JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2003, p. 1712–1716 Vol. 41, No. 4
0095-1137/03/$08.00+0 DOI: 10.1128/JCM.41.4.1712–1716.2003

Distinctiveness of *Mycobacterium tuberculosis* Genotypes from Human Immunodeficiency Virus Type 1-Seropositive and -Seronegative Patients in Lima, Peru

Niyaz Ahmed,[1] Luz Caviedes,[2] Mahfooz Alam,[1] K. Rajender Rao,[1] Vartul Sangal,[1] Patricia Sheen,[2] Robert H. Gilman,[2,3,4] and Seyed E. Hasnain[1,5,6]*

Centre for DNA Fingerprinting and Diagnostics (CDFD), Nacharam, Hyderabad 500076,[1] National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi,[5] and Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore,[6] India; Department of Pathology, Universidad Peruana Cayetano Heredia,[2] and Asociación Benefica Proyectos en Informatica, Salud, Medicina y Agricultura (AB PRISMA),[3] Lima, Peru; and Johns Hopkins School of Public Health and Hygiene, Baltimore, Maryland[4]

Received 20 September 2002/Returned for modification 16 November 2002/Accepted 18 January 2003

Genotypic analysis of *Mycobacterium tuberculosis* isolates obtained from human immunodeficiency virus type 1 (HIV-1)-seropositive ($n = 80$) and -seronegative ($n = 25$) patients from Lima, Peru, revealed two distinct genotypes correlating with the host immune status. While the level of intrastrain diversity of DNA fingerprints of HIV-seropositive isolates was less pronounced, these isolates showed many clonal groupings.

Genotypic analyses of *Mycobacterium tuberculosis* strains have repeatedly identified a group of strains linked to a high prevalence of disease in certain populations (3, 5, 18, 20). The emergence of the AIDS pandemic has led to a major shift in our approaches towards epidemiological studies of tuberculosis (11). Studies involving restriction fragment length polymorphism analysis of *M. tuberculosis* isolates obtained from AIDS patients showed that both reinfection and new infection occur in AIDS patients (10, 19). It has been speculated that AIDS patients constitute an ecological niche for *M. tuberculosis*, allowing less virulent strains to multiply freely without the selection pressure provided by an immunocompetent host (21). Consequently, it should be possible to differentially identify pathogenic bacterial clones on the basis of epidemiological parameters related to coinfection, relapse versus recent infection, and multiple drug resistance.

The availability of the complete genomic sequence of two strains of *M. tuberculosis* together with new technologies, such as DNA microarrays and computational biology, has enabled a whole-genome perspective on genomic content, gene regulation, and *M. tuberculosis* metabolism (4, 6, 9, 12, 15, 16). Whole-genome DNA typing methods, such as fluorescent amplified fragment length polymorphisms (FAFLP) (1, 2, 13, 14), have made it feasible to robustly analyze isolates from different patient populations at the genome sequence interface (2). Here we describe the data supporting genotypic diversity of natural strains of *M. tuberculosis* from patients with and without human immunodeficiency virus type 1 (HIV-1) infection. Using FAFLP, we performed high-resolution genotypic analysis of epidemiologically related *M. tuberculosis* strains from independent patient populations with and without AIDS.

All isolates were cultured from unrelated tuberculosis patients with and without AIDS, treated at two different hospitals (Hospital Dos de Mayo and Hospital Maria Auxiliadora) in Lima, Peru, from May 1999 to September 2000. All the patients with AIDS were diagnosed for HIV-1 seropositivity with an AIDS-defining disease and generalized tuberculosis. HIV-1-seronegative patients presented with more local pulmonary tuberculosis. No data were available on the possible contacts of these patients with each other. The patients in HIV-1 and non-HIV-1 categories were mainly from Lima, Peru, but some patients were from other localities in Peru. Patient clinical history suggested that all patients belonged to different families and were not directly related to each other. These patients were from a very young and sexually active age group (average age, 30 years) with high mobility due to business or employment links and frequent travel.

