

PCR-Based Genotyping of *Mycobacterium tuberculosis* with New GC-Rich Repeated Sequences and IS6110 Inverted Repeats Used as Primers

Roman Kotłowski,[†] Isdore Chola Shamputa, Nabil Abdullah El Aila, Anna Sajduda, Leen Rigouts, Armand van Deun, and Françoise Portaels*

Mycobacteriology Unit, Prince Leopold Institute of Tropical Medicine, B-2000 Antwerp, Belgium

Received 31 March 2003/Returned for modification 2 June 2003/Accepted 27 August 2003

In the present study we attempted to develop a PCR-based epidemiological tool for the differentiation of *Mycobacterium tuberculosis* isolates. Use of the designed primers Mtb1 (5'-CCG-GCG-GGG-CCG-GCG-G) and Mtb2 (5'-CGG-CGG-CAA-CGG-CGG-C) targeting frequently repeated 16-bp sequences in combination with primers sited at the inverted repeats flanking IS6110 allowed differentiation of *M. tuberculosis* isolates.

Mycobacterium tuberculosis is the most frequent cause of bacterial disease in humans. Tuberculosis kills approximately 2 million people each year, and more than 8 million people suffer from the disease annually (16). The burden of *M. tuberculosis* infections is greatest in low-income countries, where recurrent disease in previously treated persons is most prevalent. DNA fingerprinting techniques allow differentiating between endogenous reactivation and reinfection with a new strain in case of recurrence (4, 14). In the present study we applied new frequently repeated GC-rich sequences and inverted repeats flanking IS6110 as primers in a PCR-based genotyping method of *M. tuberculosis* isolates.

The most common 16-bp oligonucleotides were found in whole *M. tuberculosis* genomes (2, 5) by using Clone Manager (version 7; Scientific and Educational Software) with successive replacement of S characters in the string SSS-SSS-SSS-SSS-SSS-S with C or G letters according to the highest rate of occurrence of direct and inverted repeats. As a result, the Mtb1 (5'-CCG-GCG-GGG-CCG-GCG-G) sequence was found and was subsequently used in pDRAW32 (version 1.0; AcaClone Software) for comparison to the whole genome, with the indication of four possible mismatches. Repeated oligonucleotides were subsequently extended by successive additions of single nucleotides from the 5' and 3' ends until the repeat number was not lower than 30 (9). Independent searching for at least 16-bp repeats in the genome of *M. tuberculosis* H37Rv by using Reputer (version 2.74-linux.i586; AG Praktische Informatik, Universität Bielefeld [<http://www.genomes.de>]) confirmed our results (presented in Table 1). Sequence number 6 (5'-CGG-CGG-CAA-CGG-CGG-CA) containing the 16-bp sequence of the Mtb2 oligonucleotide was previously published (12), but to the best of our knowledge its use as a primer for PCR typing of *M. tuberculosis* in combination with primers

designed within IS6110 inverted repeats is reported here for the first time.

Bacterial suspensions were prepared by scraping of three to four loops from Löwenstein-Jensen slants and suspension of the bacilli in Tris-EDTA buffer. For DNA extraction from mycobacteria, a standardized method based on heat inactivation for 20 min at 85°C, followed by lysis with lysozyme and sodium dodecyl sulfate-proteinase K, extraction by cetyltrimethylammonium bromide-NaCl, and chloroform-isoamyl alcohol purification was used (15). In a simple PCR for *M. tuberculosis* strain differentiation, selected GC-rich Mtb1 and Mtb2 sequences were separately used as primers in combination with primers IS1 (5'-CGG-ACT-CAC-CGG-GGC-GGT-TCA) and IS2 (5'-CGG-ACA-TGC-CGG-GGC-GGT-TCA) designed within inverted repeats of IS6110. For the PCR, 3- μ l aliquots of PvuII-digested genomic DNA was added to a 47- μ l PCR mixture containing 25 pmol of each primer, 1 U of AmpliTaq DNA polymerase (Roche Molecular Systems, Brussels, Belgium), 0.2 mM concentrations of each deoxyribonucleotide triphosphate, 10 mM Tris-HCl (pH 8.4), 1.65 mM MgCl₂, 50 mM KCl, and 0.1% Triton X-100 and overlaid with mineral oil. Cycling conditions were as follows: denaturation at 94°C for 5 min, followed by amplification for 35 cycles of 94°C for 1 min, 68°C (or 66°C) for 1 min, and 72°C for 1 min, followed by a final extension at 72°C for 10 min. The annealing temperature at 68°C used in the PCR for Mtb1 primer is the optimal hybridization temperature for mycobacteria (1). For Mtb2 primer, with 14 GC nucleotides, a lower temperature at 66°C for efficient DNA stringency is required. A total of 20 μ l of amplified DNA was subjected to electrophoresis through a 2% agarose gel, detected by ethidium bromide staining, and visualized under UV light.

