

Mycobacterium tuberculosis complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay

Jian Zhang,¹ Edgar Abadia,¹ Guislaine Refregier,¹ Silva Tafaj,² Maria Laura Boschiroli,³ Bertrand Guillard,⁴ Antoine Andremont,⁵ Raymond Ruimy⁵ and Christophe Sola^{1,6}

Correspondence

Christophe Sola
christophe.sola@u-psud.fr

¹IGEPE Team, Institute of Genetics and Microbiology, UMR8621, Universud, CNRS Université Paris-Sud 11, Campus d'Orsay, F-91405 Orsay-Cedex, France

²National TB Reference Laboratory, University Hospital of Lung Diseases 'Shefqet Ndroqi', Tirana, Albania

³Agence Française de Sécurité Sanitaire des Aliments, Maisons-Alfort, France

⁴Institut Pasteur du Cambodge, Pnom-Penh, Cambodia

⁵Microbiology Laboratory, Bichat-Claude Bernard Hospital, Paris, France

⁶Unité de Génétique Mycobactérienne, Institut Pasteur, Paris, France

The aims of the present study were to implement a microbead-based 'spoligotyping' technique and to evaluate improvements by the addition of a panel of 25 extra spacers that we expected to provide an increased resolution on principal genetic group 1 (PGG 1) strains. We confirmed the high sensitivity and reproducibility of the classical technique using the 43 spacer panel and we obtained perfect agreement between the membrane-based and the microbead-based techniques. We further demonstrated an increase in the discriminative power of an extended 68 spacer format for differentiation of PGG 1 clinical isolates, in particular for the East African-Indian clade. Finally, we define a limited yet highly informative reduced 10 spacer panel set which could offer a more cost-effective option for implementation in resource-limited countries and that could decrease the need for additional VNTR (variable number of tandem repeats) genotyping work in molecular epidemiological studies. We also present an economic analysis comparing membrane-based and microbead-based techniques.

Received 21 October 2009
Accepted 26 November 2009

INTRODUCTION

The current global plan for tuberculosis (TB) control suggests that it will be difficult to eradicate the disease by 2050 (Dye & Williams, 2008). This issue is particularly

relevant to large nations with high TB infection rates, such as Brazil, Russia, India, China and Mexico, but it is also relevant to resource-poor countries in Asia and Africa, which are facing the combined threat of human immunodeficiency virus and TB (Uplekar & Lonroth, 2007). TB transmission chains may be interrupted in resource-poor countries by improvements in public health programmes, medical diagnostics facilities and provision of adequate drug treatment (Dye & Williams, 2008). The Millennium Development Goals of the United Nations Organization has called for the detection of 70 % of active cases through the development of rapid diagnostics methods, and the attainment of 80 % cure rates, through means such as expansion of the Directly Observed Treatment of Short Course (DOTS+) programme. However, the expanding threat of multidrug resistant (MDR) TB (resistance to

Abbreviations: CAS, Central Asian; CRISPR, clustered regularly interspaced palindromic repeats; DR, direct repeat; EAI, East African-Indian; EDC, *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide; LAM, Latin-American and Mediterranean; DVR, direct variable repeat; HGDI, Hunter and Gaston index; IGEPE, infection, genetics, emerging pathogens, evolution; MDR, multidrug resistant; MTC, *Mycobacterium tuberculosis* complex; PGG, principal genetic group; S : N, signal to noise ratio; SNP, single nucleotide polymorphism; TMAC, tetra-methyl ammonium chloride; VNTR, variable number of tandem repeats.

A table and figure of spoligotyping data are available as supplementary material with the online version of this paper.

isoniazid and rifampicin) and now of extremely drug resistant (XDR) TB, which is defined as a MDR-TB that is resistant to any fluoroquinolones plus at least one of the three injectable second-line drugs (amikacin, capreomycin, kanamycin) (Manissero & Fernandez de la Hoz, 2006), has made the implementation of public health measures, and molecular epidemiological investigations using rapid and high throughput molecular methods, a crucial priority to achieve these goals.

TB is caused by the *Mycobacterium tuberculosis* complex (MTC), which can be defined genetically as a single bacterial species within the order *Actinomycetales* and encompasses several ecotypes or subspecies that include *M. tuberculosis*, *Mycobacterium bovis*, *Mycobacterium africanum*, '*Mycobacterium canettii*', *M. microti*, *M. caprae* and *Mycobacterium pinipedii* (Rastogi *et al.*, 2001; Smith *et al.*, 2006). Rapid molecular diagnostics methods have been developed during the two last decades; however, these methods have been slow to be implemented worldwide (Drobniewski *et al.*, 2003). Spoligotyping, comprises a PCR-based reverse line blot hybridization method that assesses the genetic diversity of the direct repeat (DR) locus (Kamerbeek *et al.*, 1997). The DR locus spans up to 5 kb and represents 0.1% of the MTC genome. It is a member of the clustered regularly interspaced palindromic repeats (CRISPR) loci genetic family and is found in all Archaea and around 40% of Eubacteria (Jansen *et al.*, 2002). CRISPR loci are believed to have a functional role against bacteriophage challenge, and at the evolutionary genetics level contain remnants and 'memories' of previous bacteriophage encounters. Due to their genetic polymorphisms, CRISPR loci may be excellent epidemiological markers (Sorek *et al.*, 2008).

The MTC DR locus is made up of a limited number of direct variable repeats (DVRs), each repeat is made up of a conserved 36 bp sequence (the DR) and of a variable 35–41 bp sequence (the spacer). A total of 43 spacers were initially chosen, comprising 6 spacers from *M. bovis* BCG and 37 spacers from the *M. tuberculosis* H37Rv genomes (Groenen *et al.*, 1993; Kamerbeek *et al.*, 1997). Only 10 of these spacers are required to differentiate many of the MTC (Sebban *et al.*, 2002) and this set could be useful as a preliminary 'fast lane' screen. Spoligotyping has been found to be congruent with other molecular techniques, such as multilocus sequence typing (Baker *et al.*, 2004), single nucleotide polymorphism (SNPs) and large sequence polymorphism analysis (Filliol *et al.*, 2006; Gagneux & Small, 2007; Gutacker *et al.*, 2006). However, for more in-depth molecular epidemiological analysis, higher discriminatory genotyping techniques are necessary. These rely on mini-satellite type sequences known as MIRUs (mycobacterial interspaced repeat units) or VNTRs (variable number of tandem repeats). A 24 locus VNTR typing scheme was adopted widely (Supply *et al.*, 2006); it was found to be much more discriminatory than spoligotyping, but not fully concordant, hence there is still a need to perform the two techniques to obtain optimal discrimination. Thus, as

spoligotypes are considered both as phylogenetic and epidemiological markers (Asiimwe *et al.*, 2008; Sola *et al.*, 2001; Tanaka & Francis, 2006; Valcheva *et al.*, 2008), extensive spoligotype databases have proved useful in determining and cataloguing worldwide MTC genetic diversity (Brudey *et al.*, 2006). However, concerns have been raised by some investigators about the potential consequences of suboptimal spoligotyping results, or interpretations of results, on the quality of international databases.

