Identification of an IS6110 insertion site in *plcD*, the unique phospholipase C gene of *Mycobacterium bovis*

Cristina Viana-Niero,¹ Cesar Alejandro Rosales Rodriguez,² Fabiana Bigi,³ Marcos Santos Zanini,⁴ José Soares Ferreira-Neto,² Angel Cataldi³ and Sylvia Cardoso Leão¹

Correspondence

Sylvia Cardoso Leão
sylvia@ecb.epm.br

¹Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de São Paulo, Rua Botucatu 862 3º andar, São Paulo, CEP 04023-062, Brazil

²Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, Av. Professor Dr Orlando Marques de Paiva 87, São Paulo, CEP 05508-900, Brazil

³Instituto de Microbiología y Zoología Agrícola del CICVyA – Instituto Nacional de Tecnología Agropecuaria (INTA), Castelar, Buenos Aires, Argentina

⁴Centro de Ciências Agrárias, Universidade Federal do Espírito Santo, PO Box 16, Alegre, Espírito Santo, Brazil

The IS6110 repetitive element is present in multiple copies in most *Mycobacterium tuberculosis* complex bacteria, except for *Mycobacterium bovis* strains, which usually contain a single copy of IS6110 located on a 1·9 kb *PvuII* fragment of the direct repeat region. IS6110 transposition can disrupt coding regions and is a major force of genomic variation. In a previous work it was demonstrated that phospholipase C genes are preferential loci for IS6110 transposition in *M. tuberculosis* clinical strains. Bacterial phospholipase C enzymes participate in pathogenic mechanisms used by different organisms, and have been implicated in intracellular survival, cytolysis and cell-to-cell spread. Four phospholipase C genes (*plcA*, *plcB*, *plcC* and *plcD*) were detected in the genomes of *M. tuberculosis*, *Mycobacterium africanum*, *Mycobacterium microti* and '*Mycobacterium canettii*'. *M. bovis* and the vaccine strain *M. bovis* Bacillus Calmette–Guérin contain only the *plcD* gene. In the present work, the existence of IS6110 insertions within *plcD*, the unique phospholipase C gene of *M. bovis*, has been investigated by PCR, Southern blot hybridization and sequencing analysis. In 18 (7·3 %) of 245 isolates analysed, the *plcD* gene was interrupted by the insertion of one copy of IS6110, which in all cases was transposed in the same orientation and at the same position, 1 972 894, relative to the genome of *M. bovis* AF2122/97. These 18 isolates were distributed in 6 different spoligotype patterns and contained 4 to 8 IS6110 copies. In contrast, strains showing an intact *plcD* gene contained one (87 %), two (9·4 %) or three (2·4 %) IS6110 copies, and only a single isolate (1·2 %) had four IS6110 copies. The implications of *plcD* gene disruption in *M. bovis* have not been fully investigated, but no differences in the organ distribution of the disease were detected when animals infected with strains from the same spoligotype patterns bearing *plcD*::IS6110 and intact *plcD* were compared.

Received 7 October 2005

Accepted 21 December 2005

INTRODUCTION

Genomes of *Mycobacterium tuberculosis* complex bacteria contain four highly related genes encoding phospholipase C enzymes (Cole *et al.*, 1998). Three genes, *plcA*, *plcB* and *plcC*, are organized in tandem (*plcABC* locus). The fourth gene, *plcD*, is located in a different region. These four genes are present in most strains of *M. tuberculosis*, *Mycobacterium*

africanum, '*Mycobacterium canettii*' and *Mycobacterium microti*. Other complex members, the classical virulent *M. bovis* and the vaccine strain *M. bovis* Bacillus Calmette–Guérin, have only the *plcD* gene, because an 8·9 kb fragment containing the *plcABC* locus, the RD5 region, is consistently absent from strains of these species (Behr *et al.*, 1999; Brosch *et al.*, 2002; Gordon *et al.*, 1999). In the laboratory strain *M. tuberculosis* H37Rv, *plcD* is truncated by insertion of a copy of IS6110 and by deletion of a 7·9 kb fragment, designated RvD2 (Brosch *et al.*, 2002; Gordon *et al.*, 1999).

Abbreviations: DR, direct repeat; DVR, direct variable repeat.