The principle of the presented PCR genotyping method is depicted in Fig. 1. The selection criteria for the Mtb1 and Mtb2 primers were their high GC content and high copy number in *M. tuberculosis* and *M. bovis* (7) genomes (Table 1). The distribution of possible regions for hybridization of Mtb1, Mtb2, and IS6110 sequences in the genomes of *M. tuberculosis* (H37Rv and CDC1551) is presented in Fig. 2. A Clone Manager version 7-aided comparison analysis showed that the en-

* Corresponding author. Mailing address: Prince Leopold Institute of Tropical Medicine, Mycobacteriology Unit, Department of Microbiology, Nationalestraat 155, 2000 Antwerp, Belgium. Phone: 32-3-247-63-17. Fax: 32-3-247-63-33. E-mail: portaels@itg.be.

[†] Present address: Faculty of Chemistry, Gdansk University of Technology, 80-952 Gdansk, Poland.

TABLE 1. Oligonucleotide repeats found in mycobacterial genomes

| Genome | No. of repeats in genomes obtained with primer: | |
|---------------------------------|---|----------------------------|
| | Mtb1 (5'-CCGGCGGGGCCGCGG) | Mtb2 (5'-CGGCGGCAACGGCGGC) |
| <i>M. tuberculosis</i> H37Rv | 157 | 152 |
| <i>M. tuberculosis</i> CDC 1551 | 151 | 150 |
| <i>M. bovis</i> AF2122/97 | 173 | 149 |
| <i>M. leprae</i> (3) | 0 | 0 |

ture Mtb1 and Mtb2 sequences cannot hybridize with the genomes of *Corynebacterium efficiens* YS-314, *Escherichia coli* K-12, *Staphylococcus epidermidis* ATCC 12228, *Bacillus anthracis* AMES, *Clostridium perfringens*, *Chlamydia pneumoniae* J138, *Staphylococcus aureus* MW2, and *Helicobacter pylori* J99 (ftp://ftp.ncbi.nih.gov/GenBank/genomes/Bacteria/).

In experimental studies, 27 *M. tuberculosis* isolates were analyzed. Ten colonies isolated from one sputum sample were subjected to spoligotyping (Isogen; Bioscience, Maarsse, The Netherlands) and IS6110-restriction fragment length polymorphism (RFLP) assay (13). One of the 10 colonies tested (colony 8) showed a small variation in spoligotyping and IS6110-