Multiplexing is an efficient way to increase the throughput and efficiency of genotyping results and many recent techniques, such as multiple ligation-mediated assays, are showing promise in this area (Bergval *et al.*, 2008). A DNA chip-based spoligotyping method for the strain identification of *M. tuberculosis* isolates has also been recently developed (Song *et al.*, 2007). All these techniques are particularly well suited to SNP analysis and may be used to assess between 100 to 1000 probes in a single tube (Hardenbol *et al.*, 2003). However, at the lower range of the analysis (50–200 probes), macro-arrays and more recently the suspension microarray format are proving sufficiently powerful and economical genotyping methods (Kaufhold *et al.*, 1994; Goguet de la Salmoniere *et al.*, 2004; Dunbar *et al.*, 2003).

A single study by the Centers for Disease Control and Prevention, USA, evaluated the possibility of transferring spoligotyping to a multiplex microbead-based suspension array format (Cowan *et al.*, 2004). However, economical restrictions have meant that this has not yet been widely implemented. Yet as the cost of medical diagnosis devices drop, and newer and even more innovative systems will be developed, this versatile microbead suspension format could prove useful even in resource-limited settings because: (i) it is a simple and powerful multianalyte system, (ii) the use of the microspheres is easy with a 96-well microplate format, (iii) it is flexible and has the opportunity to include new or customized probes, (iv) it has high throughput, (v) it is a time-saving technique, (vi) it increases the quality of results. Such a molecular diagnostic and genotyping device could be fully dedicated to public health-oriented MTC genotyping projects and would provide effective support for epidemiological studies in national TB reference centres, and other TB diagnostic laboratories that have the capacity for mycobacterial cultures and DNA preparation.

In this article, we describe the implementation of and improvements to the spoligotyping technique in different formats, including suggestions for an increased discrimination format or simplifications to create a reduced cost-effective form suitable for a wide range of settings.

METHODS

Chemicals, buffers, microbeads and oligonucleotides. *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide (EDC) was used for

oligonucleotide-microbead coupling (Pierce Chemical) in 0.1 M MES buffer solution, pH 4.5 (Sigma-Aldrich). Hybridization solutions of $1.5\times$ and $1\times$ TMAC (tetra-methyl ammonium chloride) were prepared using 5 M TMAC (Sigma-Aldrich), 20 % sarkosyl (sodium lauroylsarcosinate), Tris/HCl and EDTA, and were prepared as described by Cowan *et al.* (2004). Oligonucleotides (25mers) were modified (van Embden *et al.*, 2000) from described sequences (Kamerbeek *et al.*, 1997) (Table 1). Panel 1 used a standardized 43 spacer format (Sebban *et al.*, 2002). Panel 2 (25 more spacers) was chosen to improve the discrimination for the MTC *M. africanum* subspecies and for the East African–Indian (EAI) clade (van Embden *et al.*, 2000). The 10 spacers of panel 3 included spacer number 7, 9, 18, 22, 29, 31, 32, 33, 34 and 43 (Kamerbeek nomenclature), and were chosen based on previous bioinformatics results (Sebban *et al.*, 2002). Primers 5'-biotinylated DRa (GGTTTTGGGTCTGACGAC) and DRb (CCGAGAGGGGACGGAAC) were used for the spoligotyping PCR. All capture oligonucleotides (Eurogentec) were supplied with a 5' terminal amino group modification using a 12 carbon spacer linker to increase the free space between the microspheres and the oligonucleotides for a better hybridization.

Microbead coupling. Luminex xMap microbeads (L100-C101 to L100-C168) were resuspended by vortexing (20 s) and sonication (20 s). Microbeads (2.5 million) were transferred to an Eppendorf Protein LoBind tube and pelleted at 8000 g for 1–2 min. The supernatant was discarded and the pellet was resuspended in 25 μ l 0.1 M MES buffer (pH=5.4). Oligonucleotides were suspended in ultrapure water at 1 mM and 100 pmol oligonucleotide capture probe previously diluted to 0.1 mM was added to the microbeads. A fresh EDC solution (10 mg ml⁻¹) was prepared and 1.25 μ l was added to each Eppendorf tube. Incubation was carried out in the dark at room temperature for 30 min. Subsequently 1.25 μ l EDC was added to each tube and incubation at room temperature was carried out in the dark for a further 30 min. The microbeads were then washed with 0.5 ml Tween 20 (0.02 %) followed by 0.5 ml SDS (0.1 %), and then resuspended in 50 μ l ($1\times$) TE (10 mM Tris, 1 mM EDTA, pH 8) with vortexing and sonication. Coupled microbeads were stored in the dark at 4 °C until required. A mixture of appropriate beads, depending on the assay, was prepared by combining equal volumes of each selected coupled set. Microbead concentration was checked by haemocytometer. Bead coupling efficiency was checked for the first designed capture probe using hybridization with a biotinylated-antisense detection probe followed by the multianalyte dual-laser detection.

Spoligotyping PCR protocol. PCR (25 μ l total) was performed for 20 cycles (Kamerbeek *et al.*, 1997) using thermolysates as samples. This was expected to generate 10^6 amplification copies of each spacer for each tube.

