

Molecular Epidemiology of Tuberculosis in a Rural Area of High Prevalence in South India: Implications for Disease Control and Prevention

Sujatha Narayanan,¹ Sulochana Das,¹ Renu Garg,² Lalitha Hari,¹ Vijay Bhaskara Rao,¹ Thomas R. Frieden,² and P. R. Narayanan^{1*}

Tuberculosis Research Centre, Chennai,¹ and the World Health Organization, Stop Tuberculosis Unit, South East Asian Regional Office, New Delhi,² India

Received 28 May 2002/Returned for modification 9 July 2002/Accepted 10 September 2002

Molecular and conventional epidemiologic techniques were used to study the mechanisms and risk factors for tuberculosis transmission in a rural area with high prevalence in south India. Restriction fragment length polymorphism analysis with IS6110 and direct repeat probes was performed with 378 *Mycobacterium tuberculosis* isolates from patients. Forty-one percent of *M. tuberculosis* isolates harbored a single copy of IS6110. Of 378 patients, 236 had distinct strains; 142 (38%) shared a strain with other patients, indicating recent infection. Older patients, those detected by a house-to-house community survey, and those hospitalized in a sanatorium were more likely to have had a recent infection. These findings suggest that the majority of the tuberculosis cases in south India were due to reactivation; therefore, efforts to control tuberculosis should be sustained.

Each year, nearly 3.5 million new cases of infectious tuberculosis (TB) occur worldwide (11). Molecular epidemiologic studies of *Mycobacterium tuberculosis* strains by using DNA fingerprinting have provided valuable information about the mechanisms of transmission of TB. Most of these studies, however, have been conducted in countries with a low incidence of tuberculosis and usually in urban settings with pockets of high transmission. In such settings, 35 to 50% of new TB cases are the result of recent person-to-person or exogenous transmission (1, 12, 25). Few molecular epidemiologic studies have been conducted in countries with a high incidence of TB (10, 16).

India has the highest number of incident TB cases in the world. The limited numbers of molecular epidemiologic studies conducted in India were laboratory based and comprised small numbers of patients (9, 24). Risk factors for exogenous transmission of TB have not been documented in the country. This study combined molecular and conventional epidemiologic techniques to investigate the mechanisms and risk factors for TB transmission in a rural area of high prevalence in south India.

The study site was Tiruvallur District, where the world's largest *Mycobacterium bovis* BCG trial was conducted (28). The total population of the study area is 580,000. The incidence of smear-positive TB is 76 per 100,000 in the population (29). The study subjects were all TB patients undergoing treatment according to the Revised National Tuberculosis Control Program guidelines (21) at one of the 17 primary health centers in the study area from July 1999 to June 2000. Within a week of commencing treatment, two sputum samples were

collected from each patient for mycobacterial culture on Löwenstein-Jensen (LJ) medium. The samples that yielded positive cultures were further confirmed to be typical *M. tuberculosis* strains (2) and tested for drug susceptibility. The DNA extraction and restriction fragment length polymorphism (RFLP) analysis were performed by standard protocols (9, 18, 24, 30). Two investigators independently analyzed the patterns by visual inspection. For quality control, RFLP was repeated with an additional specimen for 75 patients (20%). The RFLP patterns were identical in both specimens of each of the 75 patients; hence, only one specimen from each patient was included for further analysis. Standard definitions were used to classify patients according to the type and category of TB (33). Patients who demonstrated identical RFLP patterns by IS6110 and direct repeat (DR) probes were considered to belong to the same cluster. Patients in a cluster were interviewed to investigate epidemiologic linkages. Patients were considered epidemiologically linked if they reported acquaintance with one or more patients in the same cluster. Data were analyzed with Epiinfo version 6.04 (Centers for Disease Control and Prevention, Atlanta, Ga.). Chi-square and Fisher exact tests were used to interpret univariate analyses.

Of 437 culture-positive patients treated during the study period, isolates from 378 were available for RFLP analysis. Characteristics of patients for whom RFLP data were unavailable ($n = 59$) were similar to those of patients for whom RFLP data were available ($n = 378$), but the former were more likely to have multidrug-resistant (MDR) TB (10% versus 2%; $P = 0.03$). Fifty-seven of the 378 (15%) patients included in the analysis could not be interviewed to investigate epidemiologic linkages.

