JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2005, p. 1797–1806 Vol. 43, No. 4
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.4.1797–1806.2005

Analysis of *Mycobacterium tuberculosis* Genotypes in Madrid and Identification of Two New Families Specific to Spain-Related Settings

Darío García de Viedma,[1]* Emilio Bouza,[1] Nalin Rastogi,[2] and Christophe Sola[2]

Department of Clinical Microbiology and Infectious Diseases, Gregorio Marañón Hospital, Madrid, Spain,[1] and Unité de la Tuberculose et des Mycobactéries, Institut Pasteur de Guadeloupe, Pointe-à-Pitre, Guadeloupe[2]

Received 2 July 2004/Returned for modification 21 September 2004/Accepted 25 October 2004

In Spain, tuberculosis (TB) patterns are changing because of the recent increase in the number of cases among immigrants. To establish the composition of circulating *Mycobacterium tuberculosis* strains before the effects of foreign strains appear, this study focused on molecular characterization of 233 patient isolates using spoligotyping. The spoligotyping data were further analyzed using an international database, SpolDB4. The results obtained showed that the general features of the *M. tuberculosis* population in Spain are coherent with those of other European countries, with the Latin American and Mediterranean group, and with the Haarlem 3 and T1 families as the most prevalent genotypes. The Spanish isolates clustered mostly with genotypes which had previously been isolated in countries linked with Spain. We also describe and fully characterize two novel *M. tuberculosis* families, Madrid1 and Madrid2, which are specific to Spain-related settings. The data reported here provide a solid reference when monitoring changes in the composition of the *M. tuberculosis* population in Spain as a consequence of the increasing rate of TB in the foreign population.

The potential role of international social movements in modifying the patterns and transmission dynamics of tuberculosis (TB) has been studied in different countries (2, 8, 12). In Spain, the increase in the number of immigrants is still a recent phenomenon, which has become marked only in the last few years (10). Therefore, we thought it could be useful to obtain a baseline reference for the clonal composition of the circulating *Mycobacterium tuberculosis* strains in Madrid to assess the genetic structure of the global TB bacillus population in Madrid at a time when the effects of immigration were still quite moderate.

Traditionally, tuberculosis transmission could be understood only through contact tracing, using the classical methods of conventional epidemiology and the "stone-in-the-pond" principle (26). The advent of molecular epidemiology has shed light on the recent transmission rates of tuberculosis in the community, with the finding of higher transmission rates than suspected (25). The definition of "clusters," groups of strains showing identical genetic characteristics, and their application to the assessment of recent tuberculosis transmission dynamics are the subjects of intense research. Furthermore, clusters are now widely accepted as representing phylogenetically significant information on the population structure of tubercle bacilli (and on their history) in a given setting (5). Indeed, the prevalences of different clones of *M. tuberculosis* vary from one region to another, leading to the interest in analyzing the worldwide population structure of tubercle bacilli (18).

The creation of international databases has revealed the clonal structure of *M. tuberculosis* populations in different geographic settings and has also defined superfamilies that are specific to certain countries (4, 14, 15). In this regard, the existence of an international database on spoligotyping (for spacer oligonucleotide typing), a PCR-based method that relies on a genetic locus called the direct repeat (11), which is highly polymorphic worldwide, has allowed some of the major superfamilies of *M. tuberculosis* to be described (6, 7).

The objective of this study was to describe the clonal composition of *M. tuberculosis* in Spain (in the Madrid area) in order to provide a reference which could be used to monitor potential changes in the genetic structure of the global population of the Spanish *M. tuberculosis* isolates. Indeed, the importation of TB cases to Spain through foreign-born patients from countries with a high TB burden may considerably affect circulating *M. tuberculosis* clones in the near future. Consequently, we decided to genotype 233 clinical isolates recruited during a 2-year period and to compare the results with those of an international database, SpolDB4, that previously contained some, although few, genetic data on *M. tuberculosis* clinical isolates in Spain.

MATERIALS AND METHODS

Strains. All the *M. tuberculosis* strains were isolated at the Department of Clinical Microbiology and Infectious Diseases of the Gregorio Marañón Hospital in Madrid during a 2-year period (from January 2001 to December 2002). *M. tuberculosis* isolates were identified by using Accuprobe specific probes (Gene Probe, San Diego, Calif.). Only one *M. tuberculosis* strain per patient was selected for study. Madrid is divided into 11 health service areas, and we studied all the *M. tuberculosis* strains cultured from all the tuberculosis cases in area I, which is served by our hospital. This area is one of the biggest and most populated in Madrid, covering 1,142 km^2 with 637,028 inhabitants. In Madrid, the numbers of cases of tuberculosis reported for the years 2001 and 2002 were 1,155 and 1,130, respectively; therefore, our sample represents ~10% per year of the total TB cases in Madrid. The number of cases in Madrid represents a relevant part of the total number of cases in Spain (7,374 and 7,493 for the same years).

Clinical specimens were processed according to standard methods and inoculated in Lowenstein-Jensen slants and in MGIT (Becton Dickinson, Sparks, Md.) medium.

* Corresponding author. Mailing address: Servicio de Microbiología y Enfermedades Infecciosas, Hospital Gregorio Marañón, C/ Dr Esquerdo 46, 28007 Madrid, Spain. Phone: 91 4265104. E-mail: dgviedma@microb.net.

Molecular typing. Spoligotyping was performed according to the manufacturer's instructions and the previously published procedure (11, 23).

IS*6110* restriction fragment length polymorphism (RFLP) analysis was performed according to the international standardization guidelines (24).

Mycobacterial interspersed repetitive unit–variable-number tandem repeat (MIRU-VNTR) typing was performed as described elsewhere (22). MIRU-VNTR is a PCR-based typing method which assigns the number of tandem repeats in 12 independent loci (MIRUs) which are polymorphic in *M. tuberculosis*. The 12 PCRs specific for each MIRU locus were performed with primers and conditions detailed elsewhere (21). PCR products were separated by electrophoresis in MS-8 2% agarose gels (Pronadisa, Madrid, Spain), and the molecular sizes of the amplicons were calculated by comparing their mobilities with a 100-bp ladder marker (Gibco BRL), using the ChemiDoc system (Bio-Rad Laboratories, Richmond, Calif.) and Diversity database software (Bio-Rad). The number of repetitive units for each MIRU was calculated by means of the references included in the MIRU provisional web page (http://www.ibl.fr/mirus/mirus.html). The MIRU type consists of a 12-number code that indicates the number of tandem repeats found for each of the MIRU loci.

Phylogenetic comparisons between genotypes were performed using Bionumerics version 3.0 (Applied Maths, St. Marten-Latems, Belgium).

