Comparison of Restriction Fragment Length Polymorphism with the Polymorphic Guanine-Cytosine-Rich Sequence and Spoligotyping for Differentiation of *Mycobacterium tuberculosis* Isolates with Five or Fewer Copies of IS6110[▽]

Laura Flores,¹ Leah G. Jarlsberg,¹ Elizabeth Y. Kim,¹ Dennis Osmond,² Jennifer Grinsdale,³ Masae Kawamura,³ Ed Desmond,⁴ Philip C. Hopewell,¹ and Midori Kato-Maeda^{1*}

Francis J. Curry National Tuberculosis Center, Division of Pulmonary and Critical Care Medicine, San Francisco General Hospital, Department of Medicine, University of California, San Francisco, San Francisco, California¹; Department of Epidemiology and Biostatistics University of California, San Francisco, San Francisco, California²; and Tuberculosis Control Section, San Francisco Department of Public Health,³ and Microbial Diseases Laboratory, California Department of Public Health,⁴ Richmond, California

Received 18 August 2009/Returned for modification 17 October 2009/Accepted 14 December 2009

The use of IS6110 as a marker for molecular epidemiological studies is limited when a *Mycobacterium tuberculosis* isolate has five or fewer copies of IS6110. Restriction fragment length polymorphism analysis with a highly polymorphic GC-rich repetitive sequence located in the plasmid pTBN12 (PGRS RFLP) and spoligotyping (based on the polymorphism of the DR region) are two frequently used secondary typing methods. The aim of this study was to compare the performance of these two methods in a population-based study in San Francisco. We included all patients with culture-positive tuberculosis from 1999 to 2007 with IS6110 RFLP results presenting five or fewer bands. PGRS RFLP and spoligotyping were performed using standardized methods. We determined the concordance between the two methods regarding cluster status and the risk factors for an isolate to be in a cluster with each of the methods. Our data indicate that both methods had similar discriminatory power and that the risk factors associated with clustering by either method were the same. Although the cluster/unique status was concordant in 84% of the isolates, patients were clustered differently depending on the method. Therefore, the methods are not interchangeable, and the same method should be used for longitudinal studies.

Genotyping of *Mycobacterium tuberculosis* together with conventional epidemiologic investigations is used to elucidate the dynamics of tuberculosis transmission in populations (8, 18). Restriction fragment length polymorphism analysis (RFLP) of insertion sequence 6110 (IS6110) is the most frequently used and best-validated genotyping method (28). Most isolates of *M. tuberculosis* contain between 0 and 25 copies of IS6110 (19). Polymorphisms of IS6110 are based on the variation in the number of copies and the molecular weights of the fragments in which IS6110 is located, as determined by Southern blot hybridization. In population-based studies, isolates that share the same IS6110 RFLP pattern are considered to be clustered and are assumed to be epidemiologically linked (that is, to be part of a chain of recent transmission), although the link may be indirect and remote in time (9, 14, 24, 25). In contrast, cases with isolates having IS6110 RFLP patterns not shared by other isolates within the population are considered to have resulted from reactivation of latent infection, presumably acquired either outside the population or prior to the initiation of genotyping.

IS6110 RFLP has been used to confirm suspected transmis-

sion links and to identify previously unidentified links (9, 10, 17, 20, 27). The usefulness of IS6110 RFLP in this regard depends on the ability of the method to discriminate among strains of *M. tuberculosis*. However, it has been demonstrated that there is a significantly higher rate of clustering and a lower proportion of patients with confirmed epidemiological links among isolates with five or fewer copies of IS6110, suggesting lower discriminatory power (6, 22, 29). Additionally, the transposition rate of the IS6110 element has been shown to be lower in *M. tuberculosis* isolates with a low number of IS6110 copies than in other isolates, thereby limiting the discriminatory capacity of IS6110 RFLP (6). For this reason, a secondary typing method, such as RFLP using the polymorphic guanine-cytosine-rich sequence (PGRS), is generally used to increase the discriminatory power of the IS6110 RFLP in isolates with five or fewer bands (22). PGRS is a highly polymorphic GC-rich repetitive sequence located in the plasmid pTBN12 (6). The PGRS method is also based on the RFLP technique. Compared with IS6110 RFLP, PGRS RFLP produces many more bands with different intensities, complicating its reading.