RFLP analysis by IS6110 identified 4 strains (1%) with no copy, 154 strains (41%) with one copy, 99 strains (26%) with 2 to 5 copies, and 121 strains (32%) with >6 copies. Of the 378 patients, 225 (60%) carried strains shared by at least 1 other

* Corresponding author. Mailing address: Tuberculosis Research Centre, Mayor V. R. Ramanathan Rd. (Sputank Rd.), Chetput, Chennai 600 031, India. Phone: 91 (044) 8265425. Fax: 91 (044) 8262137. E-mail: nrparanj@md2.vsnl.net.in.

TABLE 1. IS6110 and DR RFLP patterns among 142 TB patients in a cluster in Tiruvallur, South India, July 1999 to June 2000

Lane no.	Pattern (no. of bands)		No. of patients in cluster
	IS6110 probe	DR probe	
1	No copy	B041 (4)	2
2	A (1)	B001 (4)	2
3	A (1)	B002 (5)	25
4	A (1)	B003 (5)	6
5	A (1)	B039 (5)	7
6	A (1)	B058 (5)	2
7	A (1)	B007 (6)	2
8	A (1)	B010 (6)	10
9	A (1)	B014 (6)	5
10	A (1)	B018 (6)	4
11	A (1)	B027 (6)	6
12	A (1)	B038 (6)	3
13	A (1)	B046 (6)	2
14	A (1)	B054 (6)	2
15	A (1)	B009 (7)	3
16	A (1)	B032 (7)	8
17	A (1)	B036 (7)	5
18	A (1)	B043 (7)	2
19	A (1)	B077 (7)	3
20	A (1)	B088 (7)	7
21	E (2)	B014 (6)	3
22	H (2)	B071 (4)	2
23	H (2)	B119 (6)	2
24	M (2)	B014 (6)	2
25	E (3)	B014 (6)	2
26	N (3)	B032 (7)	2
27	B (4)	B002 (5)	4
28	B (4)	B027 (6)	2
29	C (4)	B002 (5)	2
30	H (4)	B002 (5)	2
31	D (5)	B002 (5)	3
32	F (5)	B002 (5)	3
33	G (5)	B002 (5)	2
34	P (5)	B002 (5)	2
35	A (7)	B002 (5)	3

patient. These 225 patients were in 17 clusters; the largest cluster contained 142 patients, whose DNA fingerprints displayed a common characteristic single band. Overall, RFLP by IS6110 identified 271 distinct patterns among 378 isolates.

Combined RFLP analysis with IS6110 and the DR probes identified 236 patients with distinct patterns, and 142 patients (38%) carried a strain shared with at least one other patient. These 142 patients were in one of the 35 clusters (Table 1 and Fig. 1). The number of patients in each cluster ranged from 2

to 25. Excluding 78 patients who had a single copy of IS6110, there were 34 patients (9%) who had a clustered isolate by both probes.

Patients who had been identified by house-to-house survey were more likely to be part of a cluster than those who had been identified when they sought care at a health facility (49% versus 35%; $P = 0.06$). Patients who had a relapse were insignificantly more likely to be part of a cluster compared with new patients (10 of 18, 56%, versus 114 of 302, 38%; $P = 0.2$). Eight of the 10 patients with relapse who were in a cluster had isolates with a single band. Patients who had been hospitalized in a sanatorium were somewhat more likely to be in a cluster than those who had not been (47% versus 34%; $P = 0.06$). MDR, alcoholism, family size, and close contact with a pulmonary TB patient were not significant risk factors for clustering. Limiting the analysis to patients detected by house-to-house survey ($n = 66$) showed that patients in a cluster were significantly older than patients who were not (55 versus 43 years; $P = 0.01$).

One hundred thirty-one of the 142 patients who were in one or more clusters could be contacted for epidemiologic investigation. Of these 131 patients, 21 (16%) reported acquaintance with one or more patients in the same cluster. Eleven of the 35 clusters comprised four or more patients, and these were analyzed in detail (Table 2). In clusters 2 and 7, clustering was significantly associated with case detection by a house-to-house community survey. Being hospitalized in a sanatorium was a significant risk factor for clustering among patients in cluster 10.