Genetic variations in *plcABC* and *plcD* loci caused by insertions of IS6110 mobile genetic elements, and deletions originating from homologous recombination of two IS6110 elements, have been characterized in *M. tuberculosis* clinical isolates (Nguyen *et al.*, 2004; Talarico *et al.*, 2005; Vera-Cabrera *et al.*, 2001; Viana-Niero *et al.*, 2004; Warren *et al.*, 2000; Yesilkaya *et al.*, 2005), indicating that these domains represent preferential regions for the integration of this insertion element (Warren *et al.*, 2000; Yesilkaya *et al.*, 2005). The IS6110 element is present at variable copy numbers (0–25) and at different insertion sites in the chromosome of different *M. tuberculosis* strains. In contrast, *M. bovis* isolates harbour few copies of this element, frequently only a single copy located in the direct repeat (DR) region (Hermans *et al.*, 1991), suggesting that IS6110 transposition is a rare event in this species.

In this work, we investigated the existence of polymorphisms in the *plcD* gene, the unique phospholipase C gene of virulent *M. bovis*, and related these polymorphisms to IS6110 element insertions.

METHODS

Mycobacterial strains. A total of 245 *M. bovis* strains from 3 states in Brazil (São Paulo, Minas Gerais and Rio Grande do Sul) and 6 states in Argentina (Buenos Aires, Santa Fe, Entre Rios, Formosa, Rosario and La Pampa) were used in this study (Table 1). *M. bovis* isolates from São Paulo were obtained during a surveillance study of bovine tuberculosis. Federal and State Inspection Services collected samples in slaughterhouses from different regions of São Paulo State. Bacteria from lesions were cultivated in Löwenstein–Jensen and Stonebrink–Lesslie media (Centro Panamericano de Zoonosis, 1973). Two human isolates from São Paulo were obtained from a private clinical laboratory. *M. bovis* DNA of isolates from Minas Gerais, Rio Grande do Sul and Argentina were obtained from the collection of Centro de Investigaciones en Ciencias Veterinarias, Instituto Nacional de Tecnología Agropecuaria, Castelar, Argentina.

Table 1. Origin and source of isolates included in this study

Origin	Source	No. of isolates
Brazil		
São Paulo	Cattle	167
São Paulo	Human	2
Rio Grande do Sul	Cattle	13
Minas Gerais	Cattle	22
Argentina		
Santa Fe and Entre Rios	Human	20
Santa Fe, Formosa and Buenos Aires	Cattle	8
Buenos Aires	Cat	3
Rosario and Buenos Aires	Pig	7
La Pampa	Wild boar	2
Buenos Aires	Deer	1
Total		245

All isolates were previously identified as *M. bovis* by bacteriological criteria (Vestal, 1975; Wayne, 1984) and/or spoligotyping (Kamerbeek *et al.*, 1997). *M. bovis* AN5 was used as a control for PCR and hybridization experiments.

PCR and sequencing. The PCR of the *plcD* gene was set up using primers D1 5'-TTCGCGGCGAGAATTCCTAG-3' and D2 5'-TCCTGCGTGGGCATGATCTG-3' for amplification of a 1478 bp fragment, as described previously (Viana-Niero *et al.*, 2004) (Fig. 1a). Amplification of a 123 bp fragment from the IS6110 element was carried out using primers IS1 5'-CCTGCGAGCGTAGGCGTCGG-3' and IS2 5'-CTCGTCCAGCGCCGCTTCGG-3', as described by Eisenach *et al.* (1990). Amplicons were sequenced using primers IS3 5'-CCATCCG-CACCGCCCGCTCAC-3' and IS4 5'-CATCGCCGCTCTACCAGTAC-3' directed outwards from the IS6110 sequence in an automated ABI PRISM 377 sequencer (Perkin-Elmer) using the BigDye Terminator Ready Reaction Cycle Sequencing kit v 3.0 (Applied Biosystems). Sequences were compared using BLAST – basic local alignment tool (<http://www.ncbi.nlm.nih.gov/blast>).