Microbead-based multianalyte two-laser system: principle, hardware and software. The system (Luminex 200) is made up of three components: an analyzer/reader, a microplate handler and a sheath-fluid delivery component. The software used was Luminex version 2.3. The device uses oligonucleotide-coupled microspheres (diameter 5.4 μ m) as support, with oligonucleotide capture probes that have been previously covalently linked to the microspheres. If, within a PCR product, the target of the capture probe is present, it will hybridize to the microbead surface. Each set of microbeads is characterized by a combination of two dyes (one red and one near infrared) so that each set has a unique spectral address. Beads are presented individually to detection lasers through a quartz cuvette. A 635 nm red laser excites the two dyes in each bead and identifies which set it belongs to. A 532 nm green laser is the reporter laser that excites the fluorochromes bound to the PCR product hybridized to the beads and provides quantification.

Hybridization. A defined microsphere mixture was prepared by addition of each set of stocked beads in $1.5\times$ TMAC hybridization buffer to a final concentration of approximately 150 microspheres μ l⁻¹ for each set. For hybridization, we added 2 μ l PCR product into 33 μ l microsphere mixture. After a denaturation at 95 °C for 5 min, incubation was done at 52 °C for 15 min. Plates were centrifuged at 2250 g for 3 min and half of the supernatant (30 μ l approximately) was replaced with the same amount of $1\times$ TE. A reporter mix (25 μ l) was prepared by adding 4 μ l streptavidin-R-phycoerythrin (1 mg ml⁻¹) (Molecular Probes) to 996 μ l $1\times$ TMAC hybridization buffer in the plates. After a final incubation again at 52 °C for 5 min, mixtures were analysed by the microplate reader as suggested in the user's manual.

Data analysis. The distribution of the values obtained for each spacer was determined, as described by Cowan *et al.* (2004), as bimodal. A cut-off value was initially chosen by eye for each spacer. A mean negative value and mean positive value were then calculated and the signal to noise ratio (S:N) derived from the ratio of the two means. We also extrapolated a mean control value and derived a simple calculation linking the cut-off value and the mean negatives.

Statistical analysis. The Hunter and Gaston index (HGDI) was computed to show the superiority of the 68 spacers versus 43 spacers technique for the EAI and *M. bovis* clades (Hunter & Gaston, 1988).

Mycobacterial isolates, origin, DNA extraction. Two independent sets of 50 DNA extractions were obtained from MTC clinical isolates from the Bichat Claude Bernard Hospital, Paris (R. Ruimy, unpublished results). Both sets had been previously genotyped by spoligotyping using commercial membranes (Isogen Biosciences). One set was representative of the overall genetic diversity of *M. tuberculosis* found in Paris; the second was selected to include more *M. africanum* and EAI isolates, and other less defined principal genetic group (PGG) 1 and PGG 2 organisms (Sreevatsan *et al.*, 1997) based on their spoligotyping signatures determined by Soini *et al.* (2000). A further set of DNA samples extracted from a set of bovine clinical isolates from the French Food Safety Agency Laboratory for bovine tuberculosis (Agence Française de Sécurité Sanitaire des Aliments, Maisons-Alfort, France) ($n=234$) and a last set obtained from human clinical isolates from the Institut Pasteur du Cambodge ($n=118$) were also tested using the 68 spacer set format. Ten more DNA samples extracted from as many clinical isolates from a previous molecular epidemiological study performed in Albania (Tafaj *et al.*, 2009) were also included on a reduced spacer panel (panel 3). All DNAs were extracted using a thermolysate method (three cycles of freeze–thaw followed by centrifugation and removal of cell pellets). DNA samples were stored at -20 °C until use.

RESULTS AND DISCUSSION

Development of the classical 43 spacers format assay (panel 1)

We initially tested the microbead-based multianalyte dual-laser device with a standard spoligotyping format by coupling 43 oligonucleotide spacers to microbeads. A total of 50 MTC DNAs from clinical isolates obtained from Bichat Claude Bernard Hospital, Paris, were blinded and genotyped using both microbead-based and membrane-based formats (Table 1). In the microbead format the distribution value of all experimental relative fluorescence intensity (RFI) data points showed an absence of overlap between negative and positive RFI values, independent of

Table 1. List of spacer oligonucleotide sequences of MTC used in this study (capture probes)

Genome spacer no.	43 spacer format no.	Oligonucleotide sequence (5'→3')
1		ATAGAGGGTCGCCGGCTCTGGATCA
2	1	CCTCATGCTTGGGCGACAGCTTTTG
3	2	CCGTGCTTCCAGTGATCGCCTTCTA
4	3	ACGTCATACGCCGACCAATCATCAG
5		TTTTCTGACCACTTGTGCGGGATTA
6		CGTCGTCATTTCCGGCTTCAATTC
7		GAGGAGAGCGAGTACTCGGGGCTGC
8		CGTGAAACCGCCCCAGCCTCGCCG
9		ACTCGGAATCCCATGTGCTGACAGC
10		TCGACACCCGCTCTAGTTGACTTCC
11		GTGAGCAACGGCGGGCGCAACCTGG
12	4	ATATCTGCTGCCCGCCGGGAGAT
13	5	GACCATCATTGCCATTCCCTCTCCC
14	6	GGTGTGATGCGGATGGTCGGCTCGG
15	7	CTGAATAACGCGCAGTGAATTCG
16		CGAGTCCCGTCAGCGTCGTAAATC
17		GCGCCGGCCCGCGCGGATGACTCCG
18	8	CATGGACCCGGGCGAGCTGCAGATG
19	9	TAACTGGCTTGGCGCTGATCCTGGT
20	10	TTGACCTCGCCAGGAGAGAAGATCA
21	11	TCGATGTCGATGTCCCAATCGTCGA
22	12	ACCGCAGACGGCACGATTGAGACAA
23	13	AGCATCGCTGATGCGGTCCAGCTCG
24	14	CCGCCTGCTGGGTGAGACGTGCTCG
25	15	GATCAGCGACCACCGCACCCCTGTCA
26	16	CTTCAGCACCACCATCATCCGGCGC
27	17	GGATTTCGTGATCTCTTCCCGCGGAT
28	18	TGCCCGGGCGTTTAGCGATCACAAC
29	19	AAATACAGGCTCCACGACACGACCA
30	20	GGTTGCCCCGCGCCCTTTCCAGCC
31	21	TCAGACAGGTTTCGCGTCGATCAAGT
32	22	GACCAAATAGGTATCGGCGTGTTCA
33	23	GACATGACGGCGGTGCCGCACTTGA
34	24	AAGTCACCTCGCCCACACCGTCGAA
35	25	TCCGTACGCTCGAAACGCTTCCAAC
36	26	CGAAATCCAGCACCACATCCGCAGC
37	27	CGGAACTCGTCCACAGTCCCCCTT
38	28	CGTGGATGGCGGATGCGTTGTGCGC
39	29	GACGATGGCCAGTAAATCGGCGTGG
40	30	CGCCATCTGTGCCTCATACAGGTCC
41	31	GGAGCTTTCGGCTTCTATCAGGTA
42	32	ATGGTGGGACATGGACGAGCGCGAC
43	33	CGCAGAATCGCACCGGGTGCGGGAG
44	34	ATGGCACGGCAGGCGTGGCTAGGGG
45		GTGCGCCGTCGCCGTAAGTGCCCCA
46	35	TTTCGACGACAATTCGTTGACCACG
47	36	GTTACCGCTGGCGCGCATCATTTCAT
48		CGTGCACATGCCGTGGCTCAGGGGT
49		CATGCAGCATGCCGTCCCGTTTTT
50		TGCTCTTGAGCAACGCCATCATCCG
51	37	GGCAAGTTGGCGCTGGGGTCTGAGT
52	38	GCGAGGAACCGTCCCACCTGGGCCT
53	39	GGAAACGCAGCACCAGCCTGACAAT
54		GCACTGCAACCCGGAATTCCTGCAC
55		CCATATCGGGGACGGCGACGCTGCG
56		ACGCGTCGTGCCATCAGTCAGCGTC