This investigation reports several characteristics of the molecular epidemiology of TB, by using IS6110 and DR probes, which differ from previously reported findings in other settings. Forty-one percent of *M. tuberculosis* isolates harbored a single IS6110 copy. Such a high proportion of single-copy isolates has not been reported elsewhere except in south India (9, 24; I. Radhakrishnan, Y. K. Manju, R. Ajaykumar, and S. Mundayoor, Letter, J. Clin. Microbiol. **39**:1683, 2001). Other than studies in South East Asian countries, which have reported single-copy strain prevalences of up to 20% (13, 22, 34), most studies have reported less than 5% single-copy strains (5, 16, 19, 20). Not all IS6110 single-copy isolates are genetically identical or clonal in origin (18). In this study, a third of the IS6110 single-copy clusters were differentiated by the DR probe.

The proportion of clustering in this study ranged from 9 to 38%, depending on whether single-copy strains were excluded

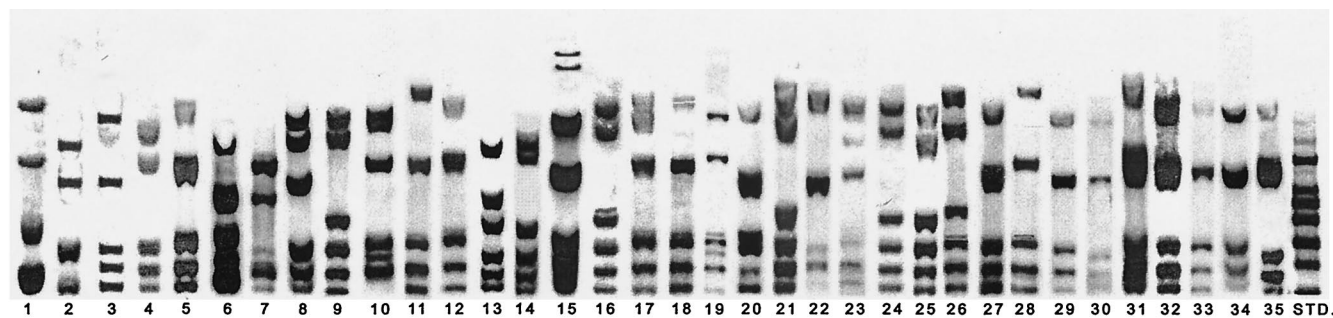


FIG. 1. DR RFLP patterns associated with a clustered *M. tuberculosis* isolate in Tiruvallur District, South India, July 1999 to June 2000. STD, standard.

TABLE 2. Characteristics of patients with unique RFLP patterns and patients in 11 clusters of ≥ 4 patients in Tiruvallur District, South India, July 1999 to June 2000^a

Cluster	No. of patients	No. of bands		Median age (yr)	% of patients							
		IS6110	DR		Epidemiologically linked	Age >45 yr	Men	Alcoholic	Employed	MDR	Diagnosed by survey	Hospitalized in a sanatorium
Distinct RFLP	236			42								
1	25	1	5	45	27	48	84	33	77	4	14	16
2	6	1	5	49	0	56	76	55	64	4	20	25
3	7	1	5	45	2	67	100	40	100	17	50*	0
4	10	1	5	49	22	57	86	60	100	0	14	29
5	5	1	6	35	0	70	80	33	67	14	10	40
6	4	1	6	31	0	20	80	0	80	0	40	20
7	4	1	6	31	0	100	100	25	100	25	0	0
8	6	1	6	43	0	67	83	33	83	0	67†	0
9	8	1	7	50	0	75	75	17	83	0	38	13
10	5	1	7	40	0	40	60	0	75	0	20	20
11	7	1	7	50	0	71	86	29	43	0	14	57*
11	4	4	5	44	0	50	100	33	100	0	0	0

^a *P* values are noted for comparison with characteristics of patients with unique RFLP patterns: *, *P* \leq 0.05; †, *P* < 0.01.

or included in the analysis, respectively. The proportion of recently transmitted cases is reported to vary widely from 10% to 58% (3, 4, 7, 8, 10, 15, 17, 20, 25, 26) in developed countries and from 20% to 36% in developing countries (12, 16, 32). In addition to the relatively low clustering, the transmission pattern may also differ in our setting. In investigations conducted in industrialized countries, an index case was often identified and transmission links could be established among most of the patients in clusters (1, 25). Despite the intensive field investigations in this study, definite epidemiological links were identified in less than a fifth of the patients who shared similar strains. Even in the largest cluster, comprising 25 patients, links could be established only between pairs. These findings suggest that several people exposed to a small risk may account for more cases than a few people exposed to a high risk (23). It is likely that the patients in this study were infected at crowded places, such as bus stops, railway stations, cinemas, and outpatient health facilities or during religious festivals.