Another secondary typing method is spacer oligonucleotide genotyping (spoligotyping), which is a PCR-based genotyping method (15) based on the polymorphism of a single direct repeat (DR) locus. It is relatively simple to perform spoligotyping and to analyze and share spoligotyping data, especially compared to PGRS RFLP data. Spoligotyping has been widely used for primary and secondary typing (1) and strain identifi-

* Corresponding author. Mailing address: San Francisco General Hospital, 1001 Potrero Ave, Box 0841, San Francisco, CA 94110. Phone: (415) 206-8121. Fax: (415) 695-1551. E-mail: midori.kato-maeda@ucsf.edu.

[▽] Published ahead of print on 23 December 2009.

TABLE 1. Concordance between IS6110/PGRS typing and IS6110/spoligotyping

IS6110/PGRS typing result	No. (%) of isolates with IS6110 spoligotyping result		Total
	Cluster	Unique	
Cluster	40 (85.1)	7 (14.9)	47
Unique	16 (16.5)	81 (83.5)	97
Total	56	88	144

McNemar's test: $P = 0.09$, odds ratio = 0.44 (95% confidence interval, 0.15 to 1.124).

cation, and there is a large pattern database with worldwide information (3). However, it has limited discriminatory power, especially in some lineages of *M. tuberculosis* (18).

In this paper, we compare the performance of PGRS RFLP with spoligotyping for secondary typing of *M. tuberculosis* isolates with five or fewer IS6110 bands to examine the community epidemiology of tuberculosis in a population-based study. We analyze the implications of the results from each of the methods at the public health level and at the individual patient level.

MATERIALS AND METHODS

We used a population-based collection of *M. tuberculosis* DNA from an ongoing study that has been conducted in San Francisco since 1991. We included isolates from patients with culture-positive tuberculosis from 1999 to 2007. All isolates included had five or fewer IS6110 bands and had PGRS RFLP and spoligotyping results available.

The RFLP using IS6110 and PGRS was performed according to standardized methods (22, 28). The RFLP patterns were scanned using Advanced Quantifier (AQ) for Windows (P/N 200100; Bio Image Systems, Inc., Jackson, MI). The band assignment was reviewed by two independent readers and the cluster designation was confirmed visually.

Spoligotyping of *M. tuberculosis* isolated between 1999 and 2004 was performed with standardized methods (15) using a membrane with 43 oligonucleotides representing the different direct variable repeats (DVRs) immobilized in specific locations (Ocimum Biosolutions, Inc., Gaithersburg, MD). Beginning in 2004, spoligotyping of the isolates was performed by the Microbial Diseases Laboratory, California Department of Public Health, as part of the Tuberculosis Genotyping Program of the Centers for Disease Control and Prevention (CDC) (5). The same 43 DVRs were interrogated using Luminex technology for detection (7).

Definitions. Isolates clustered by IS6110/PGRS typing (i.e., by IS6110 and PGRS RFLP patterns) are defined as two or more isolates with an identical genotype (same number of bands and less than 3% band size difference for each band in the IS6110 and PGRS RFLP patterns) from separate patients within a 1-year period prior to the date the second specimen was collected. For each cluster, the initial case was operationally considered to be the source case and not considered to be the result of recent transmission and rapid progression (within the 1-year window) to active tuberculosis. All other cases in the cluster were considered to be secondary to the initial case.

Isolates clustered by IS6110/spoligotyping (i.e., IS6110 RFLP pattern and spoligotype) are defined similarly to the previous term, but the isolates have an identical IS6110 RFLP pattern and an identical spoligotype identified within the 1-year window.

Analysis. We used a chi square test to determine the risk factors for an isolate (patient) to be in a cluster as determined by each of the two methods. We used McNemar's test for paired data to determine if, among those patients with risk factors known to be associated with clustering, clustering was more likely by one method or the other.

RESULTS

Between 1999 and 2007, there were 1,011 clinical isolates of *M. tuberculosis* with IS6110 RFLP results available. Of these, 192 (18.9%) had five or fewer IS6110 bands; 180 (90.1%) of these had

TABLE 2. Risk factors for cluster by genotyping method

Risk factor	Cluster by:			
	IS6110/PGRS typing		IS6110/spoligotyping	
	Odds ratio	95% confidence interval	Odds ratio	95% confidence interval
U.S. birth	11.97	4.35–32.9	9.53	4.02–22.5
HIV positivity	2.46	1.15–5.23	3.38	1.58–7.19
Homelessness	7.14	3.27–15.5	9.93	4.55–21.6
Smear positivity	1.05	0.51–2.15	0.88	0.44–1.76
Chest X ray with cavitation	1.40	0.56–3.52	0.81	0.32–2.06
Pulmonary tuberculosis	5.84	1.30–26.1	7.94	1.78–35.4

PGRS RFLP determination, and the spoligotype was available for 170 (94.4%). The isolates from 1999 ($n = 26$) were used as a reference for the cluster designation (see the definitions above), leaving 144 isolates to provide the comparison between PGRS RFLP pattern and spoligotyping as secondary typing methods.