Database comparison and nomenclature. The SpolDB4 Information System is an automated Access-based labeling and matching system for spoligotyping that will be described elsewhere (C. Delfino, N. Rastogi, and C. Sola, unpublished data). All of the data were also imported into the Bionumerics database format. The main improvement in SpolDB4 relative to SpolDB3 is that the comparison of newly introduced spoligotyping patterns to define new spoligotyping-based clusters with the database is now a fully automated process. SpolDB4 allows the simultaneous (i) introduction of any Excel file containing many hundreds to thousands of octal spoligotyping data (3), (ii) comparison of the new file with all of the spoligotypes contained in the database, and (iii) detection and labeling of new or preexisting matching orphan alleles based on the creation of an incremented shared type (ST) number if, and only if, at least two spoligotypes are identical. Before the introduction of the spoligotyping data set studied here, SpolDB4 contained a total of 26,384 isolates divided into 1,528 shared types and totalling 23,936 isolates plus 2,448 orphans. After the introduction of this file (n = 233), 15 new clusters were defined (ST1529 to ST1543), and SpolDB4 was reorganized into 24,156 isolates in shared types and 2,461 orphan alleles (for superfamily and family nomenclature, see http://www.cdc.gov/ncidod/eid/vol8no11/02-0125.htm) (6, 7). At the end of recruitment in April 2004, SpolDB4 included ~40,000 isolates split into 1,939 shared types and ~3,530 orphan profiles. A synthetic analysis of SpolDB4 is in progress and will be reported elsewhere. The building of SpolDB4 was made possible thanks to ~80 participating laboratories and as many coinvestigators worldwide.

Epidemiological data. For all cases with Madrid1 and Madrid2 *M. tuberculosis* isolates, epidemiological data were retrospectively collected from the Madrid Tuberculosis Register.

RESULTS

Analysis of *M. tuberculosis* genotypes. (i) General features. Before the introduction of the Spain233 file, SpolDB4.0 contained a total of 26,384 isolates divided into 1,528 shared types and totaling 23,936 isolates plus 2,448 orphan isolates (http://www.cdc.gov/ncidod/EID/vol8no11/02-0125-Table.htm) (6, 7). After the introduction of this file (n = 233), 15 new clusters were defined, and SpolDB4 was reorganized into 24,156 isolates in shared types and 2,461 orphan alleles.

If we consider the 233 Spanish isolates alone, 71% were clustered (166 of 233), and the total number of orphan strains was 67 of 233 (29%) (Fig. 1 and Table 1). When we examined the Spanish isolates together with those in SpolDB4, 89% of the isolates were clustered (208 of 233) and 25 strains remained orphans (11%).

In terms of population genetics (Fig. 1), nine clonal complexes of seven or more strains were present in this study: ST50 (Haarlem 3 family; n = 24), ST 53 (superfamily T1; n = 22), ST42 (LAM9 family; n = 14), ST47 (Haarlem 1 family; n = 10), ST58 (n = 10), ST33 (LAM3; n = 9), ST17 (LAM2; n = 9), ST20 (LAM1; n = 7), and ST209 (n = 7).

(ii) Specific features. Some shared types previously proposed as specific to Spain were also found in the Madrid spoligotyping data set (ST105, one new isolate, and ST106, four new isolates) (Table 1 and Fig. 1). Some isolates shared patterns with clusters which were already described in countries which have historical links with Spain, i.e., ST183 and ST222 (found in Peru and Mexico) and ST1227 and ST215 (found in Mexico and Texas) (16) (Table 1).

Of the 27 clusters found for the Spanish isolates, three had not been found in SpolDB4. For the time being, these clusters seem to be specific to the study setting and involve three microclusters of two strains each (Fig. 1): (i) ST 1541, which is close to ST33 (LAM3 family), and ST34 (S family); (ii) ST1543 which is close to ST106, already described in Spain (19); and (iii) ST1531.

After the inclusion of the 233 spoligotypes in SpolDB4, new isolates clustered with some previously orphan profiles, which were found in countries that could be considered historically linked to Spain. These include ST1538 (The Netherlands); ST1537 (Italy; Sicily); ST1535 and ST1536 (Brazil); ST1530, ST1532, ST1534, and ST1542 (United States); and ST1529 and ST1533 (Austria) (Table 1).

The main specific feature for the analysis of Spanish isolates is the presence of two highly prevalent genotypes that are potentially specific Spain-related genotypes. These are ST209 and ST58 (in this study, designated Madrid1 and Madrid2), with seven and nine isolates (Table 2), respectively. Both types had already been described in the international database with no identifiable trend concerning their origins, and now their distribution is statistically overrepresented in Latin countries, as shown in Table 2.

Characterization of the Madrid1 and Madrid2 families. An epidemiological survey did not reveal epilinks among most of the representatives of the Madrid1 or Madrid2 family. Only two cases in Madrid1 were clearly epidemiologically related (02112067 and 02116152) (Table 3). Madrid1 and Madrid2 strains were isolated from Spanish cases, except one case from Madrid1 (02101855), which corresponded to a 2-year-old Chinese child, and one case from Madrid2, a 25-year-old male who was born in Peru (Table 3).

In order to characterize more precisely the Madrid1 and Madrid2 families, additional molecular typing by IS*6110* RFLP and MIRU-VNTR was performed with the 16 isolates belonging to these families.

For ST209 (Madrid1), the RFLP analysis indicated that six isolates showed highly similar patterns, and the remaining isolate showed a pattern with lower similarity (Fig. 2). With regard to MIRUs, all seven isolates had identical genotypes (224326143323) (Fig. 2).

For ST58 (Madrid2), the RFLP analysis indicated that most isolates showed highly similar patterns, with six to eight IS*6110* copies (Fig. 2). With regard to MIRUs, all nine isolates showed highly similar MIRU patterns, with eight of the nine isolates showing an identical number of repeats in 11 of 12 MIRUs and differences only in MIRU 40 (Fig. 2).

The dendrogram obtained following the combined analysis of IS*6110* RFLP, spoligotyping, and MIRU data showed that