Another important observation of this study is that clustering was higher among older patients. Most other studies have reported recent transmission in the younger age groups (14, 26, 32). It is commonly believed that transmission due to endogenous reactivation increases with age. In this study setting, there are at least two likely explanations for the higher clustering among older patients. First, there are certain strains endemic to the region and therefore available for transmission over a considerable period of time among residents living in neighboring communities for several years. Ninety percent of the patients in this study lived in the same place for more than 5 years. Second, the source of infection could be a reactivated strain from an older patient, which could subsequently have been recently transmitted to individuals belonging to one age group, because of peer interaction and crowded living conditions. The same reasons could also account for the higher percentage of clustering attributable to the door-to-door survey carried out with 17% of the population in this study.

A higher proportion of clustering was observed among patients who had a relapse. Furthermore, a majority of the patients who had a relapse and were in a cluster harbored an IS6110 single-copy strain. Nearly half the patients who gave a

history of hospitalization in a TB sanatorium in the past 5 years were in a cluster. While epidemiologic links among these patients could not be identified, nosocomial transmission may be a likely cause, as has been reported elsewhere (14).

Previous studies have reported higher clustering among drug-resistant patients than drug-sensitive ones (1, 6, 14, 26, 27). We found no association between drug resistance and clustering. However, the rate of MDR TB was higher among patients in whom RFLP patterns were not available; there may thus be a bias causing the possible underestimation of the proportion of MDR TB patients who were in clusters.

In conclusion, the findings of this study suggest that the majority of the TB cases in south India occur predominantly due to reactivation. With the large-scale introduction of directly observed treatment—short course (DOTS) in 1998, the transmission of infection is likely to be reduced due to the higher cure and lower relapse rates. However, the reactivation of latent infection will continue to perpetuate new cases for years to come despite effective DOTS implementation. Therefore, DOTS implementation in India must be supplemented by intensive and early case detection of TB, infection control measures to prevent TB transmission at outpatient and inpatient health facilities, and community education. Long-term strategies such as developing prophylactic tools capable of preventing infection and progression to disease should also be undertaken. Similar studies are required to characterize the molecular epidemiology of *M. tuberculosis* in other regions of India and to learn whether the high percentage of IS6110 single-copy strains identified in this study is unique to south India. Further research is needed to understand the role of virulence in reactivation of *M. tuberculosis* strains.

This report was funded in part by a grant from the United States Agency of International Development provided through the World Health Organization.

We thank Nirupa Charles and Jagarajamma for patient interviews. We acknowledge the support from the Department of Bacteriology and Ramanujam of the Department of Immunology for processing the samples; P. G. Gopi and R. Subramani and their staff for data management; the field staff for their support; and T. Santha Devi, C. Kollappan, and Sadacharam for coordination of field activities. We

acknowledge editorial assistance from Jaya Shreedhar and Bandana Malhotra and secretarial assistance rendered by R. Senthil Nathan and V. Shanthi.