Southern blotting and fingerprinting analysis. Southern blotting was performed as described by van Embden *et al.* (1993). Briefly, 2 µg DNA were digested using *PvuII* and subjected to electrophoresis in 0.8% agarose in 0.5 × TBE pH 8.0, at 2 V cm⁻¹. A mixture of *PvuII*-digested supercoiled DNA ladder (Invitrogen) and *HaeIII*-digested ϕX174 DNA (Invitrogen) was used as an internal DNA size marker. DNA was blotted onto nylon membranes (Hybond-N+; Amersham Biosciences), and probed, in different experiments, with probes complementary to the *plcD* gene and IS6110, obtained by PCR using primers D1/D2 or IS1/IS2 as described previously. Blots were also hybridized with a probe complementary to the DR region (5'-GTCGTCAGACCCAAAACCCG-GAGAGGGGACGGAAAC-3') (Hermans *et al.*, 1991), and with an internal size marker. All probes were covalently labelled with peroxidase by the glutaraldehyde method, and detected using the ECL Direct system (Amersham) according to the manufacturer's instructions. Membranes were exposed to X-ray films (X-OMAT; Kodak). Fingerprint autoradiograms were superimposed with internal marker autoradiograms for normalization, and were analysed using Bio-Numerics software version 4.0 (Applied Maths).

RESULTS

Analysis of *plcD* integrity by *plcD* PCR

All 245 isolates were tested by *plcD* PCR. A total of 41 isolates from Argentina and 186 isolates from Brazil (13 from Rio Grande do Sul, 22 from Minas Gerais and 151 from São Paulo) produced amplicons of 1478 bp. In contrast, 18 (7.3%) isolates from São Paulo produced PCR amplicons of approximately 3 kb, which were larger than expected. In all cases, reamplification of these large amplicons using primers IS1 and IS2, specific for the IS6110 sequence, generated 123 bp products, confirming the insertion of this element into *plcD* gene (*plcD*::IS6110).

Corroboration of IS6110 insertion in *plcD* by fingerprinting

DNA from 42 isolates from São Paulo was available for fingerprinting with the *plcD*-complementary probe. For 12 isolates with *plcD*::IS6110, hybridization with this probe revealed a four-banded pattern, with bands of approximately 900, 1150, 3900 and 4200 bp (Fig. 2a). For 30 isolates

Fig. 1. (a) Schematic representation of the *plcD* locus from *M. bovis* AF2122/97 (Garnier *et al.*, 2003). (b) Schematic representation of isolates with *plcD*::*IS6110*. Gene names are shown in (a). Small white arrows (not to scale) above the line represent primers D1 and D2. *PvuII* restriction sites are represented as small vertical arrows. The small black horizontal arrow in (b) represents the direction of transcription of *IS6110* transposase. a, b, c, d and e are fragments of 4200, 538, 3900, 1150 and 900 bp, respectively, detected by hybridization of Southern blots with a *plcD*-derived probe.

that generated 1478 bp *plcD* amplicons, hybridization patterns produced by the *plcD*-derived probe revealed 3 bands of 538, 3900 and 4200 bp, also observed with *M. bovis* AN5 control strain, confirming the presence of an intact *plcD* gene (Fig. 2a).

Hybridization experiments with the *IS6110* complementary probe (*IS6110* RFLP) were performed on the 42 isolates from São Paulo. For an additional 55 isolates from Minas Gerais, Rio Grande do Sul and Argentina, the *IS6110* copy number was estimated from previous experiments and publications (Zanini *et al.*, 2001, 2005). The 12 isolates with *plcD*::*IS6110* contained four or more *IS6110* copies

(Fig. 2b). In contrast, 74 (87 %) from 85 isolates without insertion in the *plcD* had a single *IS6110* copy, which in 73 isolates was detected as a 1.9 kb band and in 1 case was detected as a 1.1 kb band (Fig. 2b). Eight (9.4 %) isolates had two *IS6110* copies, two (2.4 %) isolates had three copies and only one (1.2 %) isolate presented four copies (Fig. 2b, data not shown).