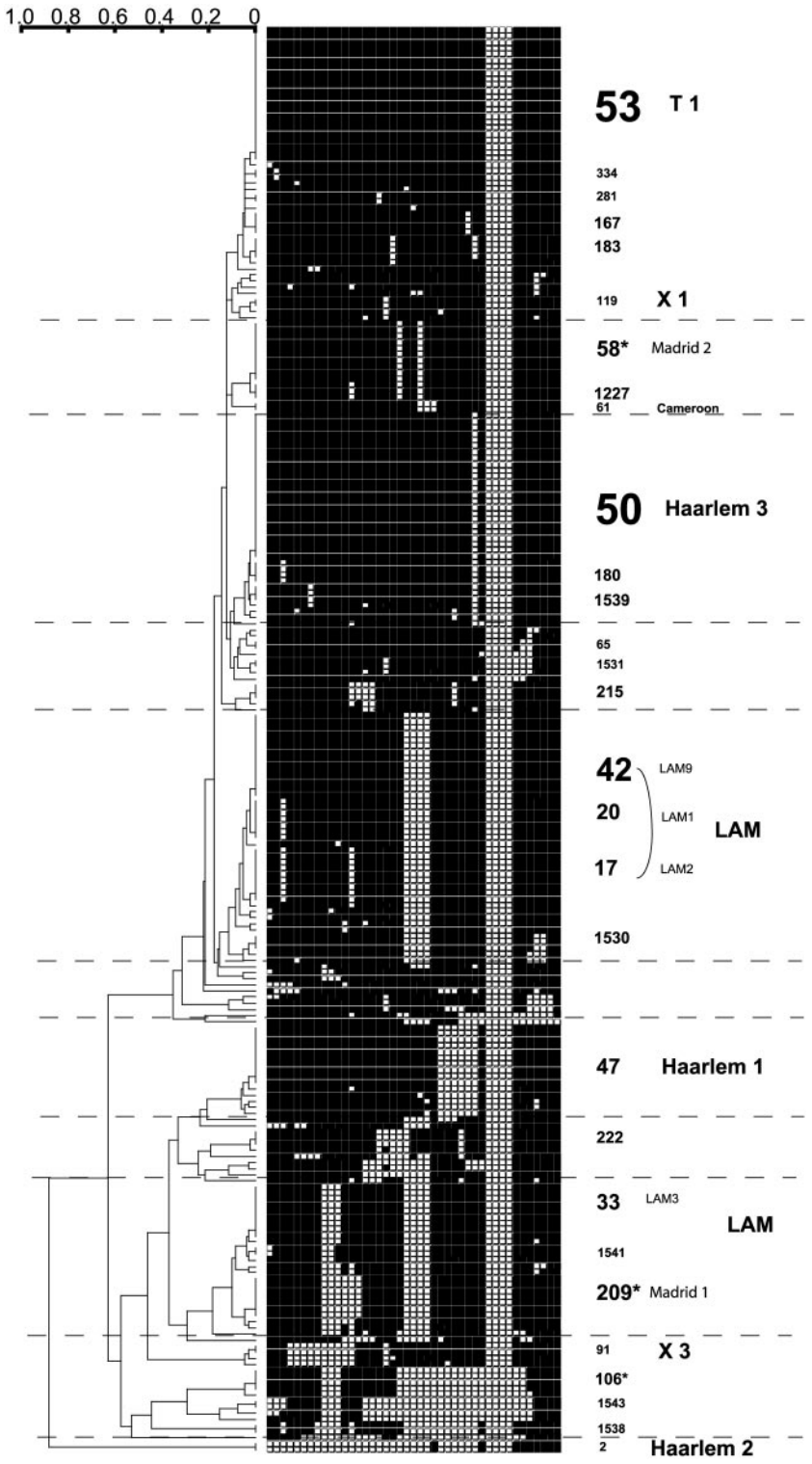

FIG. 1. Similarity dendrogram obtained from the spoligotypes of the Spanish isolates in the spoligotyping results data set ($n = 233$). The scale on the left indicates the genetic distances among the *M. tuberculosis* strains (0 corresponds to identical patterns). On the right, the shared type numbers and family names, when applicable, for the clustered strains are shown. The text size is proportional to the sizes of the clusters. The asterisks indicate the shared types specific to Spain.

TABLE 1. Full results of spoligotyping (in octal) and shared types designation after comparison with SpolDB4 (cf., Material and Methods)

Key no.	Octal	ST[a]	Key no.	Octal	ST[a]
ESP06200101009376	777777777760771	53	ESP06200202050573	777777774020771	47
ESP06200101020433	777777777760771	53	ESP06200202055049	777777774020771	47
ESP06200101023277	777777777760771	53	ESP06200101089376	777777774020731	62
ESP06200101056013	777777777760771	53	ESP06200202059245	777777764020771	45
ESP06200101085093	777777777760771	53	ESP06200202061681	777777764020731	315
ESP06200202019099	777777777760771	53	ESP06200202102117	777777754020771	883
ESP06200202024274	777777777760771	53	ESP06200101092024	777777743760771	61
ESP06200202036688	777777777760771	53	ESP06200202107683	777777743760771	61
ESP06200202039316	777777777760771	53	ESP06200202010617	777777737760771	86
ESP06200202047949	777777777760771	53	ESP06200202054595	777777717760731	513
ESP06200202048407	777777777760771	53	ESP06200202004266	777777677760771	291
ESP06200202053247	777777777760771	53	ESP06200101003070	777777607760771	42
ESP06200202053870	777777777760771	53	ESP06200101008719	777777607760771	42
ESP06200202061308	777777777760771	53	ESP06200101054345	777777607760771	42
ESP06200202074454	777777777760771	53	ESP06200101055333	777777607760771	42
ESP06200202075361	777777777760771	53	ESP06200101057188	777777607760771	42
ESP06200202075478	777777777760771	53	ESP06200101083438	777777607760771	42
ESP06200202077762	777777777760771	53	ESP06200101088055	777777607760771	42
ESP06200202117846	777777777760771	53	ESP06200101091667	777777607760771	42
ESP06200202125364	777777777760771	53	ESP06200202021339	777777607760771	42
ESP06200202129429	777777777760771	53	ESP06200202064688	777777607760771	42
ESP06200202130066	777777777760771	53	ESP06200202066225	777777607760771	42
ESP06200101098626	777777777760731	52	ESP06200202067541	777777607760771	42
ESP06200101063856	777777777760711	78	ESP06200202075028	777777607760771	42
ESP06200202024633	777777777760671	245	ESP06200202095786	777777607760771	42
ESP06200101107485	777777777760631	888	ESP06200202079111	777777607760711	1530* (b)
ESP06200101071003	777777777760471	65	ESP06200202079302	777777607760711	1530* (b)
ESP06200101074946	777777777760471	65	ESP06200202079886	777777607760711	1530* (b)
ESP06200101128853	777777777740071	574	ESP06200202092934	777777607760611	0
ESP06200101003631	777777777720771	50	ESP06200202012793	777777607660771	891
ESP06200101007044	777777777720771	50	ESP06200202045685	777777607400000	534
ESP06200101047262	777777777720771	50	ESP06200202120710	777777604360771	0
ESP06200101062125	777777777720771	50	ESP06200101000678	777777557760771	58
ESP06200101070061	777777777720771	50	ESP06200101033212	777777557760771	58
ESP06200101079889	777777777720771	50	ESP06200101091361	777777557760771	58
ESP06200101088714	777777777720771	50	ESP06200202034006	777777557760771	58
ESP06200101105330	777777777720771	50	ESP06200202047009	777777557760771	58
ESP06200101113340	777777777720771	50	ESP06200202053146	777777557760771	58
ESP06200202005469	777777777720771	50	ESP06200202054204	777777557760771	58
ESP06200202006561	777777777720771	50	ESP06200202058421	777777557760771	58
ESP06200202008484	777777777720771	50	ESP06200202079522	777777557760771	58
ESP06200202009964	777777777720771	50	ESP06200202101807	777777557760771	58
ESP06200202082582	777777777720771	50	ESP06200101092552	777777377760771	40
ESP06200202086379	777777777720771	50	ESP06200101055575	777777377720771	183
ESP06200202096731	777777777720771	50	ESP06200101085969	777777377720771	183
ESP06200202098161	777777777720771	50	ESP06200101126302	777777377720771	183
ESP06200202103151	777777777720771	50	ESP06200202124015	777777377720771	183
ESP06200202110601	777777777720771	50	ESP06200202096784	777776777760771	119
ESP06200202111536	777777777720771	50	ESP06200202102995	777776777760771	119
ESP06200202112395	777777777720771	50	ESP06200202112069	777776777760601	137
ESP06200202127291	777777777720771	50	ESP06200101058898	777776777760071	1531**
ESP06200202128087	777777777720771	50	ESP06200202034992	777776777760071	1531**
ESP06200202132814	777777777720771	50	ESP06200101105130	777776776160601	0
ESP06200101063541	777777777720771	1529*(a)	ESP06200101104707	777776775760771	1532* (c)
ESP06200202010085	777777777660771	167	ESP06200202018759	777775777760771	281
ESP06200202050232	777777777660771	167	ESP06200202061753	777775777760771	281
ESP06200202051391	777777777660771	167	ESP06200202079147	777775077560771	0
ESP06200202060878	777777777660771	167	ESP0619970970436	777774077560771	222
ESP06200101077041	777777777320771	457	ESP06200202058276	777774077560771	222
ESP06200000131	777777774020771	47	ESP06200202079188	777774077560771	222
ESP06200101017492	777777774020771	47	ESP06200202089423	777766777760731	0
ESP06200101035352	777777774020771	47	ESP06200202044915	777766777760071	0
ESP06200101055858	777777774020771	47	ESP06200202114420	777763777720771	1533* (d)
ESP06200101070485	777777774020771	47	ESP06200202024348	777761007600771	0
ESP06200101076401	777777774020771	47	ESP06200202118557	777760007760771	254
ESP06200101079938	777777774020771	47	ESP06200101081257	777760007600771	0
ESP06200202040612	777777774020771	47	ESP06200202075509	777737777700771	0