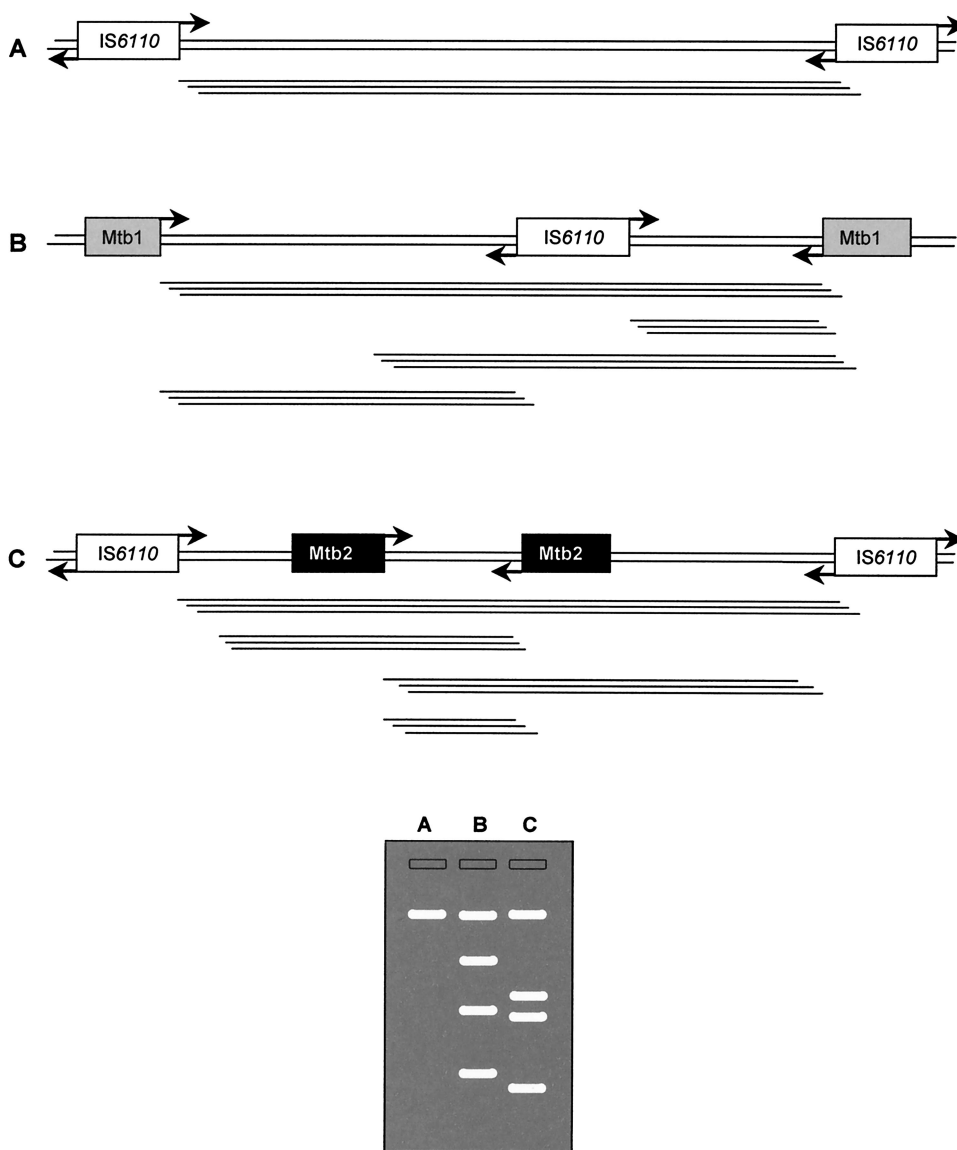


FIG. 1. Theoretical illustration of our genotyping method based on PCR amplification of DNA regions between two closely located IS6110 copies (A), between IS6110 and Mtb1 (B), and between IS6110 and Mtb2 (C).