Based on only IS6110 RFLP pattern, there was a total of 96 secondary cases from 10 different clusters. The number of secondary cases per cluster varied from 1 to 60. When IS6110 and PGRS typing were used, the number of total secondary cases decreased to 47, and they were distributed in 10 different clusters. The number of secondary cases per cluster also decreased, ranging from 1 to 24 cases. There were 56 total secondary cases according to IS6110/spoligotyping, and these were distributed in 13 different clusters. The number of secondary cases per cluster varied from 1 to 15.

For 121 (84%) of the 144 isolates, the cluster/unique status was concordant, and for 23 (16%) isolates, there was disagreement between the two methods (Table 1) (McNemar's test, $P = 0.09$). When the clustered cases were analyzed by IS6110/PGRS typing, only 10 secondary cases from five clusters were also considered secondary cases by IS6110/spoligotyping. When clustered cases were analyzed by IS6110/spoligotyping, only 4 secondary cases from two clusters were also considered secondary cases by IS6110/PGRS typing. Stated differently, although the cluster/unique status of the isolates was concordant in 84% of the cases, the patients were clustered with different patients depending on the method.

Both genotyping methods identified the same risk factors for being in a cluster, based on bivariate analysis (Table 2). To determine whether any of the risk factors associated with clustering in Table 2 were more likely to be identified by one method or the other, we first stratified the data by the presence of a risk factor and then tested the resulting 2×2 table by McNemar's test for paired data (Table 3). We found that the number of isolates clustered by IS6110/PGRS typing and unique by IS6110/spoligotyping was not statistically different from the number of strains that were unique by IS6110/PGRS genotyping and clustered by IS6110/spoligotyping. This may be due to the lack of power to detect differences because of the low number of discordant cases. Taken together, the results from Tables 2 and 3 suggest that both methods will result in the same risk factors associated with clustering.

TABLE 3. McNemar's test for paired data to determine if any of the factors associated with cluster were more likely to result in isolates clustered by one method than the other

Risk factor for cluster	No. of isolates			Odds ratio	95% confidence interval	McNemar's <i>P</i> value
	Total	Cluster by <i>IS6110</i> /PGRS typing but unique by <i>IS6110</i> /spoligotyping	Cluster by <i>IS6110</i> /spoligotyping but unique by <i>IS6110</i> /PGRS typing			
Foreign birth	62	3	6	0.5	0.08–2.34	0.5
Pulmonary TB	122	6	15	0.4	0.13–1.09	0.08
Smear positivity	54	3	5	0.6	0.09–3.08	0.72
HIV positivity	40	3	8	0.37	0.06–1.56	0.23
Homelessness	60	3	10	0.3	0.053–1.6	0.09
X ray with cavities	23	2	1	2	0.1–118	1
Any risk factor	130	7	16	0.44	0.15–1.12	0.09

DISCUSSION

The results of this study indicate that a secondary typing method applied to isolates of *M. tuberculosis* having five or fewer *IS6110* hybridizing bands improves the discriminatory power of *IS6110* RFLP. However, although there were concordant results between the two methods for 84% of the isolates overall (85.1% concordance for clustered isolates and 83.5% concordance for unique isolates), *IS6110*/spoligotyping identified more isolates in clusters (secondary cases) and more clusters than *IS6110*/PGRS typing. Unfortunately, there is no way to determine which of these results is “correct” because there is no “gold standard” method against which to compare these two methods. Adding more genotyping methods will increase the discriminatory power (fewer clustered cases) but will not necessarily establish which combination of methods is the “correct combination.” In fact, when we analyzed isolates clustered by *IS6110*/PGRS typing but not by *IS6110*/spoligotyping with mycobacterial interspersed repetitive unit (MIRU) genotyping data ($n = 18$) (26), we found that in all cases, the MIRU genotyping result was in agreement with the *IS6110*/PGRS typing result. When we analyzed isolates clustered by *IS6110*/spoligotyping but not by *IS6110*/PGRS typing with MIRU genotyping data ($n = 7$), we found that in all cases the MIRU information was in agreement with *IS6110*/spoligotyping data. Although the MIRU data were available for a small sample, these results exemplify the complexity of the interpretation of the genotyping data.