Table 1. cont.

Genome spacer no.	43 spacer format no.	Oligonucleotide sequence (5'→3')
57		AACACTTTTTTTGAGCGTGGCGCGG
58		GGGCATCGATCATGAGAGTTGCGTT
59		CTGGCGACGATTTTCGCTGTTGTGG
60		AGCACCTCCCTTGACAATCCGGCAG
61		GGCCTAAGGGTGCTGACTTCGCCTG
62	40	ACGACGAGCAGCGGCATACAGAGCC
63	41	TTGCATCCACTCGTCGCCGACACGG
64	42	TGGTAATTGCGTCACGGCGCGCCTG
65	43	ACCATCCGACGCAGGCACCGAAGTC
66		CACACCACAGCCACGCTACTGCTCC
67		ACACCGCCGATGACAGCTATGTCCG
68		CTTCGCGCGGTGTTTCGGCCGTGCC

the probe detected. Negative RFI values varied between spacers and were significantly high for spacer 29 and low for spacer 15. We observed only 2/2150 discrepant data points among 50 samples tested between the two techniques. These were linked to spacer 15 for the samples 2007-2 and 2007-13, and called negative by membrane-based and positive by microbead-based spoligotyping. As the values of spacer 15 were above the cut-off value for the microbead-based assay, we concluded that for these samples the spacer 15 was present. A repeat run using a fresh membrane confirmed the presence of the spacer 15 in the two samples. These discrepant results were not wholly unexpected as this spacer has previously been shown to give a low reproducibility in membrane-based methods probably as a result of genomic structural configurations (van Embden *et al.*, 2000). Indeed, this problem could possibly be explained by a published study showing that the DVR26 [van Embden *et al.* (2000) numbering, or DR16 in the Kamerbeek *et al.* (1997) numbering] contains a 4 bp deletion in the adjacent DR (van Embden *et al.*, 2000) and that the deletion in DR16 could decrease the relative amount of spacer 15 amplification, which ultimately lead to a lower S:N ratio. Nevertheless, despite the lower signal we could clearly distinguish between positive and negative values confirming the power of the microbead-based technique. We also observed a low S:N for spacer 15 on beads as previously seen in membrane-based assays. The addition of a washing step increased S:N for all spacers (see experiment 2 compared to experiment 1 in Table 2). Additional typing experiments confirmed that the cut-off value remained constant with each set of beads and that the method was fully reproducible (H37Rv and *M. bovis* BCG, and a negative control, were systematically included).

Obtaining good results using the classical membrane-based spoligotyping assay critically depends on the quality of a membrane. Although theoretically reusable up to 10–20 times, commercial membranes tend to provide lesser quality results after 6–8 uses and this phenomenon was responsible of some of the discrepancies we observed initially. In this work, these instances were few and were all

solved after retesting using a new membrane. It is also clear that coupling efficiency (i.e. between the beads and the probes) may also influence the quality of the results. However, we did not experience problems in this regard. We nevertheless advise recalculation of cut-off values for individual batches of coupled beads.

RFI values of all spacers for some samples were lower than those obtained for other samples. This result was most likely caused by the use of inferior quality DNA leading to suboptimal PCR amplification as in this study we used DNA extracted from thermolysates, which are known to often generate low-quality samples. Yet even using these suboptimal extractions the discriminative power resulting from this procedure was high, suggesting that even with unpurified (low DNA concentration) thermolysates the method can be considered as highly robust. Recent results on a much larger set of samples (more than 4000) showed that cetyl-trimethylammonium-extracted DNA gave higher S:N than thermolysate DNA, further simplifying analysis (data not shown).

Extended spoligotyping using 25 more spacers (panel 2)

A subset of PGG 1 isolates, as defined by Sreevatsan *et al.* (1997) and Soini *et al.* (2000), from the first set of 50 DNAs, combined with a further set that included a majority of PGG 1 clinical isolates ($n=402$ total), was chosen to test an improved spoligotyping format. Indeed, published studies have suggested that PGG 1 spoligotyping could be improved by the addition of spacers; this is not the case with 'Haarlem' clinical isolates (PGG 2) (van Embden *et al.*, 2000). Positive and negative values could be clearly distinguished without overlap, with the apparent exception of spacer 48 for which five isolates belonging to the Central Asian (CAS) clade provided trimodal signals (Supplementary Fig. S1 available with the online journal). This observation suggested intermediate hybridization signals; however, sequencing did not confirm the presence of this spacer. This result remains for the time-being