The superimposition of *plcD* and *IS6110* RFLP films showed that both probes hybridized with fragments of the same length, about 1150 bp, in strains with *plcD*::*IS6110*, further substantiating the presence of an *IS6110* copy in the *plcD* gene (Fig. 2a, b). Hybridization with DR-derived probe and

Fig. 2. Fingerprint analysis of *M. bovis* isolates using probes complementary to *plcD* (a) and *IS6110* (b), analysed using BioNumerics v. 4.0. Specific bands hybridizing with each probe are shown on the left and hybridization patterns on the right. *Bands that co-hybridized with *IS6110*- and DR-derived probes. #Bands that co-hybridized with *IS6110* and *plcD*-derived probes. A–E, *IS6110*-RFLP patterns of isolates with intact *plcD*; F–J, *IS6110*-RFLP patterns of isolates with *plcD*::*IS6110*. A, *M. bovis* AN5; B, 1 isolate representative of 73 isolates with a single *IS6110* copy located in the 1.9 kb DR region; C, D, E, single isolates with 1, 2 and 4 *IS6110* copies, respectively, and without insertion in *plcD*; F, representative of 3 isolates from spoligotype sb0881 and 1 each from spoligotypes sb1033, sb0121 and ST1853; G, single isolate from spoligotype sb0134; H, representative of 3 isolates from spoligotype sb0881; I, single isolate from spoligotype sb0881; J, single isolate from spoligotype sb0295.

superimposition with IS6110 RFLP films showed that both probes always hybridized with fragments of the same length, confirming the presence of one IS6110 copy in the DR region in all isolates tested (Fig. 2b).

Distribution of spoligotype patterns in isolates with *plcD*::IS6110

Considering all 245 isolates, a total of 42 different spoligotypes were observed (data not shown). The 18 isolates with *plcD*::IS6110 were distributed in 6 previously nominated spoligotypes: sb0881, sb0121, sb1033, sb0134, sb0295 (from the *M. bovis* spoligotype database at <http://www.mbovis.org/>) and ST1853 (from the SpolDB4 database) (Table 2). A total of 111 additional isolates with intact *plcD* genes, as shown by *plcD* PCR, were allocated in these same 6 spoligotype patterns (Table 2).

Localization of IS6110 insertion site by sequencing

Sequence analysis of the 3 kb amplicons corroborated the presence, in all 18 strains, of the IS6110 element at the same

position, 1 972 894, relative to the genome of *M. bovis* AF2122/97 (Garnier *et al.*, 2003), in the same orientation, and disclosed a DR of three nucleotides CTC:CTC at the ends of the insertion sequence (Fig. 1b).

DISCUSSION

Insertion of a copy of IS6110 in the *plcD* gene was detected in 7.3 % of the isolates studied here. All isolates bearing *plcD*::IS6110 presented an indistinguishable four-banded hybridization pattern with the *plcD*-derived probe and presented the same insertion site in the *plcD* gene by DNA sequencing. Isolates were distributed in six spoligotype patterns and five IS6110 RFLP profiles. Eight different strains could be recognized by a combination of both methods (Table 2).

The observation that all *plcD*::IS6110 isolates originated from the same region (São Paulo, Brazil) and shared two IS6110 hybridizing bands carrying particular IS6110 insertions (*plcD*::IS6110 and an IS6110 copy in the DR region) suggests that they represent a unique evolutionary event

Table 2. Distribution of spoligotype and IS6110 RFLP patterns in isolates with *plcD*::IS6110 and intact *plcD*

MG, Minas Gerais; RS, Rio Grande do Sul; SP, São Paulo; NE, not evaluated.

No. of isolates	Origin	<i>plcD</i> PCR	IS6110 RFLP	Spoligotype	Pattern
3	SP	<i>plcD</i> ::IS6110	F	1101111101111110111101111000000111111100000	sb0881
3	SP	<i>plcD</i> ::IS6110	H		
1	SP	<i>plcD</i> ::IS6110	I		
6	SP	<i>plcD</i> ::IS6110	NE		
2	SP	<i>plcD</i>	1*		
1	SP	<i>plcD</i> ::IS6110	J	11011111011111101111011111111111110100000	sb0295
5	RS	<i>plcD</i>	1		
12	MG	<i>plcD</i>	1		
28	SP	<i>plcD</i>	1†		
1	SP	<i>plcD</i> ::IS6110	F	110111110111111011110111111111111111100000	sb0121
4	RS	<i>plcD</i>	1		
49	SP	<i>plcD</i>	1‡ or 2§		
1	SP	<i>plcD</i> ::IS6110	F	11011010000011101111111111111111110111100000	sb1033
5	SP	<i>plcD</i>	NE		
1	SP	<i>plcD</i> ::IS6110	F	1101111001111110111101111000111111110100000	ST1853
3	SP	<i>plcD</i>	NE		
1	SP	<i>plcD</i> ::IS6110	G	11000111011111101111111111111111111111100000	sb0134
3	SP	<i>plcD</i>	1*		