Our results confirmed the findings by Yang et al., who found that *IS6110*/PGRS typing was more discriminative (fewer clustered cases) than *IS6110*/spoligotyping (32). However, the published study did not include either information about the risk factors for clustering or data about patients being clustered with the same patient depending on the method.

Fortunately, because in most population-based studies the proportion of cases with isolates that have five or fewer copies of *IS6110* is low, the impact of these cases in the study of the overall transmission of tuberculosis in a community will be low. For example, in San Francisco, the proportion of cases with five or fewer copies of *IS6110* is less than 20%, and the analysis of risk factors for tuberculosis due to recent transmission is not affected by the inclusion of these cases (22).

The results from this study should be considered from two perspectives. The first is related to the use of these markers in

population-based studies to describe the community pathogenesis of tuberculosis (tuberculosis resulting mainly from reactivation of latent infection or as a result of recent transmission with rapid progression to active tuberculosis) and to identify risk factors for acquiring infection with rapid progression, as determined by clustered isolates (secondary cases) and the features of the secondary cases (2, 12, 13, 21, 24). Our study demonstrated that *IS6110*/spoligotyping and *IS6110*/PGRS genotyping identified roughly similar proportions of clustered (secondary) cases as well as the same risk factors for clustering.

The second perspective is related to the individual patient with tuberculosis. We demonstrated that although 84% of the isolates had concordant results, the presumed secondary cases resulting from the index case were different, depending on the genotyping method used. Although being in a different cluster will not affect the management of the patient (i.e., patients with risk factors for being in a cluster will always be considered priority for contact investigation) (4), it may have an impact on the contact investigation procedure, as public health personnel may use the genotyping data to inform the direction and intensity of the contact investigation to search for epidemiological links and to find other contacts that may be infected or have active tuberculosis.

The inconsistent results obtained by these two secondary methods may be due in part to the different molecular clocks of the PGRS elements and the DR region from which different genetic polymorphisms originate. The DR region evolves mainly by *IS6110*-mediated mutation, homologous recombination between repeat sequences that leads to deletion of DVRs, strand slippage that leads to duplication of DVRs, and point mutation (30). In isolates with low *IS6110* copy numbers, the DR region evolves mainly by homologous recombination (11). It has been estimated that the rate of change of the DR region is lower than that of *IS6110*, but the rate will be different depending on the genetic background of the strain (i.e., the strain family) (30). In San Francisco, most isolates with five or fewer copies of *IS6110* bands are strains from the Euro-American lineage, which is known to have sufficient spoligotyping polymorphism to be discriminatory (L. Flores et al., personal communication). The evolution of PGRS-containing regions includes duplication, recombination, and strand slippage (16); however, these mutations rarely induce changes that can be observed in RFLP using PGRS elements (23, 31). As with spoligotyping, it has been estimated that the rate of change of PGRS is lower than that of *IS6110* (23, 31).

As noted before, the main limitation of this study is the lack of a gold standard. Therefore, we cannot determine which of the two genotyping combinations (IS6110/PGRS typing versus IS6110/spoligotyping) is better for identifying secondary cases.

However, the discriminatory powers of the two methods are similar in our population, and therefore, the methods are comparable for making inferences in a population-based study. The differences we found do, however, indicate that the methods are not interchangeable and that for longitudinal studies, one method or the other should be used consistently.

ACKNOWLEDGMENTS

We express our appreciation to the staff of the San Francisco Department of Public Health, Tuberculosis Control Section, the Mycobacteriology Section, San Francisco Department of Public Health Laboratory and the Microbial Diseases Laboratory, California Department of Public Health.

This study was supported by a grant from the National Institutes of Health (AI 034238).