Table 2. S : N in microbead-based spoligotyping

Spacer no.	Mean positive value		Mean negative value		Ratio positive/negative	
	Expt 1	Expt 2	Expt 1	Expt 2	Expt 1	Expt 2
1	167.24	284.34	11.96	11.73	13.99	24.25
2	161.29	235.01	12.04	11.31	13.4	20.78
3	760.28	1018.26	11.63	11.77	65.4	86.54
4	500.49	819.09	16.8	19.32	29.79	42.39
5	352.63	704.86	11	13.03	32.06	54.09
6	951.49	1277.66	41.97	40.75	22.67	31.35
7	338.79	538.73	9.82	10.89	34.51	49.47
8	318.95	473.64	11.7	17.21	27.26	27.52
9	418.68	642.67	11.68	10.71	35.84	59.98
10	680.99	1083.87	11.14	12.53	61.14	86.51
11	453.21	795.24	13.29	14.66	34.09	54.26
12	274.63	482.34	12.22	11.82	22.47	40.79
13	731.68	1126.56	12.57	12.65	58.2	89.06
14	600.85	930.69	13.11	12.88	45.83	72.25
15	69.54	129.7	10.41	7.84	6.68	16.54
16	633.79	931.32	23.97	22.62	26.44	41.18
17	291.74	362.74	11.67	10.47	25.01	34.66
18	346.54	635.74	10.03	11.23	34.55	56.59
19	527.32	643.43	12.46	11.08	42.31	58.09
20	447.5	663.24	14.13	14.2	31.68	46.71
21	240.61	324.52	16.59	17.72	14.5	18.31
22	610.98	835.27	11.35	12.34	53.83	67.68
23	249.54	417.09	11.55	12.09	21.6	34.5
24	100	227.17	11.2	11.54	8.93	19.69
25	1075.07	1779.39	11.9	10.75	90.34	165.52
26	434.72	907.52	15.05	19.83	28.89	45.76
27	463.85	670.16	14.23	9.86	32.59	67.99
28	447.45	783.62	13.54	10.46	33.06	74.9
29	815.04	1329.41	50.69	71.79	16.08	18.52
30	980.74	1476.71	13.25	11.47	74.02	128.74
31	510.83	763.22	13.05	14.5	39.14	52.64
32	528.83	704.21	12.73	10.89	41.53	64.65
33	141.13	170.63	10.2	13.46	13.84	12.68
34	792.5	704	11.85	12.12	66.88	58.07
35	663.88	944.96	11.85	9.06	56.02	104.25
36	228.92	476.88	19.35	11.59	11.83	41.15
37	474.3	773.67	27.21	30.08	17.43	25.72
38	401.23	804.47	16.42	11.5	24.44	69.95
39	322.51	608.45	13.78	12.44	23.41	48.92
40	326.17	655.08	11.42	8.17	28.55	80.21
41	645.46	884.4	13.79	13.58	46.82	65.11
42	347.11	622.72	15.38	7.33	22.58	84.92
43	209.97	384.95	17.33	20	12.11	19.25

unexplained. Among the 402 DNAs assessed using the 68 spacer spoligotyping, 351 DNAs from essentially two new origins (Agence Française de Sécurité Sanitaire des Aliments for *M. bovis* and Institut Pasteur du Cambodge for the EAI clade) were clustered into seven clades [*M. africanum*, *M. bovis*, Beijing, EAI, CAS, Latin-American and Mediterranean (LAM) and T (all other modern isolates)]. Results provided a greater discrimination of *M. africanum*, EAI and *M. bovis* than had been described by Sreevatsan *et al.* (1997). The most significant results are

shown in Table 3 (clustering results for 351 clinical isolates) and in Supplementary Table S1 (available with the online journal) (unique patterns description). A computation of the HGDI on the samples belonging to the *M. bovis* and to the EAI clades, when using the 68 versus 43 spacers format, shows an increase of HGDI from 0.83 to 0.85 (for *M. bovis*) and from 0.93 to 0.96 (for EAI). According to the current available requirements (HGDI >0.95 for a molecular marker to be used in epidemiology), spoligotyping under the 68 spacers format would then

Table 3. Comparison of cluster number obtained with the 43 spacers format as compared with the 68 spacers format

	No. of isolates	No. of clusters	
		43 format	68 format
<i>M. africanum</i>	13	3	5
<i>M. bovis</i>	224	13	19
EAI	63	12	21
Beijing	30	1	2
CAS	7	1	2
LAM	8	1	1
T	6	2	2
Total	351	33	52

T, Modern isolates (all other undifferentiated isolates).

become a suitable method to run population-based molecular epidemiological studies in South-East Asia where the EAI clade is prevalent. Our sampling on *M. africanum* is currently too low to conclude on the superiority of the 68 versus 43 spoligotyping format.

The introduction of 25 additional spacers, derived from *M. bovis* and *M. tuberculosis* sequences (Beggs *et al.*, 1996), provided the greatest discriminatory power so far achieved within the MTC complex, with the exception of the '*M. canettii*' subspecies whose typing would require 26 more spacers (van Embden *et al.*, 2000). In previous works, we and other investigators have suggested that the introduction of 25 new spacers could improve genotyping, especially for large animal and human-adapted MTC PGG 1 subspecies (Brudey *et al.*, 2004; van der Zanden *et al.*, 2002; Javed *et al.*, 2007). In this study, we confirm that this improvement is significant for the study of the clinical isolates belonging to the EAI clade that is prevalent in South-East Asia.

Development of a simplified 10 spacers format assay (panel 3)

Other genetic diversity studies on MTC using IS6110-RFLP, VNTR, SNPs or region of differences diversity studies have identified a phylogeographical population structure in this monomorphic bacterial species (Baker *et al.*, 2004; Dos Vultos *et al.*, 2008; Gagneux & Small, 2007; Gutacker *et al.*, 2006; Kremer *et al.*, 1999; Sola *et al.*, 2003; van Embden *et al.*, 1993). A previous bioinformatics study has shown that a limited set of spoligotype spacers ($n=10$, spacers number 7, 9, 18, 22, 29, 31, 32, 33, 34 and 43) could be sufficient to discriminate between the most prevalent MTC phylogeographical types (Sebban *et al.*, 2002). Using these spacers we also implemented a simplified microbead-based spoligotyping assay format, which could be used as a relatively inexpensive 'first screen' genotyping technique for MTC. This assay could also be used as an external quality control for laboratories still working with mem-

brane-based spoligotyping, or as a first step toward further multiplexing. Other markers such as rifampicin-resistance mutations (*rpoB*) could be included in these assays. The use of a 10 spacers assay (panel 3) would naturally be cheaper, and still remain an informative assay that could be suited to resource-limited countries. Improving the capacity and transfer of rapid molecular diagnostics methods, whilst decreasing prices of emerging technologies, are undoubtedly desirable objectives towards good public-health strategies and achieving a more sustainable decrease in preventable deaths caused by poverty-related infectious diseases.