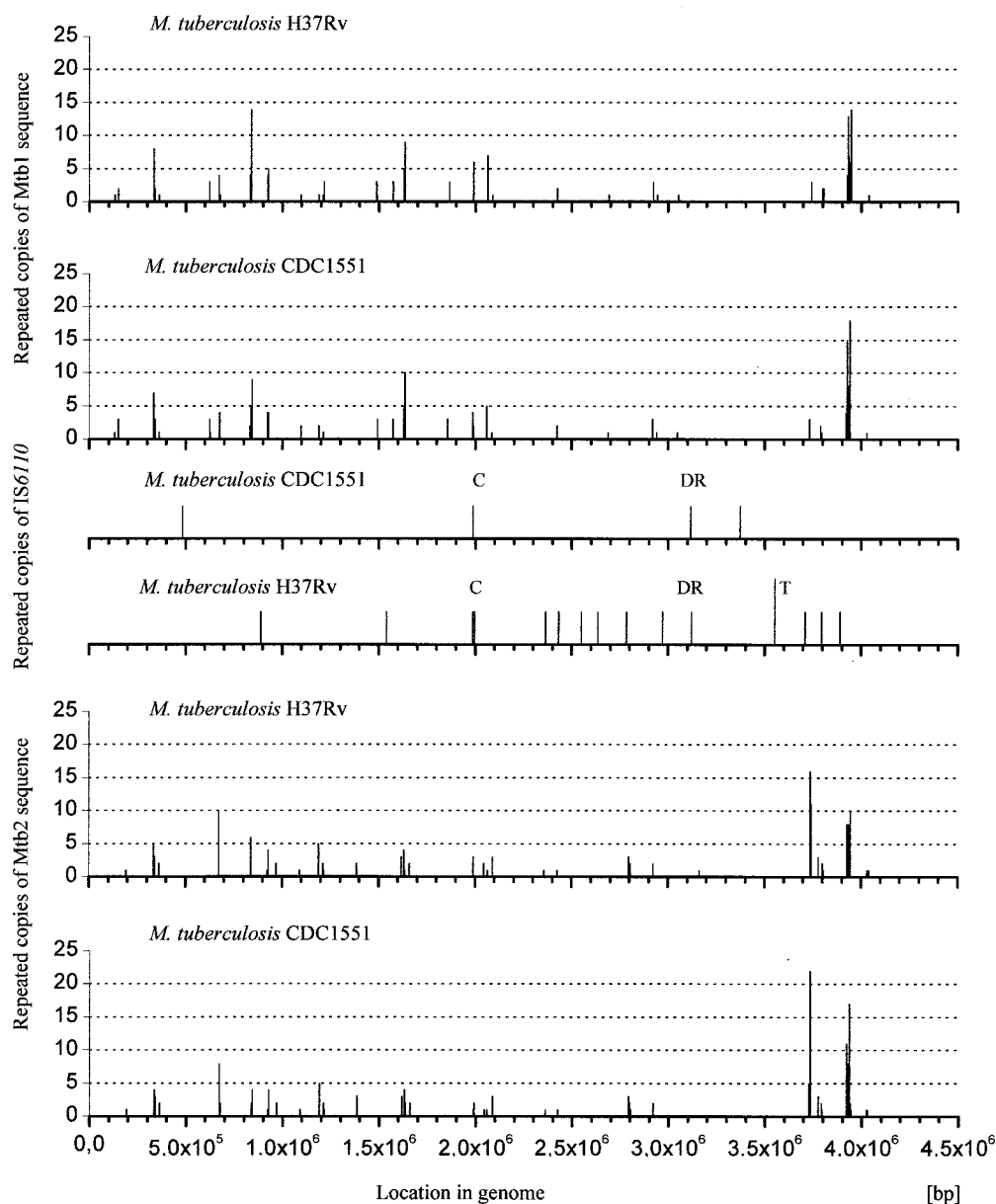


FIG. 2. Stacked histograms presenting the distribution of Mtb1 and Mtb2 sequences and the insertion sequence IS6110 in the genomes of *M. tuberculosis* (H37Rv and CDC1551). C, location of IS6110 at 36-nucleotide position upstream of Rv1758 and MT1805 genes coding for cutinase; DR, location of IS6110 within direct-repeat region; T, IS6110 tandem repeats.

RFLP: a deficiency in spacer 25 in the spoligotype and transposition of one IS6110 copy from a 1.4-kb *Pvu*II fragment to a 6.1-kb *Pvu*II fragment (Fig. 3). These data suggest the existence of a mixed bacterial population in the sample tested. By using the primer combination Mtb1-IS1-IS2 or Mtb2-IS1-IS2, clone 8 could be also differentiated by our new PCR typing method (Fig. 3). In addition, 13 of 17 *M. tuberculosis* tested strains showed identical affiliation into six clusters with our PCR-based genotyping method and IS6110-RFLP assay (Fig. 4, lanes 1 to 6 and 14 to 20). The other seven isolates showed

different banding patterns by both methods (Fig. 4, lanes 7 to 13).

The PCR-fingerprinting method presented here can be useful for laboratories performing *M. tuberculosis* epidemiology studies. The major advantage of this method versus DRE-PCR (6) is the high annealing temperature that improves reproducibility of results. In addition, like the IS6110-inverse PCR (10) and ligation-mediated PCR (11), this method can be helpful in the localization of IS6110 copies throughout the genomes of *M. tuberculosis* by sequencing amplimers, thereby providing