*One isolate evaluated.
†Thirteen isolates evaluated.
‡Four isolates evaluated.
§Two isolates evaluated.

followed by clonal expansion and divergence of spoligotype and IS6110 RFLP patterns. This hypothesis is reinforced by the similarities of spoligotype and IS6110 RFLP patterns, but is somewhat contradicted by the recognition of 111 additional isolates belonging to the same spoligotype patterns but showing intact *plcD* genes (Table 2). Forty-two of these isolates were also evaluated by IS6110 RFLP. All had a copy of IS6110 inserted in the 1.9 kb fragment of the DR region but not in the 1.4 kb fragment (data not shown). Therefore, an independent evolution of the different genetic events cannot be excluded.

The DR region is a polymorphic locus in the genome of *M. tuberculosis* complex members comprising a cluster of directly repeating sequences of 36 bp, separated by unique spacer sequences of 36–41 bp. One repeat sequence and the following spacer sequence have been named DVR (direct variable repeat). Polymorphisms in this region can originate from transposition of the IS6110 element inside the DR region or successive deletions from a primordial DR locus, probably mediated by homologous recombination between IS6110 elements, neighbouring or distant directly repeating sequences and/or by slippage during DNA replication (Fang *et al.*, 1998; Groenen *et al.*, 1993; van Embden *et al.*, 2000; Warren *et al.*, 2002). IS6110-dependent mechanisms of evolution would not be frequent in *M. bovis* isolates due to the low-copy number of IS6110 elements. In this context, spoligotype pattern sb0121 showed the most intact structure, and patterns sb0295, sb0881 and ST1853 could have originated from a common ancestor with the same DR structure as sb0121 by successive loss of DVRs in dependent or independent events. Spoligotype patterns sb1033 and sb0134 retain DVR21, which is missing in the other spoligotype patterns, and therefore were probably generated by independent events. Except for spoligotype pattern ST1853, which was found only in isolates from Brazil, the other five spoligotype patterns were previously observed in *M. bovis* isolates from France, the Netherlands, Belgium and Australia, according to data from *M. bovis* spoligotype database at <http://www.mbovis.org/>. They were not found in 600 *M. bovis* isolates from Argentina (data not shown). Spoligotype patterns sb0121 and sb0134 were predominant in a study on *M. bovis* isolates from Europe (Haddad *et al.*, 2001). Due to the wide geographic distribution of these spoligotype patterns it is also possible that they have evolved convergently through preferential deletion of specific DVRs (Fang *et al.*, 1998; Warren *et al.*, 2002). This would justify the existence of isolates with and without insertion in the *plcD* gene belonging to the same spoligotype pattern.

The majority of *M. bovis* isolates have a single copy of IS6110 stably inserted in the 1.9 kb fragment of the DR region (van Embden *et al.*, 2000). On rare occasions, this element replicates and inserts into different regions of the genome. Transposition can be assisted by activation of adjacent promoters, and transposition frequency is higher if the insertion sequence is inserted into transcriptionally active sites (Wall *et al.*, 1999). *plcD* is a transcriptionally active site

(Raynaud *et al.*, 2002; Viana-Niero *et al.*, 2004) and, therefore, insertion into this gene can potentially activate IS6110 transposition events. In addition to the two IS6110 copies common to all *plcD*::IS6110 isolates, seven isolates had a further two common IS6110 copies on 4.3 kb and 2.5 kb *PvuII* fragments and one of these isolates subsequently acquired another IS6110 copy (patterns F and G). The other five *plcD*::IS6110 isolates belong to three IS6110 RFLP patterns that also had additional common IS6110 copies on 6.3, 4 and 2.2 kb *PvuII* fragments (patterns H and I), and on 2.2 and 1.7 kb *PvuII* fragments (patterns I and J). Isolates of spoligotype pattern sb0881 showed higher IS6110 transposition activity, comprising three different IS6110 RFLP patterns, from which at least one (F) could have evolved independently from the other two (H and I), suggesting parallel evolution of the IS6110 banding pattern in these strains.