M. tuberculosis isolates were cultured on Lowenstein-Jensen medium and characterized as described previously (8). Standard reference strains *M. tuberculosis* H37Rv (virulent strain) and *M. tuberculosis* Erdman were used, and *Mycobacterium bovis* AN5, *M. bovis* ATCC 27290, *M. bovis* ATCC 27291, *M. bovis* BCG, seal bacillus, *Mycobacterium microti*, and *Mycobacterium africanum* were also studied as controls at various levels. Genomic DNA was prepared from all the isolates by a standard method described elsewhere (2) and processed for FAFLP analysis (1, 2, 13).

Comparative analysis of FAFLP data was performed, using the whole-genome sequence of *M. tuberculosis* H37Rv (9). On the basis of the results of computer modeling with certain restriction enzymes, the *M. tuberculosis* H37Rv sequence data (9) were grouped into various size categories (in base pairs). The Genotyper software (Applied Biosystems) was trained on these categories to allow comparison of FAFLP fragment data of clinical isolates. On the basis of the presence or absence of monomorphic and polymorphic bands or peaks, different FAFLP profiles were identified as amplitypes. Bands were sized and genotyped for all the isolates within the user-defined categories (allele bins) of marker size (in base pairs). Presence

* Corresponding author. Mailing address: Centre for DNA Fingerprinting and Diagnostics (CDFD), Ecil Rd., Nacharam, Hyderabad 500076, India. Phone: 91 40 27155604. Fax: 91 40 27155610. E-mail: director@cdfd.org.in.