Continued on facing page

TABLE 1—*Continued*

Key no.	Octal	ST[a]	Key no.	Octal	ST[a]
ESP06200202029726	777737774020771	218	ESP06200202005040	773777777720771	1539* (j)
ESP06200202053152	777737607760771	93	ESP06200202023798	773767777720771	0
ESP06200101073952	777737557760771	1227	ESP06200101066272	771777777760771	1069
ESP06200202124576	777737557760771	1227	ESP06200202124050	760001400000171	29
ESP06200202136436	777737557760771	1227	ESP06200101043848	757777777760771	154
ESP06200101109124	777737477400001	0	ESP06200202075503	757777777320771	433
ESP06200202016612	777723777360771	0	ESP06200101120091	757777607760611	0
ESP06200202108205	777721607560731	1534* (e)	ESP06200202111140	737737777760731	0
ESP062001010130614	777703777360771	215	ESP06200101114557	700036777760771	91
ESP06200202092941	777703777360771	215	ESP06200202016149	700036777760771	91
ESP06200202111055	777703777360771	215	ESP06200202103753	700036776760771	0
ESP06200101082979	777677607760771	770	ESP06200202089427	700036377760771	0
ESP06200202039161	777667607760771	0	ESP06200202033204	677777777720771	180
ESP06200202116155	777603405760471	136	ESP06200202059756	677777777720771	180
ESP06200202053242	777577607760771	1535* (f)	ESP06200202116151	677777777720771	180
ESP06200202089621	776777707720771	0	ESP06200202136780	677777777720771	180
ESP06200101101275	776377607760771	1536* (g)	ESP06200101008451	677777607760771	20
ESP06200101018901	776177777760771	156	ESP06200101098784	677777607760771	20
ESP06200101014943	776177607760771	33	ESP06200101104470	677777607760771	20
ESP06200101018215	776177607760771	33	ESP06200101113508	677777607760771	20
ESP06200101024635	776177607760771	33	ESP06200202048022	677777607760771	20
ESP06200101071603	776177607760771	33	ESP06200202115601	677777607760771	20
ESP06200101104985	776177607760771	33	ESP06200202118380	677777607760771	20
ESP06200101111774	776177607760771	33	ESP06200101002434	677737607760771	17
ESP06200202044179	776177607760771	33	ESP06200101022121	677737607760771	17
ESP06200202105761	776177607760771	33	ESP06200101022580	677737607760771	17
ESP06200202110027	776177607760771	33	ESP06200101085378	677737607760771	17
ESP06200202006621	776177607760731	130	ESP06200101093934	677737607760771	17
ESP06200101015929	776177400000171	106	ESP06200202006486	677737607760771	17
ESP06200101072339	776177400000171	106	ESP06200202029095	677737607760771	17
ESP06200202028011	776177400000171	106	ESP06200202034480	677737607760771	17
ESP06200202090934	776177400000171	106	ESP06200202121465	677737607760771	17
ESP06200101101178	776177400000071	0	ESP06200202063828	577777777760771	334
ESP06200101072015	776160000000071	105	ESP06200202101528	577777777760771	334
ESP06200202037345	776137607760771	211	ESP06200202018681	437777774320731	0
ESP06200202061812	776137607760731	1537* (h)	ESP06200101124628	377777777760771	7
ESP06200101022818	776137607760711	0	ESP06200202054848	377777607760771	177
ESP06200101059429	776127400000171	0	ESP06200202098021	377377607760771	1540* (k)
ESP06200202016105	776037607760771	0	ESP06200202089426	376377777760771	884
ESP06200101033451	776017607760771	209	ESP06200202035411	376377607760771	1541**
ESP06200101113046	776017607760771	209	ESP06200202049608	376377607760771	1541**
ESP06200202006451	776017607760771	209	ESP06200202110289	177776777760601	1542* (l)
ESP06200202017030	776017607760771	209	ESP06200101091768	176160000000071	0
ESP06200202101855	776017607760771	209	ESP06200101055443	076160000000071	1543**
ESP06200202112067	776017607760771	209	ESP06200202115336	076160000000071	1543**
ESP06200202116152	776017607760771	209	ESP06200101078029	047777607760771	0
ESP06200202039149	774137600020771	1538* (i)	ESP06200202100200	037677777760771	1278
ESP06200202115484	774137600020771	1538* (i)	ESP06200101047458	000000004020771	2
ESP06200101050817	773777777720771	1539* (j)	ESP06200202035466	000000004020771	2
ESP06200101103661	773777777720771	1539* (j)			

[a] *, new shared type; **, new specific shared type (not found elsewhere); (a), match with a strain from Austria; (b), match with a strain from the United States; (c), match with a strain from the United States; (d) match with a strain from Austria; (e), match with a strain from the United States; (f), match with a strain from Brazil; (g), match with a strain from Brazil; (h), match with a strain from Italy; (i), match with a strain from Belgium (Moroccan immigrant); (j), match with a strain from Malaysia; (k), match with a strain from Indonesia; (l), match with a strain from the United States.

the isolates of the Madrid1 and Madrid2 families clustered in two similarity groups (Fig. 2).