The *plcD* locus has been identified as a preferential site for IS6110 insertion in *M. tuberculosis* (Gordon *et al.*, 1999; Ho *et al.*, 2000; Sampson *et al.*, 1999). Yesilkaya *et al.* (2005) identified 9 distinct IS6110 insertion sites within the *plcD* gene in 161 clinical *M. tuberculosis* isolates, and Kong *et al.* (2005) reported the existence of 27 different insertion sites within the *plcD* gene in 144 clinical *M. tuberculosis* isolates. However, insertions at the same position of the *plcD* gene in different *M. tuberculosis* isolates were also reported (Kong *et al.*, 2005; Viana-Niero *et al.*, 2004; Warren *et al.*, 2000). The insertion position identified here, which is different from insertion positions reported in *M. tuberculosis* isolates, was the same in all *M. bovis* isolates, regardless of differences in spoligotype and IS6110 RFLP patterns. Together these findings suggest that IS6110 insertions at the same site in the *plcD* gene could potentially originate from independent events. Independent occurrences of IS6110 insertion at *ipl* loci (*iplA* and *iplB*) in different *M. tuberculosis* isolates were also reported (Fang *et al.*, 2001). Further research, using additional markers or sequencing will disclose if *M. bovis* with insertion in *plcD* has arisen from an ancestral clone or resulted from independent events of transposition.

Deletions in the *plcD* locus have been reported in *M. tuberculosis* clinical isolates (Talarico *et al.*, 2005; Viana-Niero *et al.*, 2004). Lari *et al.* (2001) showed that 15 out of 45 isolates of *M. tuberculosis* studied presented the RvD2 deletion. We have shown that deletions of *plc* genes in *M. tuberculosis* isolates originate from homologous recombination between two copies of IS6110 (Viana-Niero *et al.*, 2004). Deletions in *plcD* gene were not found in any of 245 *M. bovis* isolates studied here, as the primer sites at each end of the *plcD* gene were always present, irrespective of the presence of the IS6110 element. The low-copy number of IS6110 elements in *M. bovis* could explain the absence of this occurrence, as proximity of two copies of the insertion element in the same orientation is necessary for homologous recombination, and this proximity is not expected to occur in *M. bovis*.

No difference in the organ distribution of lesions was observed when animals infected with strains with *plcD*::IS6110 were

compared to animals infected with strains with intact *plcD* of the same spoligotype pattern (data not shown). Therefore, it seems that the *plcD* product is not involved in the organ dissemination of *M. bovis*. Despite the well-known role of phospholipases C in the virulence of several intracellular bacterial pathogens (Berka *et al.*, 1981; Logan *et al.*, 1991; Vazquez-Boland *et al.*, 1992), the role of these enzymes in the pathogenicity of mycobacteria has not yet been completely understood. In a previous study (Viana-Niero *et al.*, 2004) we demonstrated that lack of transcription of all four *plc* genes did not impair the ability of *M. tuberculosis* isolates to cause disease in their hosts. Raynaud *et al.* (2002) concluded that the *plcD* gene did not contribute to the virulence of *M. tuberculosis* but could act in association with other *plc* genes. Yang *et al.* (2005) demonstrated that *M. tuberculosis* bearing *plcD* gene mutations generated more extra-pulmonary tuberculosis than strains with intact *plcD*. As yet unresolved questions are whether insertions in *plcD* represent a biological advantage for the bacterium, and what the precise role of this gene in bacterial survival and pathogenicity is.

ACKNOWLEDGEMENTS

Jorge Luiz Mello Sampaio is acknowledged for sending DNA from two *M. bovis* human isolates from São Paulo. Christophe Sola and Nalin Rastogi are acknowledged for tracking the ST1853 spoligotype pattern in the SpolDB4 database developed at the Pasteur Institute of Guadeloupe. This work was supported by Centro Argentino Brasileiro de Biotecnologia (CABBIO) Proyecto 13, and grant 480382/01-8 from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Cristina Viana-Niero is the recipient of a fellowship from FAPESP (04/07394-5). Angel Cataldi and Sylvia Cardoso Leão are members of RELACTB (Latin American and Caribbean Tuberculosis Network).