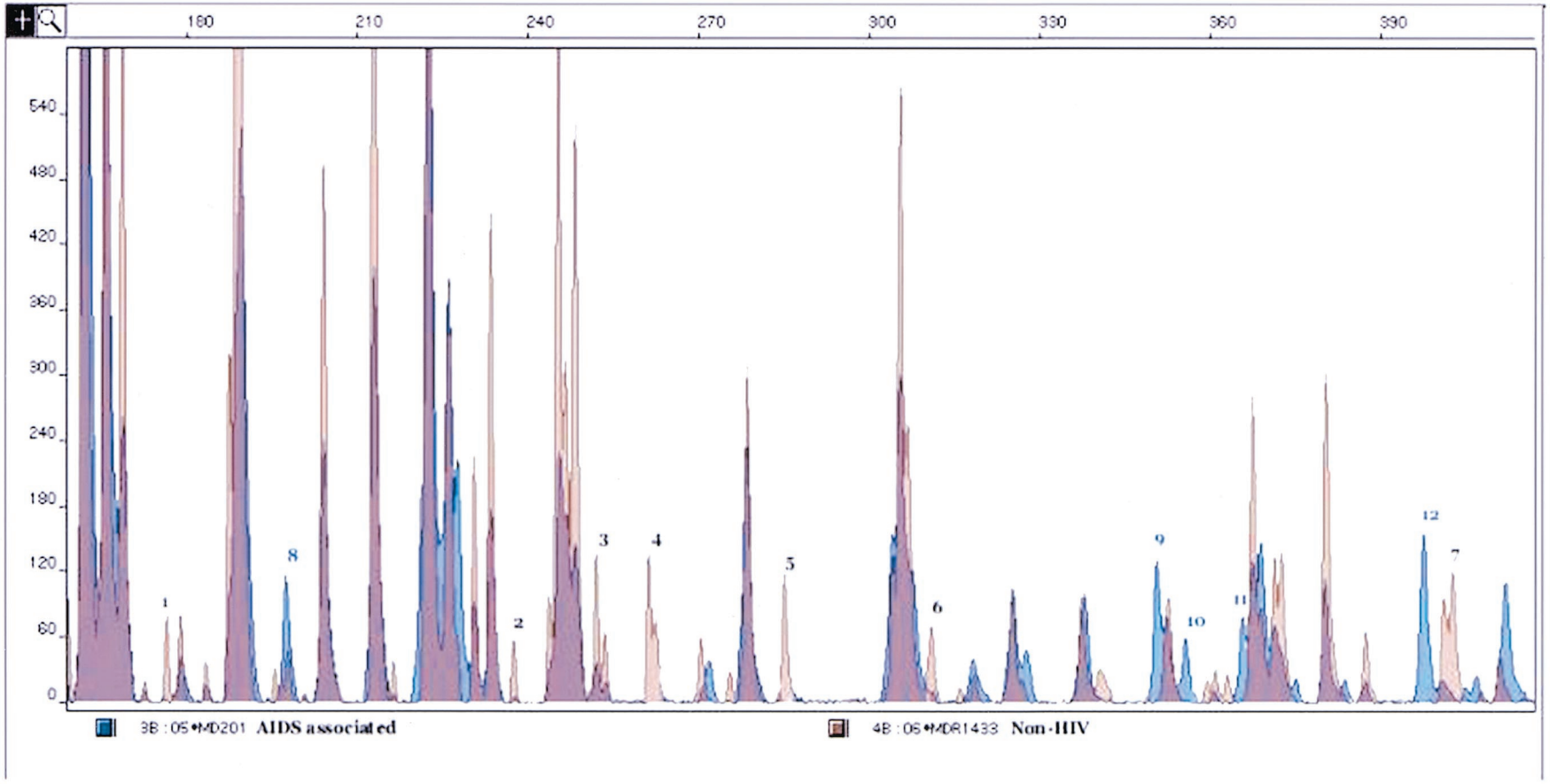

FIG. 1. Genescan-derived FAFLP profiles from isolates from patients with and without AIDS in Lima, Peru. One representative isolate from each patient group was used. FAFLP profiles with similar gel mobility conditions with equal data points (one data point being the relative position of a DNA fragment on the two-dimensional gel image) were color coded and superimposed to visualize differentially amplified fragments (visible as peaks; peak height indicates the quantity of amplicon generated, and peak position indicates size [in base pairs]). Fragments unique to each of the two isolates are marked with numbers at the top of the peak. The horizontal scale indicates size (in base pairs), while the vertical scale signifies level of fluorescence incorporated (peak intensity).