DISCUSSION

In this study, we have defined the genetic structure of the population of circulating *M. tuberculosis* strains in Madrid based on a snapshot study of 233 clinical isolates. In our city, the numbers of cases of tuberculosis reported for the years 2001 and 2002 (1,155 and 1,130 cases, respectively) represent a relevant part of the total number of cases in Spain (7,374 and 7,493 for the same years). Therefore, we studied the clonal composition of *M. tuberculosis* strains in Madrid to obtain an idea of the clonal structure of tuberculosis in Spain at a time which coincides with marked increases in the number of immigrants coming to Spain and in the number of cases of tuberculosis that can be attributed to imported cases (from 6.7% in 1997 to 99 to 29.4% in 2002 and 2003). Different studies have analyzed the impact of immigration on the transmission dynamics of tuberculosis (2, 8, 12). However, most of these

TABLE 2. Characteristics of new Madrid 1 (ST209/LAM10) and Madrid 2 (ST58) families

Family	Spoligotype (binary)	ST no.	IS copy no.[f]	Total no.	Origin[a]	Reference[j]
Madrid 1	■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	209	18	2	CUB	Diaz et al. (3a)
	■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	209	ND	3	ESP	March et al. (12a)
	■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	209	16	2	FXX	Maugein et al., unpublished data
	■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	209	ND	1	USA (Tx)	Soini et al. (17)[b]
	■■■■■■■■□□□□□□■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	209	16–17	7	ESP	This study
Madrid 2	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	5	BRA	Rossetti, unpublished data (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	BRA	Dellagostin, unpublished data (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	5	CUB	Diaz et al. (3a)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	DZA	RIIPIA, unpublished data (SpolDB4)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	2	ESP	March et al. (12a)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	5	2	FXX (Bdx)	Maugein, unpublished data (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	2	FXX (Par)[h]	Goguet et al. (9)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	GUF(BRA)[g]	IPG (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	ITA	Bonora et al. (1a)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	4	NLD	RIVM (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	Type U, 7 bands (6–10)[c]	23	USA (NYS)	Driscoll et al., unpublished data (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	7	USA (Tx)	Soini et al. (17)[d]
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	VEN	Takiff et al., unpublished data (SpolDB3)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	3	AUT	Prodinger et al. (15a)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	USA (Mi)[i]	Cowan et al. (1)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	3	USA (Tx)	Quitugua et al. (16)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	NA	1	MEX	Quitugua et al. (16)
	■■■■■■■■■■■■■■■■■■■□■■□■■■■■■■■■■□□□□■■■■■■■	58	6 to 8[e]	9	ESP	This study

[a] BRA, Brazil; CUB, Cuba; DZA, Algeria; FXX, Metropolitan France; Bdx, Bordeaux; Par, Paris; ITA, Italy; NLD, The Netherlands; USA, United States of America; NYS, State of New York; Tx, Texas; Mi, Michigan; Mex, Mexico; ESP, Spain.

[b] Isolate belonging to the principal genetic group 2 according to Sreevatsan et al. (20a), using *katG-gyrA* polymorphism.

[c] Type U from the United States is different from the strains harboring identical seven copies in Spain.

[d] Isolate belonging to the principal genetic group 3 according to Sreevatsan et al. (20a), using *katG-gyrA* polymorphism.

[e] All strains from Spain highly similar.

[f] ND, not done; NA, not available.

[g] Brazilian immigrant; VNTR of this clinical isolate, 32533.

[h] MIRU value of this clinical isolate, 223326153324.

[i] MIRU value of this clinical isolate, 223326153324.

[j] IPG, Institut Pasteur Guadeloupe; RIIPIA, Reseau International des Instituts Pasteur et Instituts Associes; RIVM, Rijk Institute of Veterinary Medicine.

TABLE 3. General features of cases with Madrid1 and Madrid2 *M. tuberculosis* strains

Sex[a]	Age (yr)	Nationality	Site of disease	Risk factor for TB	Epilink[b]	*M. tuberculosis* strain
M	42	Spanish	Lung	HIV$^+$, IVDU[a]	No	Madrid1
M	36	Spanish	Ganglia	HIV$^+$, IVDU	No	Madrid1
M	21	Spanish	Lung	None	Yes	Madrid1
M	21	Spanish	Lung	None	Yes	Madrid1
M	76	Spanish	Lung	None	No	Madrid1
M	2	Chinese	Lung	None	No	Madrid1
M	47	Spanish	Lung	HIV$^+$, prison	No	Madrid1
F	25	Peruvian	Lung	None	No	Madrid2
M	65	Spanish	Lung	None	No	Madrid2
F	27	Spanish	Disseminated	None	No	Madrid2
M	34	Spanish	Disseminated	HIV$^+$	No	Madrid2
M	21	Spanish	Lung	None	No	Madrid2
F	74	Spanish	Lung, nervous system	None	No	Madrid2
M	59	Spanish	Lung	Alcoholism, IVDU	No	Madrid2
M	56	Spanish	Lung	HIV$^+$	No	Madrid2
M	47	Spanish	Lung	Alcoholism, IVDU	No	Madrid2

[a] M, male; F, female.
[b] Indicates whether case had epidemiological links with another case(s) within the Madrid1 or Madrid2 group.
[c] HIV, human immunodeficiency virus; IVDU, intravenous drug user.

studies were performed in countries with a low prevalence of TB, a long history of immigration, and no snapshot of the situation before the immigrant population increased. Our study, on the other hand, was carried out during a transition period for TB transmission dynamics, which makes it possible to define a baseline reference for population genetic structure in Spain. This reference can now be used to precisely monitor the effect of immigration on the patterns and transmission dynamics of TB in the coming years.

In order to understand the genetic structure of the worldwide *M. tuberculosis* population and its evolution, the creation and development of international bacterial genotyping databases which gather and share *M. tuberculosis* typing patterns is of enormous value. These genotyping databases can help us to understand both the general and particular features of tuberculosis transmission worldwide, to detect casual transmission cases, and to define *M. tuberculosis* strains specific to different geographic settings which could be further chosen as reporter strains to monitor international TB routes. SpolDB4 allows us to perform this global analysis. Unfortunately, not all countries are equally represented in global databases, and prior to this study, the Spanish data included in the SpolDB3 database were limited. As a result, the general and specific features of the *M. tuberculosis* population in Spain were not easily appreciated.