Comparison of the cost/output ratio of the microsphere-based versus the membrane-based assay for spoligotyping

Spoligotyping, either using a microbead-based multi-analyte genotyping platform or membrane-based, is a rapid diagnostic test most suitable for studying TB disease in high-burden countries. One important parameter for technology transfer, especially for resource-poor or middle-income countries, will be the cost/output ratio needed to introduce the new technique (microbead-based high throughput) compared to the established one (membrane-based low throughput). We thus evaluated the relative costs of material, reagents and staff required to genotype 5000 samples by the two techniques (based on a 'standardized' 43 spacers format). The results (Table 4) showed that running spoligotyping through a high throughput multianalyte hybridization device would be 30–40% more expensive than a standard membrane-based technique. The throughput is at least 4–8 times higher using the microbead-based rather than the membrane-based approach. It also improves the global quality in TB genotyping by (i) introducing standardized cut-off values thus improving standardization of spoligotyping, (ii) theoretically helping to discover new SNPs in CRISPR loci, (iii) increasing the potential use of newly discovered spacers (Brudey *et al.*, 2004; van der Zanden *et al.*, 2002; van Embden *et al.*, 2000). In countries where TB outbreaks are largely linked to *M. africanum*, the EAI clade or even the CAS clade, i.e. within Africa and the Indian subcontinent and Asia (Ahmed & Leblebicioglu, 2006), this is particularly relevant.

The microbead-based format allows a digital-numerical data output less prone to interpretation errors. This output is also more amenable to being managed through large international databases that can be curated for genotyping errors.

The advent of molecular epidemiological techniques using polymorphic repeated DNA markers has improved ways to investigate TB outbreaks (van Soolingen, 2001). In particular they have been used to demonstrate to clinicians, epidemiologists and public health investigators that recent TB transmission rates were higher than previously estimated (Alland *et al.*, 1994), concomitantly to the decision by the World Health Organization in 1994 to

Table 4. Economic analysis of automated versus manual spoligotyping

Parameter	Multiplexed microbead-based (automated)	Membrane-based (manual)
Device cost	Around 60 000 Euros (Luminex 200)	1200 Euros (Miniblotter, Immunetics)
Spoligotyping kits (5000 detection) (Ocimum Biosolutions)	–	3300 Euros
Full set of 43 microbeads (5000 detection), control and calibration	8600 Euros	–
Oligonucleotides costs	1200 Euros	None
PCR product volume required	5 µl	25 µl
Working time for 500 analyses	2 days	11 days
Flexibility	Yes	More or less
Operator dependent	No; however, needs a well trained technician	Yes, needs a carefully trained technician
Result interpretation	No (automated)	Yes
Database comparison required	Yes	Yes for full usefulness of information produced
Open or closed system	Open	Open (investigator may choose to build its own kit)
Development	Yes (could be applied to new SNPs)	Yes (other probes and targets can be tested)
Possibility of improved technique	Yes	Yes
Data production technician time for theoretical 5000 assays	25 days	110 days
Data reading and control technician time for 5000 assays	25 h	230 h
Total working costs (time)	200 h	1000 h
Total working costs (in Euros) if unitary working cost is set to 50 Euros h ⁻¹	10 000 Euros	50 000 Euros
Total cost analysis	80 000/5000=>16 Euros*	55 000/5000=11 Euros
Pros	High throughput, results generated directly in Excel format. Ideally suited for large centralized analysis production facilities in new emerging high-burden countries like Brazil, Russia, India, China. Can be linked to a LIMS	Robust, simple technology; however, lot of error sources, results are interpretative and require expertise, spacer 15 doubtful
Cons	More expensive at present, should be done in limited number of level 3 or 4 mycobacteriology laboratories	Less discriminative than MLVA (multilocus VNTR analysis)

LIMS, Laboratory Information Management System.

*Cost of 43 spacers-based technology.

declare TB a global emergency (Nakajima, 1993). The recent global expansion of MDR-TB and XDR-TB has renewed calls for new public health-oriented research programmes (Dye & Williams, 2008) and the introduction of more efficient and economical genotyping methods worldwide in mycobacteriology laboratories.

High-throughput multiplexed hybridization techniques are being increasingly developed for molecular diagnostics, molecular epidemiology (Dunbar *et al.*, 2003; Jiang *et al.*, 2006) and environmental studies (Baums *et al.*, 2007). The recent transfer of the 'spoligotyping' technique to a microbead format (Cowan *et al.*, 2004) provided an opportunity for incorporating this type of assay into our

laboratory and promoted a new molecular epidemiological study. Indeed the repetitive and heterogeneous nature of CRISPR loci makes them excellent targets for the development of new and powerful genotyping techniques, whilst major diagnostic genomic targets used for bacterial species identification include probes for 23S and/or 16S rRNA genes (Dunbar, 2006; Mokrousov *et al.*, 2007).

We found the multiplex microbead-based spoligotyping technique to be versatile and easily implemented and it can be adapted (classical 43 spacers format for Europe, improved 68 spacers format for Asia, simplified 10 spacers format for first-screen genotyping in resource-limited settings). It is well established that molecular epidemiolo-

gical studies in high-prevalence settings are more difficult to interpret and to run than in low-prevalence settings, requiring more analysis of samples and implementation of complex genotyping schemes. This strongly argues in favour of the widespread adoption of high-throughput methods (Hanekom *et al.*, 2008).

Further developments in this system will include study of other epidemiologically important clades (such as Beijing) and development of links to SNP-based typing methods allowing clinical, phylogenetic and epidemiological improvements in rapid molecular methods with the aim of providing simultaneous strain typing and molecular antibiogrammes to further aid in control regimens aimed against MDR and XDR-TB.