REFERENCES

- Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S. & Small, P. M. (1999). Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284**, 1520–1523.
- Berka, R. M., Gray, G. L. & Vasil, M. L. (1981). Studies of phospholipase C (heat-labile hemolysin) in *Pseudomonas aeruginosa*. *Infect Immun* **34**, 1071–1074.
- Brosch, R., Gordon, S. V., Marmiesse, M. & 12 other authors (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* **99**, 3684–3689.
- Centro Panamericano de Zoonosis (1973). Métodos de laboratorio de micobacteriología veterinaria para el aislamiento e identificación de micobacterias. In *Serie de Monografías Científicas y Técnicas*, pp. 48. Buenos Aires: Panamerican Health Organization.
- Cole, S. T., Brosch, R., Parkhill, J. & 39 other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.
- Eisenach, K. D., Cave, M. D., Bates, J. H. & Crawford, J. T. (1990). Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *J Infect Dis* **161**, 977–981.
- Fang, Z., Morrison, N., Watt, B., Doig, C. & Forbes, K. J. (1998). IS6110 transposition and evolutionary scenario of the direct repeat locus in a group of closely related *Mycobacterium tuberculosis* strains. *J Bacteriol* **180**, 2102–2109.
- Fang, Z., Kenna, D. T., Doig, C., Smittipat, D. N., Palittapongarnpim, P., Watt, B. & Forbes, K. J. (2001). Molecular evidence for independent occurrence of IS6110 insertions at the same sites of the genome of *Mycobacterium tuberculosis* in different clinical isolates. *J Bacteriol* **183**, 5279–5284.
- Garnier, T., Eiglmeier, K., Camus, J. C. & 19 other authors (2003). The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* **100**, 7877–7882.
- Gordon, S. V., Brosch, R., Billault, A., Garnier, T., Eiglmeier, K. & Cole, S. T. (1999). Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. *Mol Microbiol* **32**, 643–655.
- 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.
- Haddad, N., Ostyn, A., Karoui, C., Masselot, M., Thorel, M. F., Hughes, S. L., Inwald, J., Hewinson, R. G. & Durand, B. (2001). Spoligotype diversity of *Mycobacterium bovis* strains isolated in France from 1979 to 2000. *J Clin Microbiol* **39**, 3623–3632.
- Hermans, P. W., van Soolingen, D., Bik, E. M., de Haas, P. E., Dale, J. W. & van Embden, J. D. (1991). Insertion element IS987 from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. *Infect Immun* **59**, 2695–2705.
- Ho, T. B., Robertson, B. D., Taylor, G. M., Shaw, R. J. & Young, D. B. (2000). Comparison of *Mycobacterium tuberculosis* genomes reveals frequent deletions in a 20 kb variable region in clinical isolates. *Yeast* **17**, 272–282.
- Kamerbeek, J., Schouls, L., Kolk, A. & 8 other authors (1997). Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* **35**, 907–914.
- Kong, Y., Cave, M. D., Yang, D. & 7 other authors (2005). Distribution of insertion- and deletion-associated genetic polymorphisms among four *Mycobacterium tuberculosis* phospholipase C genes and associations with extrathoracic tuberculosis: a population-based study. *J Clin Microbiol* **43**, 6048–6053.
- Lari, N., Rindi, L. & Garzelli, C. (2001). Identification of one insertion site of IS6110 in *Mycobacterium tuberculosis* H37Ra and analysis of the RvD2 deletion in *M. tuberculosis* clinical isolates. *J Med Microbiol* **50**, 805–811.
- Logan, A. J., Williamson, E. D., Titball, R. W., Percival, D. A., Shuttleworth, A. D., Conlan, J. W. & Kelly, D. C. (1991). Epitope mapping of the alpha-toxin of *Clostridium perfringens*. *Infect Immun* **59**, 4338–4342.
- Nguyen, D., Brassard, P., Menzies, D., Thibert, L., Warren, R., Mostowy, S. & Behr, M. (2004). Genomic characterization of an endemic *Mycobacterium tuberculosis* strain: evolutionary and epidemiologic implications. *J Clin Microbiol* **42**, 2573–2580.
- Raynaud, C., Guilhot, C., Rauzier, J., Bordat, Y., Pelicic, V., Manganelli, R., Smith, I., Gicquel, B. & Jackson, M. (2002). Phospholipases C are involved in the virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **45**, 203–217.
- Sampson, S. L., Warren, R. M., Richardson, M., van der Spuy, G. D. & van Helden, P. D. (1999). Disruption of coding regions by IS6110 insertion in *Mycobacterium tuberculosis*. *Tuber Lung Dis* **79**, 349–359.
- Talarico, S., Durmaz, R. & Yang, Z. (2005). Insertion- and deletion-associated genetic diversity of *Mycobacterium tuberculosis*