REFERENCES

- Bauer, J., A. B. Andersen, K. Kremer, and H. Miorner. 1999. Usefulness of spoligotyping to discriminate IS6110 low-copy-number *Mycobacterium tuberculosis* complex strains cultured in Denmark. *J. Clin. Microbiol.* **37**:2602–2606.
- Borgdorff, M. W., M. A. Behr, N. J. Nagelkerke, P. C. Hopewell, and P. M. Small. 2000. Transmission of tuberculosis in San Francisco and its association with immigration and ethnicity. *Int. J. Tuberc. Lung Dis.* **4**:287–294.
- Brudey, K., J. R. Driscoll, L. Rigouts, W. M. Proding, A. Gori, S. A. Al-Hajjaj, C. Allix, L. Aristimuno, J. Arora, V. Baumanis, L. Binder, P. Cafrune, A. Cataldi, S. Cheong, R. Diel, C. Ellermeier, J. T. Evans, M. Fauville-Dufaux, S. Ferdinand, D. Garcia de Viedma, C. Garzelli, L. Gazzola, H. M. Gomes, M. C. Gutierrez, P. M. Hawkey, P. D. van Helden, G. V. Kadival, B. N. Kreiswirth, K. Kremer, M. Kubin, S. P. Kulkarni, B. Liens, T. Lillebaek, M. L. Ho, C. Martin, I. Mokrousov, O. Narvskaia, Y. F. Ngeow, L. Naumann, S. Niemann, I. Parwati, Z. Rahim, V. Rasolof-Razanamparany, T. Rasolonalana, M. L. Rossetti, S. Rusch-Gerdes, A. Sajduda, S. Samper, I. G. Shenyakin, U. B. Singh, A. Somoskovi, R. A. Skuce, D. van Soolingen, E. M. Streicher, P. N. Suffys, E. Tortoli, T. Tracevska, V. Vincent, T. C. Victor, R. M. Warren, S. F. Yap, K. Zaman, F. Portaels, N. Rastogi, and C. Sola. 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol.* **6**:23.
- CDC. 2005. Guidelines for the investigation of contacts of persons with infectious tuberculosis. Recommendations from the National Tuberculosis Controllers Association and CDC. *MMWR Recomm. Rep.* **54**:1–47.
- CDC. 2005. Notice to readers: new CDC program for rapid genotyping of *Mycobacterium tuberculosis* isolates. *MMWR Morb. Mortal. Wkly. Rep.* **54**:47.
- Chaves, F., Z. Yang, H. el Hajj, M. Alonso, W. J. Burman, K. D. Eisenach, F. Pronda, J. H. Bates, and M. D. Cave. 1996. Usefulness of the secondary probe pTBN12 in DNA fingerprinting of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **34**:1118–1123.
- Cowan, L. S., L. Diem, M. C. Brake, and J. T. Crawford. 2004. Transfer of a *Mycobacterium tuberculosis* genotyping method, spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multi-analyte profiling system. *J. Clin. Microbiol.* **42**:474–477.
- Daley, C. L. 2005. Molecular epidemiology: a tool for understanding control of tuberculosis transmission. *Clin. Chest Med.* **26**:217–231.
- 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 accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *N. Engl. J. Med.* **326**:231–235.
- Dewan, P. K., H. Banoung, N. Abernethy, T. Hoynes, L. Diaz, M. Wolde-mariam, T. Ampie, J. Grinsdale, and L. M. Kawamura. 2006. A tuberculosis outbreak in a private-home family child care center in San Francisco, 2002 to 2004. *Pediatrics* **117**:863–869.
- Fang, Z., N. Morrison, B. Watt, C. Doig, and K. J. Forbes. 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.
- Garcia-Garcia, M. L., A. Ponce de Leon, M. E. Jimenez-Corona, A. Jimenez-Corona, M. Palacios-Martinez, S. Balandrano-Campos, L. Ferreyra-Reyes, L. Juarez-Sandino, J. Sifuentes-Osorio, H. Olivera-Diaz, J. L. Valdespino-Gomez, and P. M. Small. 2000. Clinical consequences and transmissibility of drug-resistant tuberculosis in southern Mexico. *Arch. Intern. Med.* **160**:630–636.
- Heldal, E., U. R. Dahle, P. Sandven, D. A. Caugant, N. Brattaas, H. T. Waaler, D. A. Enarson, A. Tverdal, and J. Kongerud. 2003. Risk factors for recent transmission of *Mycobacterium tuberculosis*. *Eur. Respir. J.* **22**:637–642.
- Jasmer, R. M., J. A. Hahn, P. M. Small, C. L. Daley, M. A. Behr, A. R. Moss, J. M. Creasman, G. F. Schecter, E. A. Paz, and P. C. Hopewell. 1999. A molecular epidemiologic analysis of tuberculosis trends in San Francisco, 1991–1997. *Ann. Intern. Med.* **130**:971–978.