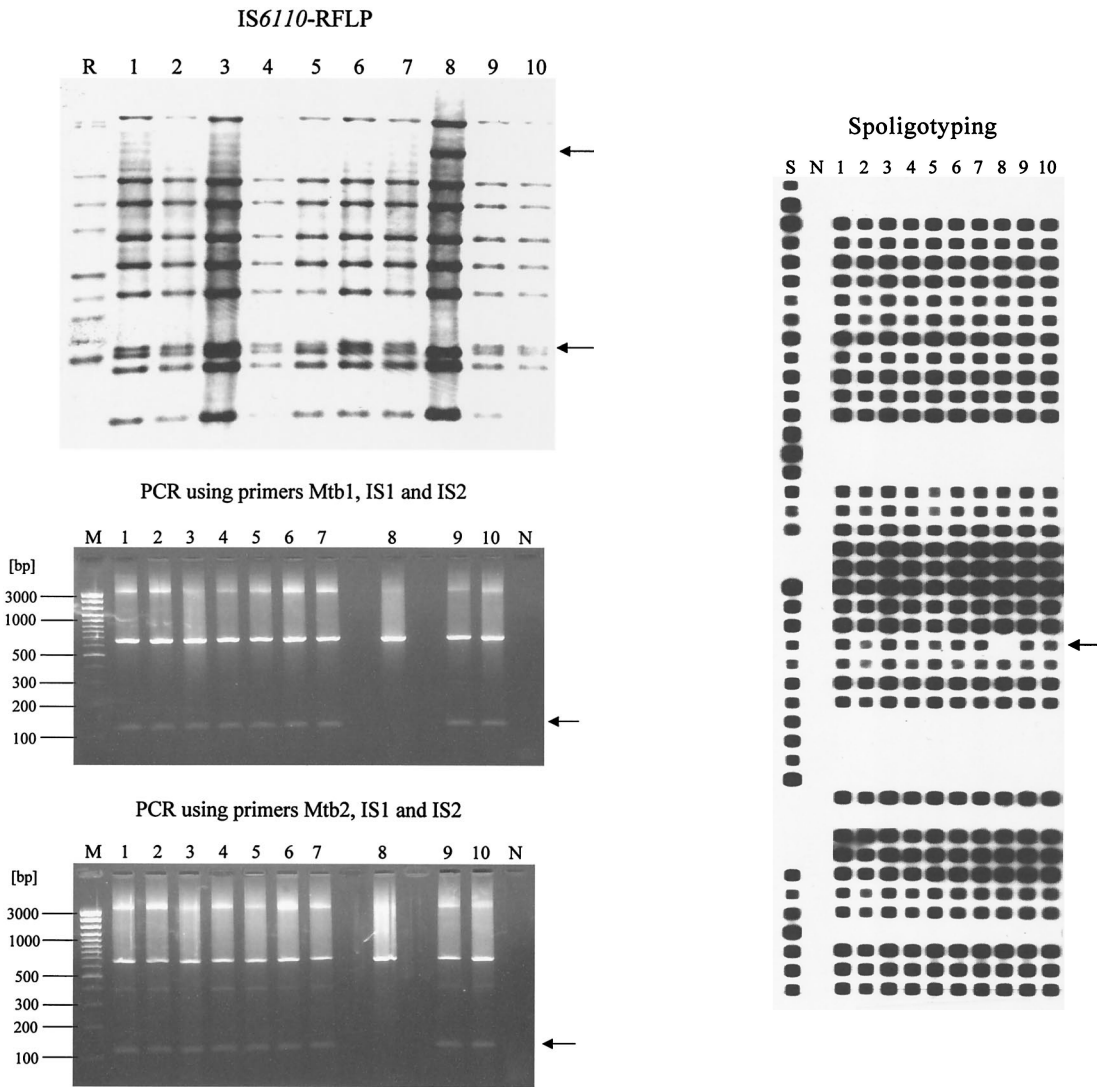


FIG. 3. Results of typing of 10 clones isolated from one sputum sample ITM 02-0544 from Bangladesh by three methods: IS6110-RFLP, spoligotyping, and PCR with the the primer combinations Mtb1-IS1-IS2 or Mtb2-IS1-IS2. Arrows in each of the typing methods indicate differences between clone no. 8 and nine other clones. R, *M. tuberculosis* Mt 14323; S, *M. tuberculosis* H37Rv; N, negative control; M, 100-bp DNA Ladder Plus (MBI Fermentas).