In conclusion the suspension microarray spoligotyping assay was efficiently implemented, improved and represents a viable and practical alternative to MIRU-VNTR as a first screen for large population-based systematic genotyping programs. An international multicentred project is currently ongoing to confirm our results on a larger scale.

ACKNOWLEDGEMENTS

This work was made possible thanks to a research grant from the Fondation bioMérieux, Lyon, France, and Dr Guy Vernet is warmly acknowledged. M. François Topin (Luminex Europe, Netherlands) is warmly acknowledged for his support. The IGEPE team was created thanks to the University Paris-Sud 11 and to the Director of the Institut of Genetics and Microbiology (IGM) (UMR8621), Dr Monique Bolotin-Fukuhara. J.Z. holds a PhD fellowship from the Presidency of the University of Paris-Sud 11. Dr Gilles Vergnaud, Head of the GPMS Research Team at IGM, is also warmly acknowledged for his help. Lastly, we are grateful to Professor B. Gicquel, Head of the Mycobacterial Genetics Unit at the Institut Pasteur, for her permanent support and helpful discussions.

REFERENCES

- Ahmed, N. & Leblebicioglu, H. (2006). India's 'gold mine' of ancestral bacilli and the looming TB-HIV pandemic. *Ann Clin Microbiol Antimicrob* 5, 31.
- Alland, D., Kalkut, G. E., Moss, A. R., Adam, R. A. M., Hahn, J. A., Bosworth, W., Drucker, E. & Bloom, B. R. (1994). Transmission of tuberculosis in New York City: an analysis by DNA fingerprinting and conventional epidemiologic methods. *N Engl J Med* 330, 1710–1716.
- Asiimwe, B. B., Koivula, T., Kallenius, G., Huard, R. C., Ghebremichael, S., Asiimwe, J. & Joloba, M. L. (2008). *Mycobacterium tuberculosis* Uganda genotype is the predominant cause of TB in Kampala, Uganda. *Int J Tuberc Lung Dis* 12, 386–391.
- Baker, L., Brown, T., Maiden, M. C. & Drobniowski, F. (2004). Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg Infect Dis* 10, 1568–1577.
- Baums, I. B., Goodwin, K. D., Kiesling, T., Wanless, D., Diaz, M. R. & Fell, J. W. (2007). Luminex detection of fecal indicators in river samples, marine recreational water, and beach sand. *Mar Pollut Bull* 54, 521–536.
- Beggs, M. L., Cave, M. D., Marlowe, C., Cloney, L., Duck, P. & Eisenach, K. D. (1996). Characterization of *Mycobacterium tuberculosis* complex direct repeat sequence for use in cycling probe reaction. *J Clin Microbiol* 34, 2985–2989.
- Bergval, I. L., Vijzelaar, R. N., Dalla Costa, E. R., Schuitema, A. R., Oskam, L., Kritski, A. L., Klatser, P. R. & Anthony, R. M. (2008). Development of multiplex assay for rapid characterization of *Mycobacterium tuberculosis*. *J Clin Microbiol* 46, 689–699.
- Brudey, K., Gutierrez, M. C., Vincent, V., Parsons, L. M., Salfinger, M., Rastogi, N. & Sola, C. (2004). *Mycobacterium africanum* genotyping using novel spacer oligonucleotides in the direct repeat locus. *J Clin Microbiol* 42, 5053–5057.
- Brudey, K., Driscoll, J., Rigouts, L., Prodinger, W. M., Gori, A., Al-Hajj, S. A. M., Allix, C., Aristimuno, L., Arora, J. & other authors (2006). *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics, and epidemiology. *BMC Microbiol* 6, 23.
- Cowan, L. S., Diem, L., Brake, M. C. & Crawford, J. T. (2004). Transfer of a *Mycobacterium tuberculosis* genotyping method, spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. *J Clin Microbiol* 42, 474–477.
- Dos Vultos, T., Mestre, O., Rauzier, J., Golec, M., Rastogi, N., Rasolof, V., Tonjum, T., Sola, C., Matic, I. & other authors (2008). Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis*. *PLoS One* 3, e1538.
- Drobniowski, F. A., Caws, M., Gibson, A. & Young, D. (2003). Modern laboratory diagnosis of tuberculosis. *Lancet Infect Dis* 3, 141–147.
- Dunbar, S. A. (2006). Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin Chim Acta* 363, 71–82.
- Dunbar, S. A., Vander Zee, C. A., Oliver, K. G., Karem, K. L. & Jacobson, J. W. (2003). Quantitative, multiplexed detection of bacterial pathogens: DNA and protein applications of the Luminex LabMAP system. *J Microbiol Methods* 53, 245–252.
- Dye, C. & Williams, B. G. (2008). Eliminating human tuberculosis in the twenty-first century. *J R Soc Interface* 5, 653–662.
- Fillioli, I., Motiwala, A. S., Cavatore, M., Qi, W., Hernando Hazbon, M., Bobadilla Del Valle, M., Fyfe, J., Garcia-Garcia, L., Rastogi, N. & other authors (2006). Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *J Bacteriol* 188, 759–772.
- Gagneux, S. & Small, P. M. (2007). Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis* 7, 328–337.
- Goguet de la Salmoniere, Y. O., Kim, C. C., Tsolaki, A. G., Pym, A. S., Siegrist, M. S. & Small, P. M. (2004). High-throughput method for detecting genomic-deletion polymorphisms. *J Clin Microbiol* 42, 2913–2918.
- Groenen, P. M., Bunschoten, A. E., van Soolingen, D. & van Embden, J. D. (1993). Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol Microbiol* 10, 1057–1065.
- Gutacker, M. M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B. N., Graviss, E. A. & Musser, J. M. (2006). Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. *J Infect Dis* 193, 121–128.
- Hanekom, M., van der Spuy, G. D., Gey van Pittius, N. C., McEvoy, C. R. E., Hoek, K. G. P., Ndabambi, S. L., Jordan, A. M., Victor, T. C., van Helden, P. D. & Warren, R. M. (2008). Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism

- genotyping for analysis of *Mycobacterium tuberculosis* Beijing strains in a setting of high incidence of tuberculosis. *J Clin Microbiol* **46**, 3338–3345.
- Hardenbol, P., Baner, J., Jain, M., Nilsson, M., Namsaraev, E. A., Karlin-Neumann, G. A., Fakhrai-Rad, H., Ronaghi, M., Willis, T. D. & other authors (2003).** Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol* **21**, 673–678.
- Hunter, P. R. & Gaston, M. A. (1988).** Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* **26**, 2465–2466.
- Jansen, R., van Embden, J. D. A., Gaastra, W. & Schouls, L. M. (2002).** Identification of a novel family of sequence repeats among prokaryotes. *OMICS* **6**, 23–33.
- Javed, M. T., Aranaz, A., de Juan, L., Bezos, J., Romero, B., Alvarez, J., Lozano, C., Mateos, A. & Dominguez, L. (2007).** Improvement of spoligotyping with additional spacer sequences for characterization of *Mycobacterium bovis* and *M. caprae* isolates from Spain. *Tuberculosis (Edinb)* **87**, 437–445.
- Jiang, H. L., Zhu, H. H., Zhou, L. F., Chen, F. & Chen, Z. (2006).** Genotyping of human papillomavirus in cervical lesions by L1 consensus PCR and the Luminex xMAP system. *J Med Microbiol* **55**, 715–720.
- Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R. & other authors (1997).** Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* **35**, 907–914.
- Kaufhold, A., Podbielski, A., Baumgarten, G., Blokpoel, M., Top, J. & Schouls, L. (1994).** Rapid typing of group A streptococci by the use of DNA amplification and non-radioactive allele-specific oligonucleotide probes. *FEMS Microbiol Lett* **119**, 19–26.
- Kremer, K., van Soolingen, D., Frothingham, R., Haas, W. H., Hermans, P. W., Martin, C., Palittapongarnpim, P., Plikaytis, B. B., Riley, L. W. & other authors (1999).** Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* **37**, 2607–2618.
- Manissero, D. & Fernandez de la Hoz, K. (2006).** Surveillance methods and case definition for extensively drug resistant TB (XDR-TB) and relevance to Europe: summary update. *Euro Surveill* **11**, E061103.
- Mokrousov, I., Limeschenko, E., Vyazovaya, A. & Narvskaya, O. (2007).** *Corynebacterium diphtheriae* spoligotyping based on combined use of two CRISPR loci. *Biotechnol J* **2**, 901–906.
- Nakajima, H. (1993).** Tuberculosis: a global emergency. *World Health Forum* **14**, 438.
- Rastogi, N., Legrand, E. & Sola, C. (2001).** The mycobacteria: an introduction to nomenclature and pathogenesis. *Rev Sci Tech* **20**, 21–54.
- Sebban, M., Mokrousov, I., Rastogi, N. & Sola, C. (2002).** A data-mining approach to spacer oligonucleotide typing of *Mycobacterium tuberculosis*. *Bioinformatics* **18**, 235–243.
- Smith, N. H., Kremer, K., Inwald, J., Dale, J., Driscoll, J. R., Gordon, S. V., van Soolingen, D., Glyn Hewinson, R. & Maynard Smith, J. (2006).** Ecotypes of the *Mycobacterium tuberculosis* complex. *J Theor Biol* **239**, 220–225.
- Soini, H., Pan, X., Amin, A., Graviss, E. A., Siddiqui, A. & Musser, J. M. (2000).** Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston, Texas, by spoligotyping. *J Clin Microbiol* **38**, 669–676.
- Sola, C., Filliol, I., Legrand, E., Mokrousov, I. & Rastogi, N. (2001).** *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR and DR-based spoligotyping suggests the existence of two new phylogeographical clades. *J Mol Evol* **53**, 680–689.
- 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.
- Song, E. J., Jeong, H. J., Lee, S. M., Kim, C. M., Song, E. S., Park, Y. K., Bai, G. H., Lee, E. Y. & Chang, C. L. (2007).** A DNA chip-based spoligotyping method for the strain identification of *Mycobacterium tuberculosis* isolates. *J Microbiol Methods* **68**, 430–433.
- Sorek, R., Kunin, V. & Hugenholz, P. (2008).** CRISPR – a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* **6**, 181–186.
- Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., Whittam, T. S. & Musser, J. M. (1997).** Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A* **94**, 9869–9874.
- Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rusch-Gerdes, S., Willery, E., Savine, E., de Haas, P., van Deutekom, H. & other authors (2006).** Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* **44**, 4498–4510.
- Tafaj, S., Zhang, J., Hauck, Y., Pourcel, C., Hafizi, H., Zoraqi, G. & Sola, C. (2009).** First insight into genetic diversity of the *Mycobacterium tuberculosis* complex in Albania obtained by multi-locus variable-number tandem-repeat analysis and spoligotyping reveals the presence of Beijing multidrug-resistant isolates. *J Clin Microbiol* **47**, 1581–1584.
- Tanaka, M. M. & Francis, A. R. (2006).** Detecting emerging strains of tuberculosis by using spoligotypes. *Proc Natl Acad Sci U S A* **103**, 15266–15271.
- Uplekar, M. & Lonroth, K. (2007).** MDR and XDR – the price of delaying engagement with all care providers for control of TB and TB/HIV. *Trop Med Int Health* **12**, 473–474.
- Valcheva, V., Mokrousov, I., Rastogi, N., Narvskaya, O. & Markova, N. (2008).** Molecular characterization of *Mycobacterium tuberculosis* isolates from different regions of Bulgaria. *J Clin Microbiol* **46**, 1014–1018.
- van der Zanden, A. G., Kremer, K., Schouls, L. M., Caimi, K., Cataldi, A., Hulleman, A., Nagelkerke, N. J. & van Soolingen, D. (2002).** Improvement of differentiation and interpretability of spoligotyping for *Mycobacterium tuberculosis* complex isolates by introduction of new spacer oligonucleotides. *J Clin Microbiol* **40**, 4628–4639.
- van Embden, J. D., Cave, M. D., Crawford, J. T., Dale, J. W., Eisenach, K. D., Gicquel, B., Hermans, P., Martin, C., McAdam, R. & other authors (1993).** Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* **31**, 406–409.
- van Embden, J. D. A., van Gorkom, T., Kremer, K., Jansen, R., van der Zeijst, B. A. M. & Schouls, L. M. (2000).** Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J Bacteriol* **182**, 2393–2401.
- van Soolingen, D. (2001).** Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J Intern Med* **249**, 1–26.