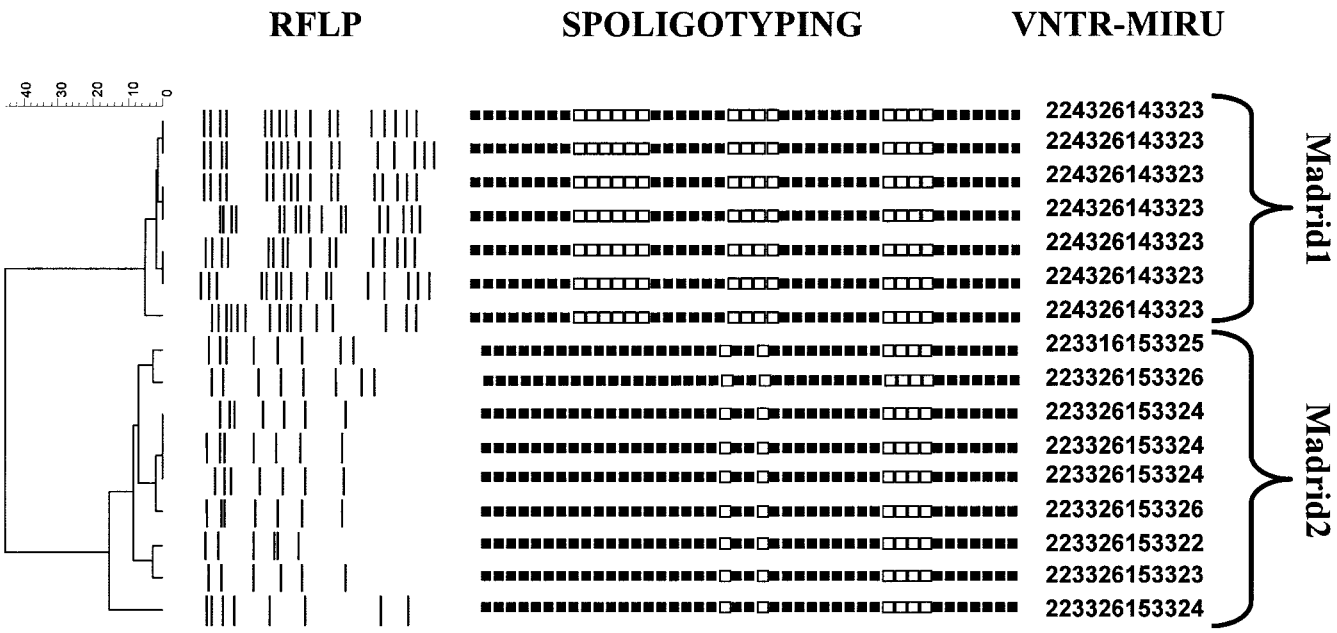

FIG. 2. Similarity dendrogram obtained after combining the RFLP, spoligotyping, and VNTR-MIRU data for the isolates belonging to the Madrid1 (ST209) and Madrid2 (ST58) families. The scale on the left indicates the genetic distances among the *M. tuberculosis* strains (0 corresponds to identical patterns) taking RFLP, spoligotypes, and MIRU types together. On the right, the numbers for the isolates are shown.

The inclusion of 233 spoligotyping patterns and some MIRU patterns from Spanish isolates in our study has increased our knowledge of this population.

If we compare Spanish spoligotyping patterns with all the data compiled in SpolDB4 from European isolates, we observe that the population structure found in our study is characteristic of a European country and shares the most prevalent European genotypes. For the Spanish isolates, ST50 (Haarlem 3 family) and ST53 (ill-defined T1) are the most prevalent genotypes ($n = 24$ and 22). The Latin American and Mediterranean (LAM) superfamily (20), with all its variants (ST17, ST20, ST42, and ST33), when taken as a whole, is the predominant superfamily ($n = 39$).

Several of our findings are consistent with a highly structured history of tuberculosis in Madrid, both historically and geographically. The inclusion of the Spanish data file revealed the following: (i) new isolates sharing patterns with genotypes previously proposed as specific to Spain or historically or geographically related countries; (ii) new isolates clustered with orphan profiles from countries which are related to Spain; and (iii) the most interesting finding, two new families of homogeneous genetic structures, which are likely to define Spain-specific clonal complexes of *M. tuberculosis*.

The first Spain-specific clade, ST209, is a new member of the LAM family of strains (designated LAM11 and Madrid1 in our study), and the second one is ST58 (designated Madrid2 in our study).

The distribution of ST209 (Madrid1) (Table 2) suggests Spanish phylogeographic specificity. Indeed, of 15 strains harboring this type, we found 10 in Spain, and the remaining isolates were found in Cuba, southwest France (Bordeaux), and Texas. Only one isolate belonging to ST209 was found in 2004 in the database of the Public Health Research Institute in the state of New York (J. Driscoll and B. Kreiwirth, personal communication). Most of the ST209 strains were shown to harbor identical or very similar 16- to 18-band IS*6110* RFLPs (Fig. 2). In Spain, most of the strains harboring ST209 were also very similar by IS*6110* RFLP, and they all had identical MIRUs (Fig. 2 and Table 2).

With regard to the other Spain-specific clade, ST58 (Madrid2), 71 strains harboring spoligotype ST58 were found in the database on 4 September 2003. They were distributed in Spain and other countries geographically or historically linked with Spain, including Brazil, Cuba, Algeria, France (two of which again originated in Bordeaux, in southwest France, near the Spanish border), French Guiana (from a Brazilian immigrant), Italy, The Netherlands, and the United States (the majority of which were found in Texas and New York [Latin-American cases]), and also in Venezuela and Mexico (Table 2). The distribution of this family, which is likely to belong to principal genetic group 3 (17), suggests a quite recent introduction and spread of a specific ancestral clone, perhaps from Austria or Italy. This clone may have spread and evolved specifically in Spanish-speaking countries and later spread to the United States.

A search of available IS*6110* RFLP data for ST58 isolates from the SpolDB4 database resulted in various distinct profiles with five to eight bands (Table 2 and Fig. 2). In Spain, however, most of the strains harboring ST58 were also identical or highly similar by IS*6110* RFLP and also showed highly similar or identical MIRU types. The MIRU type shared by four of nine ST58 isolates in Spain (223326153324) was also shared by three ST58 isolates (nonepidemiologically linked) from France and the United States (Michigan) (1, 9, 13).

A hypothetical scenario of evolution for this clonal complex suggests that ST44, which belongs to principal genetic group 3 (17), could be the origin of ST58 by the loss of spacer 19. This hypothesis relies on the finding of a high prevalence of ST44 in Austria, Italy, and the Czech Republic (39 out of 72 strains harboring this type in these three countries), countries whose history is closely linked with the former Spanish empire. However, at this stage, we cannot eliminate the possibility that the Madrid2 definition covers more than one clone and that convergence of spoligotypes occurred in two separate clones without a common ancestor, as demonstrated previously (27). However, the possibility of diverging IS*6110* RFLP profiles within a single genetic ST58 background cannot be excluded.