REFERENCES

- Alland, D., G. E. Kalkut, A. R. Moss, R. A. McAdam, J. A. Hahn, W. Bosworth, E. Drucker, and B. R. Bloom. 1994. Transmission of tuberculosis in New York City. *N. Engl. J. Med.* **24**:1710–1716.
- Allen, B. W., and F. J. Baker. 1968. *Mycobacteria*: isolation, identification and sensitivity testing. Laboratory aids series. Butterworths, London, United Kingdom.
- Barnes, P. F., H. El-Hajj, S. Preston-Martin, M. D. Cave, B. E. Jones, M. Otaya, J. Pogoda, and K. D. Eisenach. 1996. Transmission of tuberculosis among the urban homeless. *JAMA* **275**:305–307.
- Barnes, P. F., Z. Yang, S. Preston-Martin, J. M. Pogoda, B. E. Jones, M. Otaya, K. D. Eisenach, L. Knowles, S. Harvey, and M. D. Cave. 1997. Patterns of tuberculosis transmission in Central Los Angeles. *JAMA* **278**:1159–1163.
- Bauer, J., Z. Yang, S. Poulsen, and Å. B. Andersen. 1998. Results from 5 years of nationwide DNA fingerprinting of *Mycobacterium tuberculosis* complex isolates in a country with a low incidence of *M. tuberculosis* infection. *J. Clin. Microbiol.* **36**:305–308.
- Beck-Sague, C., S. W. Dooley, M. D. Hutton, J. Otten, A. Breeden, J. T. Crawford, A. E. Pitchenik, C. Woodley, G. Cauthen, and W. R. Jarvis. 1992. Hospital outbreak of multi drug resistant *Mycobacterium tuberculosis* infection: factors of transmission to staff and HIV-infected patients. *JAMA* **268**:1280–1286.
- Braden, C. R., G. L. Templeton, M. D. Cave, S. Valway, I. M. Onorato, K. G. Castro, D. Moers, Z. Yang, W. W. Stead, and J. H. Bates. 1997. Interpretation of restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates from a state with a large rural population. *J. Infect. Dis.* **175**:1446–1452.
- Burman, W. J., R. R. Reves, A. P. Hawkes, C. A. Rietmeijer, Z. H. Yang, H. El-Hajj, J. H. Bates, and M. D. Cave. 1997. DNA fingerprinting with two probes decreases clustering of *Mycobacterium tuberculosis*. *Am. J. Respir. Crit. Care Med.* **155**:1140–1146.
- Das, S., C. N. Paramasivan, D. B. Lowrie, R. Prabhakar, and P. R. Narayanan. 1995. IS6110 RFLP typing of clinical isolates of *M. tuberculosis* from patients with pulmonary tuberculosis in Madras, South India. *Tuber. Lung Dis.* **76**:550–554.
- de Lowrdes Garcia-Garcia, M., A. Palacios-Martinez, A. Ponce-de-Leon, M. E. Jimenez-Corona, A. Jimenez-Corona, S. Balandrano-Campos, H. Olivera-Diaz, J. L. Valdespino-Gomez, and P. M. Small. 2000. The role of core groups in transmitting *Mycobacterium tuberculosis* in a high prevalence community in Southern Mexico. *Int. J. Tuberc. Lung Dis.* **4**:12–17.
- Dye, C., S. Scheele, P. Dolphin, V. Pathania, and M. C. Raviglione. 1999. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence and mortality by country. W. H. O. Global Surveillance and Monitoring Project. *JAMA* **282**:677–686.
- Faussett, P. G., P. Sonnenberg, S. C. Shearer, M. C. Bruce, C. Mee, L. Morris, and J. Murray. 2000. Tuberculosis control and molecular epidemiology in a South African gold mining community. *Lancet* **356**:1066–1071.
- Fomukong, N. G., T. H. Tang, S. Al-Maamary, W. A. Ibrahim, S. Ramayah, M. Yates, F. Zainuddin, and J. W. Dale. 1994. Insertion typing of *Mycobacterium tuberculosis*: characterization of a widespread subtype with a single copy of IS6110. *Tuber. Lung Dis.* **75**:435–440.
- Frieden, T. R., P. I. Fujiwara, R. M. Washko, and M. A. Hamburg. 1995. Tuberculosis in New York City—turning the tide. *N. Engl. J. Med.* **333**:229–233.
- Genewein, A., A. Telenti, C. Bernasconi, C. Mordasini, S. Weiss, A. M. Maurer, H. Rieder, K. Schopfer, and T. Bodmer. 1993. Molecular approach to identifying route of transmission of tuberculosis in the community. *Lancet* **342**:841–844.
- Haas, W. H., G. Engelmann, B. Amthor, S. Shyamba, F. Mugala, M. Felten, M. Rabbow, M. Leichsenring, O. J. Oosthuizen, and H. J. Bremer. 1999. Transmission dynamics of tuberculosis in a high-incidence country: prospective analysis by PCR DNA fingerprinting. *J. Clin. Microbiol.* **37**:3975–3979.
