible to identify seemingly Portuguese related genotypes. These included LAM1 (SIT 20 and SIT 389), LAM6 (SIT 64), LAM9 (SIT 1106) and T1 (SIT 244), LAM8 (SIT 1759), T3 (SIT 157) and Undefined (SIT 1752 and SIT 1757) genotypes, of which SIT 244 and SIT 1106, but also SIT 1752, SIT 1757 and SIT 1759, appeared to be Portuguese specific. It is relevant to note that the LAM1 genotype SIT244 was important in only one other global setting, Bangladesh (Z. Rahim, K. Zaman, A.G.M. van der Zanden, M.J. Mullers, D. van Soolingen, N. Rastogi, and C. Sola, Abstr. 27th Annual Congress of the European Society of Mycobacteriology, abstr. 09, 2006). Also, Portugal was the major representative of the LAM9 genotype SIT 1106, whereas SIT 64 was found to be important in Brazil as well as in our sample from Lisbon.

Clinical information on a limited number of isolates confirmed imported origin; however, this data was not available for the large majority of cases. In tuberculosis incubation periods are expectedly long, it is therefore important that physicians realize the relevance in reporting patient history as a measure for disease control.

Through this study the first Beijing isolate resistant to all five first line antibiotics was detected in 2005, presumably of Eastern European origin. This is a matter of concern considering the rapid global spread and suspected association of these genotypes with high virulence (Bifani et al., 2002; Drobniewski et al., 2005).

This study was the first in Portugal stressing the importance of monitoring the population structure of *M. tuberculosis* isolates. It represents a first baseline study of the *M. tuberculosis* population structure, which is useful to guide future epidemiological studies in the country and to enlarge the picture on the global pathogen's distribution. It revealed a gamut in genotypes that went from high global distribution to restricted geographical distribution in spite of the centuries of mobility of its human hosts.

We now intend to extend the characterization of the population structure of *M. tuberculosis* strains to other regions of Portugal and to proceed with regional and international comparisons. Standard molecular genotyping methods such as spoligotyping and MIRU-VNTR genotyping as well as others will permit discerning whether clustered strains are emerging epidemic strains or strains for which further epidemiological and phylogenetic reconstruction should be undertaken.

With robust and highly reproducible genotyping tools it is now also possible from a global perspective to begin comparison of the *M. tuberculosis* population structure between historically or culturally related settings, an important step towards gaining an understanding of tuberculosis and the dynamics of this disease.

Acknowledgements

We are grateful to Dr. Nalin Rastogi and Dr. Christophe Sola (Unit of Tuberculosis and Mycobacteria, Pasteur Institute of Guadeloupe) and to Prof. Hugo L. David (Honorary Professor

of the Pasteur Institute) for fruitful discussion and help in preparing this manuscript. This investigation has been partially financed by grants from the Fundação Calouste Gulbenkian and from the Fundação para a Ciência e Tecnologia (FCT), Portugal (grant reference: POCTI/ESP/44826/2002), coparticipated by the European Community fund FEDER. D. Rapose Ribeiro's contribution was in the context of preparation for her degree in Pharmacy, Licenciatura Bietápica em Farmácia, Escola Superior de Tecnologia da Saúde de Lisboa.

References

Bifani, P.J., Mathema, B., Kurepina, N.E., Kreiswirth, B.N., 2002. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. Trends Microbiol. 10, 45–52.

Brudey, K., Driscoll, J.R., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajoj, S.A., Allix, C., Aristimuno, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J.T., Fauville-Dufaux, M., Ferdinand, S., Garcia de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H.M., Guttierez, M.C., Hawkey, P.M., van Helden, P.D., Kadival, G.V., Kreiswirth, B.N., Kremer, K., Kubin, M., Kulkarni, S.P., Liens, B., Lillebaek, T., Ho, M.L., Martin, C., Martin, C., Mokrousov, I., Narvskaia, O., Ngeow, Y.F., Neumann, L., Niemann, S., Parwati, I., Rahim, Z., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M.L., Rusch-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I.G., Singh, U.B., Somoskovi, A., Skuce, R.A., van Soolingen, D., Streicher, E.M., Suffys, P.N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T.C., Warren, R.M., Yap, S.F., Zaman, K., Portaels, F., Rastogi, N., Sola, C., 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol. 6, 23–39.

Cowan, L.S., Diem, L., Monson, T., Wand, P., Temporado, D., Oemig, T.V., Crawford, J.T., 2005. Evaluation of a two-step approach for large-scale, prospective genotyping of *Mycobacterium tuberculosis* isolates in the United States. J. Clin. Microbiol. 43, 688–695.