It could be argued that our study is a snapshot of isolates obtained from a very limited geographic area and that the definition of two genotypes proposed to be Spain related must demonstrate that these genotypes are also found in other Spanish regions. Madrid1 and Madrid2 have also been found (two Madrid1 and one Madrid2 in 112 isolates) in a study which is currently being carried out in Almería (Andalucía, in southern Spain). Furthermore, Madrid1 and Madrid2 were found in another study in Segovia between 1995 and 1999 (2 Madrid1 and 4 Madrid2 in 96 strains analyzed). Finally, these genotypes were also found throughout Spain in the national *M. tuberculosis* database from the University of Zaragoza: 15 Madrid1 and 23 Madrid2 isolates from five and eight regions other than Madrid, respectively. It is also worth noting that Madrid1 and Madrid2 have been identified in 4 of 30 isolates from elderly people with tuberculosis (>65 years of age; probable reactivations) in Madrid (M. J. Rebollo, E. Palenque, S. Samper, F. Jaen, and M. J. Garcia, XI GEM Meeting, Torrelavega, Spain, poster, 2004), which suggests endemicity of these genotypes. Since the initial comparison made in September 2003 (15 ST209 and 71 ST58 isolates found), new submissions to SpolDB4 have increased the total number of clinical isolates harboring ST209 and ST58 in October 2004 to 24 and 95, respectively.

Altogether, these data demonstrate the highly historically and geographically structured history of tuberculosis in Madrid, a history that may be assessed both retrospectively and prospectively by genotyping using polymorphic markers, such as MIRU and spoligotyping, and by accessing worldwide genetic diversity databases. The distribution of isolates within 26,617 isolates in the database representing >100 countries suggests the high specificity of ST58 and ST209 clonal complexes for Spain and Spain-related areas.

Although robust, spoligotyping may be subject to misinterpretation, and the presence of typing artifacts should always be considered when interpreting results. MIRU-VNTR is presently the best complementary technique to demonstrate the clonality of some isolates, and it has been applied in this study to the ST58 and ST209 isolates. Congruence in the identity of MIRU markers makes it seem unlikely that ST58 and ST209 were produced by convergence and that they are not part of phylogenetic families. MIRU analysis will undoubtedly enable us to further discriminate the strains in ST58 and ST209 from

other studies and may be helpful to discriminate between phylogenetic clusters and epiclusters. In parallel, contact tracing may reveal the presence of an epidemiological contact within these clusters. With the exception of two cases, an epidemiological survey in our study did not reveal epilinks among cases with Madrid1 and Madrid2 isolates. This rules out the possibility that these strains represent ongoing transmission within a specific population.

The existence of SpolDB4 and the public release of SpolDB3 should foster a further search for homogeneous clusters, as well as the description of their genetic characteristics, such as spoligotyping signature, most frequent VNTR and MIRU alleles, and IS*6110* RFLP profiles. Recently, two new families of strains have been defined thanks to the retrospective analysis of IS*6110* RFLP profiles. In Tanzania, McHugh et al. identified the Kilimanjaro 1 TB lineage, a characteristically Tanzanian clone, which may be a sublineage of the Central Asian 1-Delhi lineage (14). In the Philippines, Douglas et al. and Sola et al. described the Manila family, (also designated as the East African-Indian 2 genotype), which is part of the largest EAI superfamily (4, 6). Recently, the Cameroon family of TB bacilli was also identified (15). Undoubtedly, many more genetic families or lineages of TB bacilli, with as yet unknown specific pathogenicities, will be described in the near future, and these, together with Madrid1 and Madrid2, would be useful as reporter strains to explain the international routes of transmission of tuberculosis worldwide.

In conclusion, our study defines the general and specific features of the clonal composition of the *M. tuberculosis* population in Spain at a point in time when the immigrant population is increasing rapidly. Our data could therefore provide a useful reference for the composition of *M. tuberculosis* strains in a country before the influence of immigration on the population genetic structure becomes too important and erases historical traces of previous TB epidemics. These data will be used in forthcoming years as a reference for the real-time observation of the likely effects that a marked increase in immigration from countries with a high TB burden could have on the patterns and transmission dynamics of TB.

ACKNOWLEDGMENTS

We are grateful to Jeff Driscoll, Barun Mathema, and Barry Kreiswirth for communication of unpublished genotyping information on ST58 and ST209 clinical isolates from the state of New York. We also thank Jeanne Maugein from Bordeaux, France, for communication of unpublished information on clinical isolates; Fernando Chaves and Pablo Carrero for data from Segovia, Spain; Elia Palenque and Maria Jesus Garcia for data from elderly patients in Spain; Sofia Samper and Carlos Martin for data from the national *M. tuberculosis* database from the University of Zaragoza; and Jesús Iñigo and Elena Rodriguez for the epidemiological survey of cases with Madrid1 and Madrid2 isolates in this study. We are also thankful to Sandra Andrés and Yolanda Paredes for their support with the genotyping and to Thomas O'Boyle for his revision of the English in this article.

This work has been partially financed by grants from Comunidad de Madrid (08.2/0029.1/01) and Fondo de Investigaciones Sanitarias (02/0882; 03/0654).

This research has been approved by the appropriate research review boards. No organization with a financial interest in the subject matter was involved in the manuscript. The authors have disclosed any conflict of interest related to this article.