- Hermans, P. W. M., D. van Soolingen, E. M. Bik, P. E. W. de Haas, J. W. Dale, and J. D. A. van Embden. 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.
- Hermans, P. W. M., D. van Soolingen, J. W. Dale, A. R. J. Schuitema, R. A. McAdam, D. Catty, and J. D. A. van Embden. 1990. Insertion element IS986 from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. *J. Clin. Microbiol.* **28**:2051–2058.
- Hermans, P. W. M., F. Messadj, I. I. Guebrexabher, D. van Soolingen, P. E. W. de Haas, H. Heersam, H. de Neeling, A. Ayoub, F. Portaels, D. Frommel, M. Zribi, and J. D. A. van Embden. 1995. Analysis of the population structure of *Mycobacterium tuberculosis* in Ethiopia, Tunisia and the Netherlands: usefulness of DNA typing for global tuberculosis epidemiology. *J. Infect. Dis.* **171**:1504–1513.
- Horwitz, O., P. Q. Edwards, and A. M. Lowell. 1973. National Tuberculosis Control Program in Denmark and United States. *Health Serv. Rep.* **88**:493–498.
- Khatri, G. R., and T. R. Frieden. 2000. The status and prospectus of tuberculosis control in India. *Int. J. Tuberc. Lung Dis.* **4**:193–200.
- Palittapongarnpim, P., P. Luangsook, S. Tansuphaswadikul, C. Chuchottaworn, R. Prachaktam, and B. Sathapatayavongs. 1997. Restriction fragment length polymorphism study of *Mycobacterium tuberculosis* in Thailand using IS6110 as probe. *Int. J. Tuberc. Lung Dis.* **1**:370–376.
- Rose, G. 1992. The strategy of preventive medicine. Oxford University Press, Oxford, United Kingdom.
- Sahadevan, R., S. Narayanan, C. N. Paramasivan, R. Prabhakar, and P. R. Narayanan. 1995. Restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, India, by use of direct-repeat probe. *J. Clin. Microbiol.* **33**:3037–3039.
- 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.
- Snider, D. E., Jr., G. D. Kelly, G. M. Cauthen, N. J. Thompson, and J. O. Kilburn. 1985. Infection and disease among contacts of tuberculosis cases with drug-resistant and drug-susceptible bacilli. *Am. Rev. Respir. Dis.* **132**:125–132.
- Torrea, G., C. Offredo, M. Simonet, B. Gicquel, P. Berche, and C. Pierre-Audigier. 1996. Evaluation of tuberculosis transmission in a community by 1 year of systematic typing of *Mycobacterium tuberculosis* clinical isolates. *J. Clin. Microbiol.* **34**:1043–1049.
- Tuberculosis Research Centre. 1996. Fifteen year follow up of trial of BCG vaccines in South India for tuberculosis prevention. *Ind. J. Med. Res.* **110**:56–69.
- Tuberculosis Research Centre. 2001. Trends in the prevalence and incidence of tuberculosis in South India. *Int. J. Tuberc. Lung Dis.* **5**:142–157.
- van Embden, J. D. A., M. D. Cave, J. T. Crawford, J. W. Dale, K. D. Eisenach, B. Gicquel, P. Hermans, C. Martin, R. McAdam, T. M. Shinnick, and P. M. Small. 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J. Clin. Microbiol.* **31**:406–409.
- van Soolingen, D., M. W. Borgdorff, P. E. W. De Haas, M. M. G. G. Sebek, J. Veen, M. Dessens, K. Kremer, and J. D. A. van Embden. 1999. Molecular epidemiology of tuberculosis in The Netherlands: a nationwide study from 1993 through 1997. *J. Infect. Dis.* **180**:726–736.
- Warren, R., J. Hauman, N. Beyers, M. Richardson, H. S. Schaaf, P. Donald, and P. van Helden. 1996. Unexpectedly high strain diversity of *Mycobacterium tuberculosis* in a high-incidence community. *S. Afr. Med. J.* **86**:45–49.
- World Health Organization. 2000. Joint review of tuberculosis in India. (W. H. O./SEA/TB/224) World Health Organization, Regional Office for South-East Asia, New Delhi, India.
- Yuen, L. K. W., B. C. Ross, K. M. Jackson, and B. Dwyer. 1993. Characterization of *Mycobacterium tuberculosis* strains from Vietnamese patients by Southern blot hybridization. *J. Clin. Microbiol.* **31**:1615–1618.