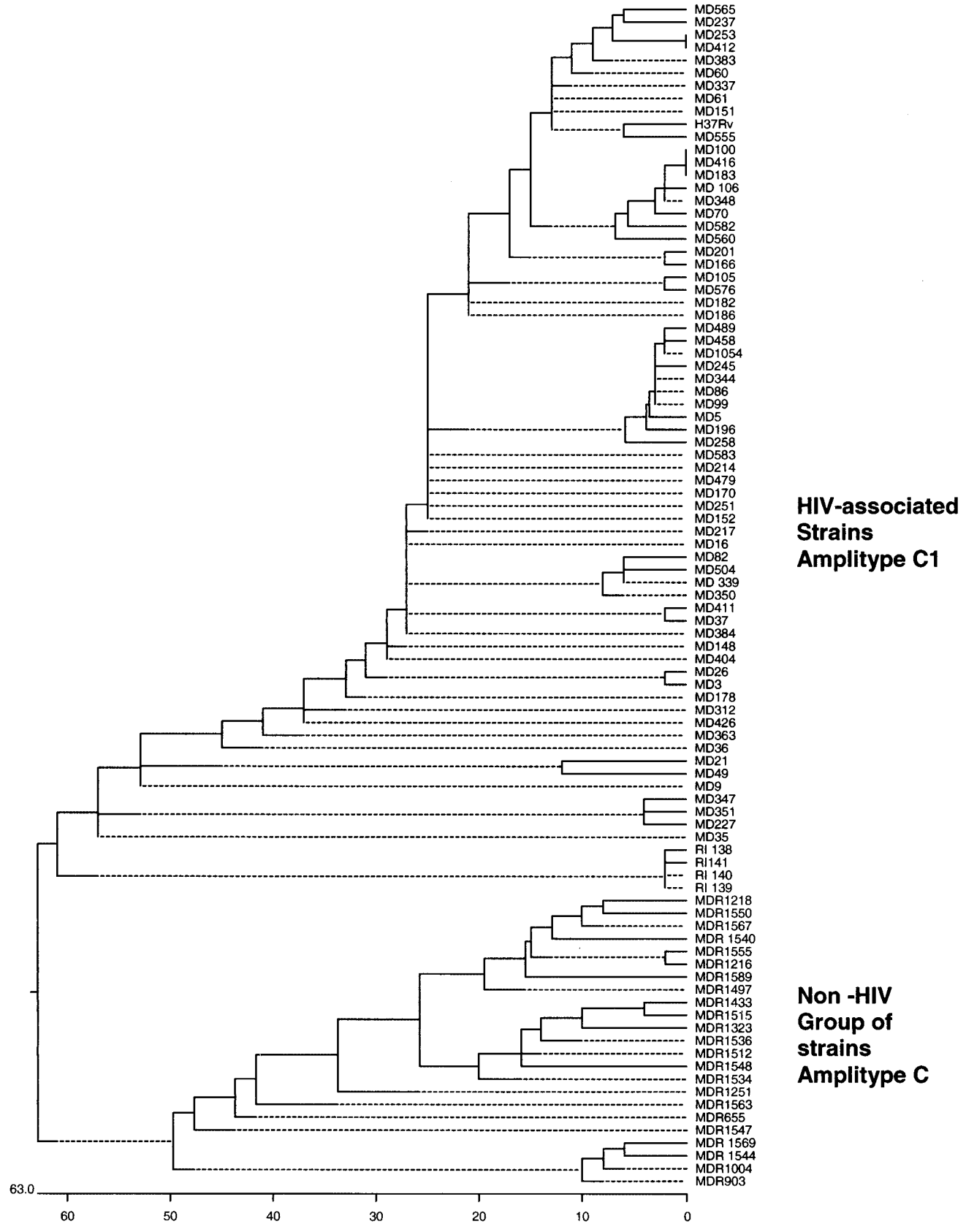

FIG. 2. Phylogenetic tree deduced from the Genotyper data based on the presence or absence of certain genomic loci in two groups of isolates. The scale at the bottom of the figure indicates genetic distance between the isolates. The sources of the isolates are indicated by the following prefixes: MD, isolates from AIDS patients from Lima; MDR, isolates from immunocompetent patients with pulmonary tuberculosis; RI, isolates from epidemiologically unrelated patients.

or absence of markers within the categories was scored by a user-defined Genotyper template that generated final output in the form of a binary table for all samples. Phylogenetic trees were generated from the binary data to establish genetic affinities and divergence among various amplitypes.

Predictive computer methods used on genome sequences of *M. tuberculosis* H37Rv and *M. tuberculosis* CDC1551 revealed a total of 136 fragments of sizes between 50 to 500 bp upon digestion with *Mse*I and *Eco*RI enzymes. Both the AIDS-associated and non-HIV-1-associated isolates were subjected to genotypic analysis, and the results were extrapolated to the computer-predicted data of the H37Rv sequence. Besides differential amplification of 12 unknown genomic regions (Fig. 1), most of the isolates revealed genotypic differences corresponding to polymorphisms mapped to a member of the PPE (proline-proline-glutamate) family of genes (Rv3343c), a conserved hypothetical protein (Rv3902c), and the phosphate transporter gene Rv0929 (PstC2). The polymorphisms observed were quite reproducible, and none of the fragment sizes varied by more than 1 bp. Some isolates in this study carried fewer IS*6110* motifs or carried only a single motif, although most of the isolates carried multiple copies of the element (6 to 19 copies). Overall, no correlation of amplitypes with IS*6110* copy number was seen in our study.

When FAFLP profiles were compared, the highest levels of relatedness were found among the isolates from the same disease group. The isolates were found to group (Fig. 2) in two distinct clusters, with a maximum genetic distance of 63%. This clustering was based on the differences in chromosomal fingerprints of the isolates from both groups, with at least 12 fragments of different sizes on genotyping (Fig. 1). The data may therefore indicate two different strain types responsible for the disease prevalence in Lima, Peru. Clonal groupings were observed among the AIDS-associated isolates; all the 25 isolates from HIV-1-negative individuals were found to be relatively heterogeneous and nonclonal.

Our observations contrast with those of Yang et al. (21) who failed to show significant differences between the IS*6110*-based genotypes of HIV-1-associated isolates and those from immunocompetent patients in Tanzania. IS*6110* typing data are often difficult to interpret due to inherent problems with this method (17). Also, the mobility of the insertion element in the chromosome does not reflect metabolic and recent evolutionary status of the strain. In contrast, FAFLP typing data has been successfully used for differentiating *M. tuberculosis* complex to the level of a subspecies (2). The presence or absence of a single fragment in FAFLP analysis cannot be viewed in the same way as an IS*6110* profile or a band shift by some other typing technique, such as pulsed-field gel electrophoresis. The DNA bands in FAFLP analysis can be tracked easily, and genome coordinates can be assigned easily. This is particularly true for a highly defined genome like *M. tuberculosis* where members of the *M. tuberculosis* complex show 99% similarity and despite identification of certain regions of differences, base substitutions in structural genes are almost negligible (6, 7, 13, 16, 22). Therefore, a single FAFLP locus may help establish the identity of a new strain if the epidemiological and clinical data support FAFLP data. This is particularly important in case of an HIV coinfection where accurate determination of a strain is more important to identify recent and resurgent infections.