REFERENCES

1. **Cowan, L. S., L. Mosher, L. Diem, J. P. Massey, and J. T. Crawford.** 2002. Variable-number tandem repeat typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS*6110* by using mycobacterial interspersed repetitive units. J. Clin. Microbiol. **40:**1592–1602.
2. **Dahle, U. R., P. Sandven, E. Heldal, and D. A. Caugant.** 2001. Molecular epidemiology of *Mycobacterium tuberculosis* in Norway. J. Clin. Microbiol. **39:**1802–1807.
3. **Dale, J. W., D. Brittain, A. A. Cataldi, D. Cousins, J. T. Crawford, J. Driscoll, H. Heersma, T. Lillebaek, T. Quitugua, N. Rastogi, R. Skuce, C. Sola, D. van Soolingen, and V. Vincent.** 2001. Spacer oligonucleotide typing of *Mycobacterium tuberculosis*: recommendations for standardized nomenclature. Int. J. Tuberc. Lung Dis. **5:**216–219.
4. **Douglas, J. T., L. Qian, J. C. Montoya, J. M. Musser, J. D. van Embden, D. van Soolingen, and K. Kremer.** 2003. Characterization of the Manila family of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **41:**2723–2726.
5. **Duchene, V., S. Ferdinand, I. Filliol, J. F. Guégan, N. Rastogi, and C. Sola.** 2004. Phylogenetic reconstruction of the *Mycobacterium tuberculosis* complex within four settings of the Caribbean region: tree comparative analysis and first appraisal on their phylogeography. Infect. Gen. Evol. **4:**5–14.
6. **Filliol, I., J. R. Driscoll, D. van Soolingen, B. N. Kreiswirth, K. Kremer, G. Valétudie, 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. Drobniewski, G. Engelmann, S. Ferdinand, D. Gascoyne-Binzi, M. Gordon, M. C. Gutierrez, W. Haas, H. Heersma, G. Källenius, E. Kassa-Kelembho, T. Koivula, H. M. Ly, A. Makristathis, C. Mammina, G. Martin, P. Moström, I. Mokrousov, V. Narbonne, O. Narvskaya, A. Nastasi, S. Ngo Niobe Eyangoh, J. W. Pape, V. Rasolofo-Razanamparany, M. Ridell, M. L. Rossetti, F. Stauffer, P. N. Suffys, H. Takiff, J. Texier-Maugein, V. Vincent, J. de Waard, C. Sola, and N. Rastogi.** 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. Emerg. Infect. Dis. **8:**1347–1350.
7. **Filliol, I., J. R. Driscoll, D. van Soolingen, B. N. Kreiswirth, K. Kremer, G. Valetudie, D. A. Dang, R. Barlow, D. Banerjee, P. J. Bifani, K. Brudey, A. Cataldi, R. C. Cooksey, D. V. Cousins, J. W. Dale, O. A. Dellagostin, F. Drobniewski, G. Engelmann, S. Ferdinand, D. Gascoyne-Binzi, M. Gordon, M. C. Gutierrez, W. H. Haas, H. Heersma, E. Kassa-Kelembho, M. L. Ho, A. Makristathis, C. Mammina, G. Martin, P. Mostrom, I. Mokrousov, V. Narbonne, O. Narvskaya, A. Nastasi, S. N. Niobe-Eyangoh, J. W. Pape, V. Rasolofo-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.** 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.
8. **Geng, E., B. Kreiswirth, C. Driver, J. Li, J. Burzynski, P. DellaLatta, A. LaPaz, and N. W. Schluger.** 2002. Changes in the transmission of tuberculosis in New York City from 1990 to 1999. N. Engl. J. Med. **346:**1453–1458.
9. **Goguet de la Salmonière, Y. O., H. M. Li, G. Torrea, A. E. Bunschoten, J. D. A. van Embden, and B. Gicquel.** 1997. Evaluation of spoligotyping in a study of the transmission of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **35:**2210–2214.
10. **Grupo de Trabajo de los Talleres de 2001 y 2002 de la Unidad de Investigación en Tuberculosis de Barcelona.** 2003. Prevencion y control de las tuberculosis importadas. Med. Clin. (Barcelona) **121:**549–557.
11. **Kamerbeek, J., L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, and J. D. A. van Embden.** 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. **35:**907–914.
12. **Lillebaek, T., A. B. Andersen, J. Bauer, A. Dirksen, S. Glismann, P. de Haas, and A. Kok-Jensen.** 2001. Risk of *Mycobacterium tuberculosis* transmission in a low-incidence country due to immigration from high-incidence areas. J. Clin. Microbiol. **39:**855–861.
13. **Mazars, E., S. Lesjean, A. L. Banuls, M. Gilbert, V. Vincent, B. Gicquel, M. Tibayrenc, C. Locht, and P. Supply.** 2001. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. Proc. Natl. Acad. Sci. USA **98:**1901–1906.
14. **McHugh, T. D., S. L. Batt, R. J. Shorten, R. D. Gosling, L. Uiso, and S. H. Gillespie.** *Mycobacterium tuberculosis* lineage: a naming of the parts. Tuberculosis, in press.
15. **Ngo Niobe-Eyangoh, S., C. Kuaban, P. Sorlin, P. Cunin, J. Thonnon, C. Sola, N. Rastogi, V. Vincent, and M. C. Gutierrez.** 2003. Genetic biodiversity of *Mycobacterium tuberculosis* complex strains from patients with pulmonary tuberculosis in Cameroon. J. Clin. Microbiol. **41:**2547–2553.
16. **Quitugua, T. N., B. J. Seaworth, S. E. Weis, J. P. Taylor, J. S. Gillette, I. I. Rosas, K. C. Jost, Jr., D. M. Magee, and R. A. Cox.** 2002. Transmission of drug-resistant tuberculosis in Texas and Mexico. J. Clin. Microbiol. **40:**2716–2724.
17. **Soini, H., X. Pan, A. Amin, E. A. Graviss, A. Siddiqui, and J. M. Musser.** 2000. Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston, Texas, by spoligotyping. J. Clin. Microbiol. **38:**669–676.

18. **Sola, C., S. Ferdinand, C. Mammina, A. Nastasi, and N. Rastogi.** 2001. Genetic diversity of *Mycobacterium tuberculosis* in Sicily based on spoligotyping and variable number of tandem DNA repeats and comparison with a spoligotyping database for population-based analysis. J. Clin. Microbiol. **39:**1559–1565.
19. **Sola, C., I. Filliol, C. Guttierez, I. Mokrousov, V. Vincent, and N. Rastogi.** 2001. Spoligotype database of *Mycobacterium tuberculosis*: biogeographical distribution of shared types and epidemiological and phylogenetic perspectives. Emerg. Infect. Dis. **7:**390–396.
20. **Sola, C., I. Filliol, E. Legrand, I. Mokrousov, and N. Rastogi.** 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS*1081*, IS*6110*, VNTR and DR-based spoligotyping suggests the existence of two new phylogeographical clades. J. Mol. Evol. **53:**680–689.
21. **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.
22. **Supply, P., E. Mazars, S. Lesjean, V. Vincent, B. Gicquel, and C. Locht.** 2000. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Mol. Microbiol. **36:**762–771.
23. **van Embden, J. D. A., T. van Gorkom, K. Kremer, R. Jansen, B. A. M. van der Zeijst, and L. M. Schouls.** 2000. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. J. Bacteriol. **182:**2393–2401.
24. **van Embden, J. D. A., M. D. Cave, J. T. Crawford, J. W. Dale, K. D. Eisenach, B. Gicquel, P. Hermans, C. Martin, R. McAdam, T. M. Shinnick, and P. M. Small.** 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. **31:**406–409.
25. **van Soolingen, D.** 2001. Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. J. Intern. Med. **249:**1–26.
26. **Veen, J.** 1992. Microepidemics of tuberculosis: the stone-in-the-pond principle. Tuberc. Lung Dis. **73:**73–76.
27. **Warren, R. M., E. M. Streicher, S. L. Sampson, G. D. Van Der Spuy, M. Richardson, D. Nguyen, M. A. Behr, T. C. Victor, and P. D. van Helden.** 2002. Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. J. Clin. Microbiol. **40:**4457–4465.