David, S., Barros, V., Portugal, C., Antunes, A., Cardoso, A., Calado, A., Sancho, L., Germano de Sousa, J., 2006. Update on the Spoligotypes of *Mycobacterium tuberculosis* complex isolates from the Fernando Fonseca Hospital (Amadora-Sintra Portugal). Rev. Port. Pneumologia XI (6), 513–531.

David, S., Portugal, C., Antunes, A., Cardoso, A., Calado, A., Barros, V., Sancho, L., 2004. Molecular Identification using Spoligotyping of strains from the *Mycobacterium tuberculosis* complex isolated from the Hospital Fernando Fonseca. Rev. Port. Pneumologia X (3), 195–204.

Douglas, J.T., Qian, L., Montoya, J.C., Musser, J.M., Van Embden, J.D., Van Soolingen, D., Kremer, K., 2003. Characterization of the Manila family of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 41, 2723–2726.

Drobniewski, F., Balabanova, Y., Nikolayevsky, V., Ruddy, M., Kuznetzov, S., Zakharova, S., Melentyev, A., Fedorin, I., 2005. Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia. JAMA 293, 2726–2731.

Filliol, I., Driscoll, J.R., Van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Anh, D.D., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniewski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kallenius, G., Kassa-Kelembho, E., Koivula, T., Ly, H.M., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolofo-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., De Waard, J.H., Sola, C., Rastogi, N., 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. Emerg. Infect. Dis. 8, 1347–1349.

Filliol, I., Driscoll, J.R., van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Dang, D.A., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniewski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kassa-Kelembho, E., Ho, M.L., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolofo-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., de Waard, J.H., Sola, C., Rastogi, N., 2003. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. J. Clin. Microbiol. 41, 1963–1970.

Filliol, I., Ferdinand, S., Negroni, L., Sola, C., Rastogi, N., 2000. Molecular typing of *Mycobacterium tuberculosis* based on variable number of tandem DNA repeats used alone and in association with spoligotyping. J. Clin. Microbiol. 38, 2520–2524.

Garcia de Viedma, D., Bouza, E., Rastogi, N., Sola, C., 2005. Analysis of *Mycobacterium tuberculosis* genotypes in Madrid and identification of two new families specific to Spain-related settings. J. Clin. Microbiol. 43, 1797–1806.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., van Embden, J., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. 35, 907–914.

McHugh, T.D., Batt, S.L., Shorten, R.J., Gosling, R.D., Uiso, L., Gillespie, S.H., 2005. *Mycobacterium tuberculosis* family: a naming of the parts. Tuberculosis (Edinb). 85, 127–136.

Niobe-Eyangoh, S.N., Kuaban, C., Sorlin, P., Thonnon, J., Vincent, V., Gutierrez, M.C., 2004. Molecular characteristics of strains of the cameroon family, the major group of *Mycobacterium tuberculosis* in a country with a high prevalence of tuberculosis. J. Clin. Microbiol. 42, 5029–5035.

Oelemann, M.C., Diel, R., Vatin, V., Haas, W., Rusch-Gerdes, S., Locht, C., Niemann, S., Supply, P., 2007. Assessment of an optimized mycobacterial interspersed repetitive- unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis. J. Clin. Microbiol. 45, 691–697.

Portugal, I., Covas, M.J., Brum, L., Viveiros, M., Ferrinho, P., Moniz Pereira, J., David, H., 1999. Outbreak of multiple drug resistant tuberculosis in Lisbon: detection by restriction fragment length polymorphism analysis. Int. J. Tuberc. Lung Dis. 3, 207–213.

Sola, C., Filliol, I., Gutierrez, M.C., Mokrousov, I., Vincent, V., Rastogi, N., 2001. Spoligotype database of *Mycobacterium tuberculosis*: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives. Emerg. Infect. Dis. 7, 390–396.

Sola, C., Filliol, I., Legrand, E., Lesjean, S., Locht, C., Supply, P., Rastogi, N., 2003. Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. Infect. Genet. Evol. 3, 125–133.

Sun, Y.J., Lee, A.S., Ng, S.T., Ravindran, S., Kremer, K., Bellamy, R., Wong, S.Y., van Soolingen, D., Supply, P., Paton, N.I., 2004. Characterization of ancestral *Mycobacterium tuberculosis* by multiple genetic markers and proposal of genotyping strategy. J. Clin. Microbiol. 42, 5058–5064.

van Soolingen, D., Hermans, P.W., de Haas, P.E., Soll, D.R., van Embden, J.D., 1991. 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. J. Clin. Microbiol. 29, 2578–2586.