phospholipase C-encoding genes among 106 clinical isolates from Turkey. *J Clin Microbiol* **43**, 533–538.

van Embden, J. D. A., Cave, M. D., Crawford, J. T. & 8 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.

Vazquez-Boland, J. A., Kocks, C., Dramsi, S., Ohayon, H., Geoffroy, C., Mengaud, J. & Cossart, P. (1992). Nucleotide sequence of the lecithinase operon of *Listeria monocytogenes* and possible role of lecithinase in cell-to-cell spread. *Infect Immun* **60**, 219–230.

Vera-Cabrera, L., Hernandez-Vera, M. A., Welsh, O., Johnson, W. M. & Castro-Garza, J. (2001). Phospholipase region of *Mycobacterium tuberculosis* is a preferential locus for IS6110 transposition. *J Clin Microbiol* **39**, 3499–3504.

Vestal, A. L. (1975). *Procedures for the Isolation and Identification of Mycobacteria*. Atlanta, GA: Centers for Disease Control.

Viana-Niero, C., de Haas, P. E., van Soolingen, D. & Leao, S. C. (2004). Analysis of genetic polymorphisms affecting the four phospholipase C (*plc*) genes in *Mycobacterium tuberculosis* complex clinical isolates. *Microbiology* **150**, 967–978.

Wall, S., Ghanekar, K., McFadden, J. & Dale, J. W. (1999). Context-sensitive transposition of IS6110 in mycobacteria. *Microbiology* **145**, 3169–3176.

Warren, R. M., Sampson, S. L., Richardson, M., van der Spuy, G. D., Lombard, C. J., Victor, T. C. & van Helden, P. D. (2000). Mapping of

IS6110 flanking regions in clinical isolates of *Mycobacterium tuberculosis* demonstrates genome plasticity. *Mol Microbiol* **37**, 1405–1416.

Warren, R. M., Streicher, E. M., Sampson, S. L., van der Spuy, G. D., Richardson, M., Nguyen, D., Behr, M. A., Victor, T. C. & van Helden, P. D. (2002). Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. *J Clin Microbiol* **40**, 4457–4465.

Wayne, L. G. (1984). Mycobacterial speciation. In *The Mycobacteria: a Sourcebook*, pp. 42–43. Edited by G. P. Kubica & L. G. Wayne. New York: Marcel Dekker.

Yang, Z., Yang, D., Kong, Y., Zhang, L., Marrs, C. F., Foxman, B., Bates, J. H., Wilson, F. & Cave, M. D. (2005). Clinical relevance of *Mycobacterium tuberculosis plcD* gene mutations. *Am J Respir Crit Care Med* **171**, 1436–1442.

Yesilkaya, H., Dale, J. W., Strachan, N. J. & Forbes, K. J. (2005). Natural transposon mutagenesis of clinical isolates of *Mycobacterium tuberculosis*: how many genes does a pathogen need? *J Bacteriol* **187**, 6726–6732.

Zanini, M. S., Moreira, E. C., Lopes, M. T. & 8 other authors (2001). *Mycobacterium bovis*: polymerase chain reaction identification in bovine lymphonode biopsies and genotyping in isolates from Southeast Brazil by spoligotyping and restriction fragment length polymorphism. *Mem Inst Oswaldo Cruz* **96**, 809–813.

Zanini, M. S., Moreira, E. C., Salas, C. E., Lopes, M. T., Barouni, A. S., Roxo, E., Telles, M. A. & Zumarraga, M. J. (2005). Molecular typing of *Mycobacterium bovis* isolates from South-East Brazil by spoligotyping and RFLP. *J Vet Med B Infect Dis Vet Public Health* **52**, 129–133.