- Kamerbeek, J., L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, and J. van Embden. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J. Clin. Microbiol.* **35**:907–914.
- Karboul, A., N. C. Gey van Pittius, A. Namouchi, V. Vincent, C. Sola, N. Rastogi, P. Suffys, M. Fabre, A. Cataldi, R. C. Huard, N. Kurepina, B. Kreiswirth, J. L. Ho, M. C. Gutierrez, and H. Mardassi. 2006. Insights into the evolutionary history of tubercle bacilli as disclosed by genetic rearrangements within a PE_PGRS duplicated gene pair. *BMC Evol. Biol.* **6**:107.
- Kline, S. E., L. L. Hedemark, and S. F. Davies. 1995. Outbreak of tuberculosis among regular patrons of a neighborhood bar. *N. Engl. J. Med.* **333**:222–227.
- Mathema, B., N. E. Kurepina, P. J. Bifani, and B. N. Kreiswirth. 2006. Molecular epidemiology of tuberculosis: current insights. *Clin. Microbiol. Rev.* **19**:658–685.
- McEvoy, C. R., A. A. Falmer, N. C. Gey van Pittius, T. C. Victor, P. D. van Helden, and R. M. Warren. 2007. The role of IS6110 in the evolution of *Mycobacterium tuberculosis*. *Tuberculosis (Edinburgh)* **87**:393–404.
- McLaughlin, S. L., P. Spradling, D. Drociuk, R. Ridzon, C. J. Pozsik, and I. Onorato. 2003. Extensive transmission of *Mycobacterium tuberculosis* among congregated, HIV-infected prison inmates in South Carolina, United States. *Int. J. Tuberc. Lung Dis.* **7**:665–672.
- Moss, A. R., J. A. Hahn, J. P. Tulskey, C. L. Daley, P. M. Small, and P. C. Hopewell. 2000. Tuberculosis in the homeless. A prospective study. *Am. J. Respir. Crit. Care Med.* **162**:460–464.
- Rhee, J. T., M. M. Tanaka, M. A. Behr, C. B. Agasino, E. A. Paz, P. C. Hopewell, and P. M. Small. 2000. Use of multiple markers in population-based molecular epidemiologic studies of tuberculosis. *Int. J. Tuberc. Lung Dis.* **4**:1111–1119.
- Richardson, M., G. D. van der Spuy, S. L. Sampson, N. Beyers, P. D. van Helden, and R. M. Warren. 2004. Stability of polymorphic GC-rich repeat sequence-containing regions of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **42**:1302–1304.
- Small, P. M., P. C. Hopewell, S. P. Singh, A. Paz, J. Parsonnet, D. C. Ruston, G. F. Schecter, C. L. Daley, and G. K. Schoolnik. 1994. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N. Engl. J. Med.* **330**:1703–1709.
- 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 *Mycobacterium tuberculosis* in patients with advanced HIV infection. *N. Engl. J. Med.* **328**:1137–1144.
- Sola, C., I. Filliol, E. Legrand, S. Lesjean, C. Locht, P. Supply, and N. Rastogi. 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.
- Valway, S. E., M. P. Sanchez, T. F. Shinnick, I. Orme, T. Agerton, D. Hoy, J. S. Jones, H. Westmoreland, and I. M. Onorato. 1998. An outbreak involving extensive transmission of a virulent strain of *Mycobacterium tuberculosis*. *N. Engl. J. Med.* **338**:633–639.
- 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.
- van Soolingen, D., P. E. de Haas, P. W. Hermans, P. M. Groenen, and J. D. van Embden. 1993. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **31**:1987–1995.
- Warren, R. M., E. M. Streicher, S. L. Sampson, G. D. van der Spuy, M. Richardson, D. Nguyen, M. A. Behr, T. C. Victor, and P. D. van Helden. 2002. Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. *J. Clin. Microbiol.* **40**:4457–4465.
- Yang, Z., F. Chaves, P. F. Barnes, W. J. Burman, J. Koehler, K. D. Eisenach, J. H. Bates, and M. D. Cave. 1996. Evaluation of method for secondary DNA typing of *Mycobacterium tuberculosis* with pTBN12 in epidemiologic study of tuberculosis. *J. Clin. Microbiol.* **34**:3044–3048.
- Yang, Z. H., K. Ijaz, J. H. Bates, K. D. Eisenach, and M. D. Cave. 2000. Spoligotyping and polymorphic GC-rich repetitive sequence fingerprinting of mycobacterium tuberculosis strains having few copies of IS6110. *J. Clin. Microbiol.* **38**:3572–3576.