We believe that the origin and molecular basis of various base modifications responsible for the clustering we found will be an important issue in the context of various stress groups and disease conditions. Our limited understanding of the genetics of *M. tuberculosis* under different stress conditions and selection pressures makes it difficult to predict which genetic polymorphisms may be of consequence. For example, while it is now relatively easy to detect genomic plasticity regions, mutations as small as a single base substitution in a single transcription factor may alter virulence in different hosts. The FAFLP patterns observed in this study should help us understand the spread and partitioning of *M. tuberculosis* genotypes among hosts with different immune status and at the same time identify informative epidemiological markers for different *M. tuberculosis* populations infecting AIDS patients.

While a statistically significant association of a clone to the prevalence of disease in a community may reflect increased adaptation and fitness of the strains, we suggest that this phenomenon should be observed in multiple patients and settings before they can be linked with certainty to a particular host population. Frequent travel and social links among the people of the economically productive and sexually active age group of Lima, Peru, studied raises concern that the *M. tuberculosis* strain type identified in this study may be circulating in other cities of Peru among mobile youth with HIV-1 infection. A nationwide surveillance study with additional FAFLP markers may be an appropriate approach to determine the spread patterns of *M. tuberculosis* in Peru.diagnosis of *H. capsulatum*.

This work was supported in part by the Department of Biotechnology of the Government of India and the International Society for Infectious Diseases (ISID).

We thank Yogesh Prasad (Applied Biosystems/Labindia) for assistance.

REFERENCES

1. **Ahmed, N., A. Bal, A. A. Khan, M. Alam, A. Kagal, V. Arjunwadkar, A. Rajput, A. A. Majeed, S. A. Rahman, S. Banerjee, S. Joshi, and R. Bharadwaj.** 2002. Whole-genome fingerprinting and genotyping of multiple drug resistant (MDR) isolates of *Pseudomonas aeruginosa* from endophthalmitis patients in India. Infect. Genet. Evol. **1:**237–242.
2. **Ahmed, N., M. Alam, A. A. Majeed, S. A. Rahman, A. Cataldi, D. Cousins, and S. E. Hasnain.** 2002. Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the *Mycobacterium tuberculosis* complex. Infect. Genet. Evol. **2:**193–199.
3. **Anh, D. D., M. W. Borgdorff, L. N. Van, N. T. N. Lan, T. van Gorkom, K. Kremer, and D. van Soolingen.** 2000. *Mycobacterium tuberculosis* Beijing genotype emerging in Vietnam. Emerg. Infect. Dis. **6:**302–305.
4. **Behr, M. A., M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolnik, S. Rane, and P. M. Small.** 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. Science **284:**1520–1523.
5. **Bifani, P. J., B. B. Plikaytis, and V. Kapur.** 1996. Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. JAMA **275:**452–455.
6. **Brosch, R., S. V. Gordon, A. Pym, K. Eiglmeier, T. Garnier, and S. T. Cole.** 2000. Comparative genomics of the mycobacteria. Int. J. Med. Microbiol. **290:**143–152.
7. **Brosch, R., S. V. Gordon, M. Marmiesse, P. Brodin, C. Buchrieser, K. Eiglmeier, T. Garnier, C. Gutierrez, G. Hewinson, K. Kremer, L. M. Parsons, A. S. Pym, S. Samper, D. van Soolingen, and S. T. Cole.** 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl. Acad. Sci. USA **99:**3684–3689.
8. **Caviedes, L., T. Lee, R. H. Gilman, P. Sheen, E. Spellman, E. H. Lee, D. E. Berg, S. Montenegro-James, and The Tuberculosis Working Group in Peru.** 2000. Rapid, efficient detection and drug susceptibility testing of *Mycobacterium tuberculosis* in sputum by microscopic observation of broth cultures. J. Clin. Microbiol. **38:**1203–1208.

9. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosi*s from the complete genome sequence. Nature **39:**3537–3544.
10. **Daley, C. L., P. M. Small, G. F. Schecter, G. K. Schoolnik, R. A. McAdam, W. R. Jacobs, Jr., and P. C. Hopewell.** 1992. An outbreak of tuberculosis with the human immunodeficiency virus: an analysis using restriction fragment length polymorphisms. N. Engl. J. Med. **326:**231–235.
11. **Edlin, B. R., J. I. Tokars, M. H. Grieco, J. T. Crawford, J. Williams, E. M. Sordillo, K. R. Ong, J. O. Kilburn, S. W. Dooley, K. G. Castro, W. R. Jarvis, and S. D. Holmberg.** 1992. An outbreak of multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome. N. Engl. J. Med. **326:**1514–1521.
12. **Gordon, S. V., R. Brosch, A. Billault, T. Garnier, K. Eiglmeier, and S. T. Cole.** 1999. Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. Mol. Microbiol. **32:**643–655.
13. **Goulding, J. N., J. Stanley, N. Saunders, and C. Arnold.** 2000. Genome-sequence-based fluorescent amplified-fragment length polymorphism analysis of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **38:**1121–1126.
14. **Janssen, P., R. Coopman, G. Huys, J. Swings, M. Bleeker, P. Vos, M. Zabeau, and K. Kersters.** 1996. Evaluation of DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. Microbiology **142:**1881–1893.
15. **Kato-Maeda, M., P. J. Bifani, B. N. Kreiswirth, and P. M. Small.** 2001. The nature and consequence of genetic variability in *Mycobacterium tuberculosis*. J. Clin. Investig. **107:**533–537.
16. **Kato-Maeda, M., J. T. Rhee, T. R. Gingeras, H. Salamon, J. Drenkow, N. Smittipat, and P. M. Small.** 2001. Comparing genomes within the species *Mycobacterium tuberculosis*. Genome Res. **11:**547–554.
17. **Siddiqi, N., M. Shamim, A. Amin, D. S. Chauhan, R. Das, K. Srivastava, D. Singh, V. D. Sharma, V. M. Katoch, S. K. Sharma, M. Hanief, and S. E. Hasnain.** 2001. Typing of drug resistant isolates of *Mycobacterium tuberculosis* from India using the IS6110 element reveals substantive polymorphism. Infect. Genet. Evol. **1:**109–113.
18. **Siddiqi, N., M. Shamim, S. Hussain, R. K. Choudhary, N. Ahmed, Prachee, S. Banerjee, G. R. Savithri, M. Alam, N. Pathak, A. Amin, M. Hanief, V. M. Katoch, S. K. Sharma, and S. E. Hasnain.** 2002. Molecular characterization of multidrug-resistant isolates of *Mycobacterium tuberculosis* from patients in north India. Antimicrob. Agents Chemother. **46:**443–450.
19. **Small, P. M., R. W. Shafer, P. C. Hopewell, S. P. Singh, M. J. Murphy, E. Desmond, M. F. Sierra, and G. K. Schoolnik.** 1993. Exogenous reinfection with multidrug-resistant *M. tuberculosis* in patients with advanced HIV infection. N. Engl. J. Med. **328:**1137–1144.
20. **van Soolingen, D. L. Qian, P. E. de Haas, J. T. Douglas, H. Traore, F. Portaels, H. Z. Qing, D. Enkhsaikan, P. Nymadawa, and J. D. van Embden.** 1995. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. J. Clin. Microbiol. **33:**3234–3238.
21. **Yang, Z. H., I. Mtoni, M. Chonde, M. Mwasekaga, K. Fuursted, D. S. Askgard, J. Bennedsen, P. E. W. de Haas, D. van Soolingen, J. D. A. van Embden, and A. B. Andersen.** 1995. DNA fingerprinting and phenotyping of *Mycobacterium tuberculosis* isolates from human immunodeficiency virus (HIV)-seropositive and HIV-seronegative patients in Tanzania. J. Clin. Microbiol. **33:**1064–1069.
22. **Yeh, R. W., P. De Leon, C. B. Agasino, J. A. Hahn, C. L. Daley, P. C. Hopewell, and P. M. Small.** 1998. Stability of *Mycobacterium tuberculosis* DNA genotypes. J. Infect. Dis. **177:**1107–1111.