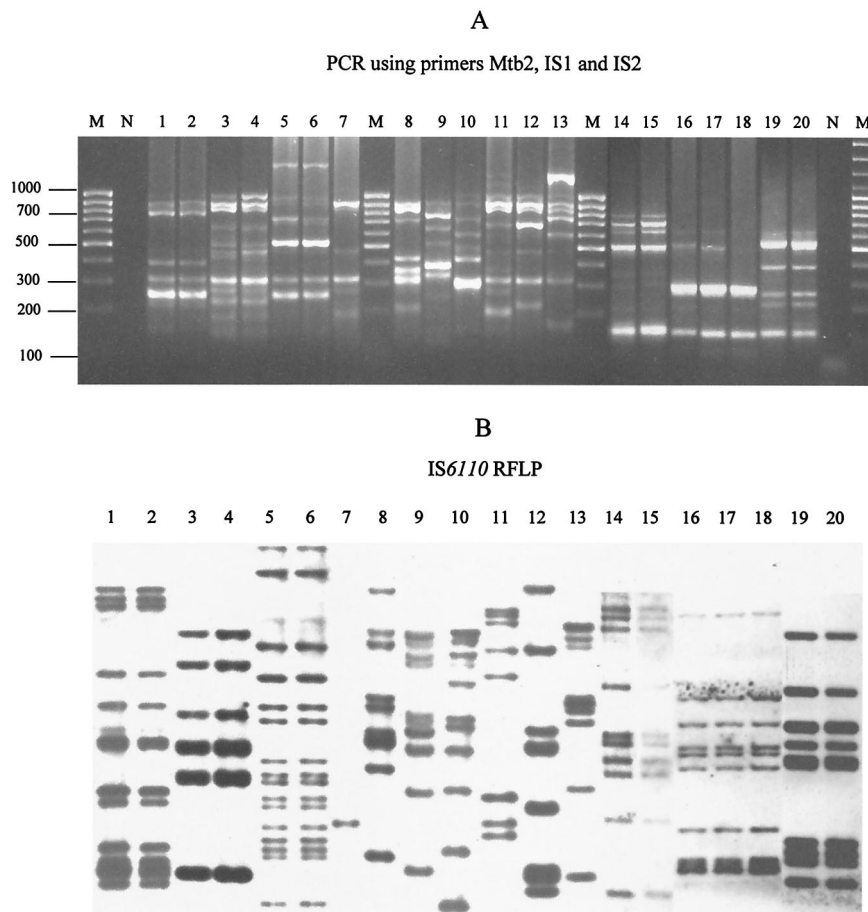


FIG. 4. Differentiation of *M. tuberculosis* strains by using our PCR-based genotyping method (A) and by IS6110-RFLP fingerprinting (B). Lanes: 1 and 2, *M. tuberculosis* ITM3288; 3 and 4, *M. tuberculosis* ITM3255; 5 and 6, *M. tuberculosis* ITM3192; 7, *M. tuberculosis* ITM3153; 8, *M. tuberculosis* ITM4454; 9, *M. tuberculosis* ITM3132; 10, *M. tuberculosis* ITM4451; 11, *M. tuberculosis* ITM4424; 12, *M. tuberculosis* ITM4337; 13, *M. tuberculosis* ITM4339; 14, *M. tuberculosis* MZ1830; 15, *M. tuberculosis* MZ1799; 16, *M. tuberculosis* KP8689; 17, *M. tuberculosis* KP2398; 18, *M. tuberculosis* KP3189; 19, *M. tuberculosis* SL19; 20, *M. tuberculosis* SL11; N, negative control; M, 100-bp DNA Ladder Plus (MBI Fermentas).

more information about possible “hot spots” for integration of IS6110 in the genomes of clinical isolates (8).

This study was funded by Damien Foundation (Brussels, Belgium). R.K. was supported by Marie Curie Individual Fellowship under research grant MCFI-1999-01398.

REFERENCES

- Athwal, R. S., S. S. Deo, and T. Imaeda. 1984. Deoxyribonucleic acid relatedness among *Mycobacterium leprae*, *Mycobacterium lepraemurium* and selected bacteria by dot blot and spectrophotometric deoxyribonucleic acid hybridization assays. *Int. J. Syst. Bacteriol.* **34**:371–375.
- Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaiia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B. G. Barrell, et al. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **11**:537–544.
- Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell. 2001. Massive gene decay in the leprosy bacillus. *Nature* **409**:1007–1011.
- du Plessis, D. G., R. Warren, M. Richardson, J. J. Joubert, and P. D. van Helden. 2001. Demonstration of reinfection and reactivation in HIV-negative autopsied cases of secondary tuberculosis: multilevel genotyping of *Mycobacterium tuberculosis* utilizing IS6110 and other repetitive element-based DNA fingerprinting. *Tuberculosis* **81**:211–220.
- Fleischmann, R. D., D. Alland, J. A. Eisen, L. Carpenter, O. White, J. Peterson, R. DeBoy, R. Dodson, M. Gwinn, D. Haft, E. Hickey, J. F. Kolonay, W. C. Nelson, L. A. Umayam, M. Ermolaeva, S. L. Salzberg, A. Delcher, T. Utterback, J. Weidman, H. Khouri, J. Gill, A. Mikula, W. Bishai, W. R. Jacobs, Jr., J. C. Venter, and C. M. Fraser. 2002. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* **184**:5479–5490.
- Friedman, C. R., M. Y. Stoeckle, W. D. Johnson Jr, and L. W. Riley. 1995. Double repetitive element PCR method for subtyping *Mycobacterium tuberculosis* clinical isolates. *J. Clin. Microbiol.* **33**:1064–1069.
- Garnier, T., K. Eiglmeier, J. C. Camus, N. Medina, H. Mansoor, M. Pryor, S. Duthoy, S. Grondin, C. Lacroix, C. Monsempe, S. Simon, B. Harris, R. Atkin, J. Doggett, R. Mayes, L. Keating, P. R. Wheeler, J. Parkhill, B. G. Barrell, S. T. Cole, S. V. Gordon, and R. G. Hewinson. 2003. The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. USA* **100**:7877–7882.
- Kivi, M., X. Liu, S. Raychaudhuri, R. B. Altman, and P. M. Small. 2002. Determining the genomic locations of repetitive DNA sequences with a whole-genome microarray: IS6110 in *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **40**:2192–2198.
- Kotłowski, R., and R. Tylingo. 2003. Oligonucleotide repeats in genomes. Department of Food Chemistry and Technology, Gdansk University of Technology, Gdansk, Poland. [Online.] <http://www.pg.gda.pl/chem/Katedry/Zywnosc/oligos.html>.
- Otal, I., S. Samper, M. P. Asensio, M. A. Vitoria, M. C. Rubio, R. Gomez-Lus, and C. Martin. 1997. Use of a PCR method based on IS6110 polymor-

- phism for typing *Mycobacterium tuberculosis* strains from BACTEC cultures. J. Clin. Microbiol. **35**:273–277.
11. **Prod'hom, G., B. Lagier, V. Pelicic, A. J. Hance, B. Gicquel, and C. Guilhot.** 1998. A reliable amplification technique for the characterization of genomic DNA sequences flanking insertion sequences. FEMS Microbiol. Lett. **158**: 75–81.
 12. **Poulet, S., and S. T. Cole.** 1995. Repeated DNA sequences in mycobacteria. Arch. Microbiol. **163**:79–86.
 13. **van Embden, J. D., M. D. Cave, J. T. Crawford, J. W. Dale, K. D. Eisenach, B. Gicquel, P. Hermans, C. Martin, R. McAdam, T. M. Shinnick, et al.** 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. **31**:406–409.
 14. **van Rie, A., R. Warren, M. Richardson, T. C. Victor, R. P. Gie, D. A. Enarson, N. Beyers, and P. D. van Helden.** 1999. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. N. Engl. J. Med. **341**:1174–1179.
 15. **van Soolingen, D., P. W. Hermans, P. E. de Haas, D. R. Soll, and J. D. van Embden.** 1991. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J. Clin. Microbiol. **29**:2578–2586.
 16. **World Health Organization.** 2002. Tuberculosis. Fact sheet no. 104. World Health Organization, Geneva, Switzerland. [Online.] <http://www.who.int/mediacentre/factsheets/who104/en/index.html>.