Molecular epidemiology of tuberculosis: achievements and challenges to current knowledge

Megan Murray[1] & Edward Nardell[2]

Abstract Over the past 10 years, molecular methods have become available with which to strain-type *Mycobacterium tuberculosis*. They have allowed researchers to study certain important but previously unresolved issues in the epidemiology of tuberculosis (TB). For example, some unsuspected microepidemics have been revealed and it has been shown that the relative contribution of recently acquired disease to the TB burden in many settings is far greater than had been thought. These findings have led to the strengthening of TB control. Other research has demonstrated the existence and described the frequency of exogenous reinfection in areas of high incidence. Much recent work has focused on the phenotypic variation among strains and has evaluated the relative transmissibility, virulence, and immunogenicity of different lineages of the organism. We summarize the recent achievements in TB epidemiology associated with the introduction of DNA fingerprinting techniques, and consider the implications of this technology for the design and analysis of epidemiological studies.

Keywords Mycobacterium tuberculosis/genetics/transmission; Tuberculosis/epidemiology; Epidemiology, Molecular; DNA fingerprinting/utilization; Tuberculosis, Multidrug-resistant/transmission; Risk factors; Cost of illness; Review literature (*source: MeSH, NLM*).

Mots clés Mycobacterium tuberculosis/génétique/transmission; Tuberculose/épidémiologie; Epidémiologie moléculaire; Empreinte génétique/utilisation; Tuberculose résistante à la polychimiothérapie/transmission; Facteur risque; Coût maladie; Revue de la littérature (*source: MeSH, INSERM*).

Palabras clave Mycobacterium tuberculosis/genética/transmisión; Tuberculosis/epidemiología; Epidemiología molecular; Dermatoglifía del ADN/utilización; Tuberculosis resistente a multidrogas/transmisión; Factores de riesgo; Costo de la enfermedad; Literatura de revisión (*fuente: DeCS, BIREME*).

Voir page 480 le résumé en français. En la página 480 figura un resumen en español.

Introduction

Styblo defined tuberculosis (TB) epidemiology as "the study of the interactions between the tubercle bacillus and man in his environment (in a community)" (*1*), and remarked that it was particularly important to study them under natural conditions without any interference in the form of direct or indirect control measures. The data that Styblo relied on to assess the burden of clinical TB included case notifications based on the examination of sputum by microscopy, bacteriological cultures, and chest radiographs. Rates of TB infection were estimated from surveys involving serial tuberculin skin tests. These tools allowed him to describe the downward trend in the incidence of TB in Europe during the 20th century, to measure the mortality associated with untreated disease, to estimate infectiousness and to measure the contribution of exogenous reinfection to TB morbidity.

Notwithstanding the work of Styblo, Gryzbowski, Comstock, Stead, and others, many important questions remain unresolved, largely because the natural history of the disease makes it so difficult to study. The armamentarium has often been inadequate for studying patterns of occurrence of tuberculosis, especially in those areas of the developing world where its toll is highest. Surveys based on the tuberculin skin test are often difficult to interpret because of cross-reactivity with BCG vaccine and environmental mycobacteria. Case notification data continue to underestimate the disease burden in areas where the prevalence of TB is high but resources for diagnosis and record-keeping are limited. Until relatively recently it has not been possible to trace pathways of TB transmission within populations.

Unresolved issues in TB epidemiology

By the early 1990s, when molecular fingerprinting first appeared, many questions in TB epidemiology remained unresolved. They included the relative contributions of reactivation and primary disease in areas of high and low prevalence, risk factors for recent infection and/or primary disease, and the occurrence and frequency of exogenous reinfection. More recent issues include the impact of human immunodeficiency virus (HIV) coinfection on transmission, the infectiousness of smear-negative tuberculosis, the relative transmissibility of different strains, and other phenotypic differences among strains of *M. tuberculosis*.

Efforts to strain-type *M. tuberculosis* failed until the 1990s, when polymorphic sites were identified in repetitive sequences in the genome (*2*, *3*). The most widely used marker is the transposable element IS*6110* (*4*), which varies in both copy number and location in the genome. This marker is generally

[1] Assistant Professor, Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA.

[2] Associate Professor, Department of Medicine, Harvard Medical School, and Chief, Pulmonary Medicine, The Cambridge Hospital, 1493 Cambridge Street, Cambridge, MA 02139, USA. Correspondence should be addressed to this author (email: enardell@pih.org).

considered to be variable enough to distinguish between unrelated strains but stable enough to remain consistent in related strains. Limitations of molecular fingerprinting include its inability to distinguish between unrelated isolates with a low copy numbers.

Soon after their development the fingerprinting techniques were used to document the transmission of *M. tuberculosis* between contacts. Daley et al. described 12 cases of TB that occurred in a housing facility in San Francisco, USA, for HIV-infected people (5). All 12 *M. tuberculosis* isolates shared a single IS*6110* fingerprint, confirming the authors' expectation that the cases were attributable to the recent transmission of the pathogen in the institutional setting. While this study suggested that TB could both spread rapidly and progress rapidly to active disease in people infected with HIV, many subsequent molecular studies documented the transmission and progression of TB among immunocompetent people as well (6, 7). Shortly after the San Francisco outbreak, Godfrey-Faussett et al. reported the clustering of two *M. tuberculosis* isolates taken from neighbours with active tuberculosis (8). This was a further demonstration of the concordance between molecular data and the results of conventional contact studies.

Assessing the burden of recent transmission

Subsequently, molecular fingerprinting has been used in many studies to confirm suspected outbreaks and to discern previously unsuspected transmission. Thus Genewein et al. fingerprinted all *M. tuberculosis* isolates notified in Berne, Switzerland, during one year and identified a cluster of 22 cases that belonged to a defined but complex social network (9). This cluster could not have been easily identified through the tracing of contacts in the standard way.

These studies made it clear that contact-tracing often failed to identify transmission networks. They also showed that many cases of tuberculosis, previously classified as reactivation disease, shared a DNA fingerprint with other contemporaneous cases. Most researchers interpret clusters as being epidemiologically linked chains of recently transmitted disease, and unique isolates as being cases of reactivation disease, resulting from remote TB infection. The finding that a large proportion of cases were in clusters challenged the conventional wisdom that the vast majority of TB cases in low-incidence countries were attributable to the reactivation of *M. tuberculosis* infection acquired in the remote past. Researchers in New York City (10) and San Francisco (11) in the USA investigated this matter by systematically enrolling consecutive cases in a population-based approach. Using molecular methods to identify clusters, they independently estimated that 40% of incident TB cases fell into clusters and were thus classified as recently acquired disease. Since these proportions of clustered cases were much higher than expected, the findings demonstrated the problem of TB transmission in urban centres in the USA and helped to invigorate previously neglected efforts in the field of TB control. Studies in Europe found that the proportion of clustered cases ranged from 16% to 46%, suggesting that transmission of *M. tuberculosis* could be an important factor even in areas of very low incidence (12, 13).

Whereas large population-based studies in low-incidence countries have estimated similar proportions of clustered and unique cases, the data from areas of higher incidence have been less consistent. The proportion of clustered cases reported from Africa (14, 15) has ranged from approximately 30% to 70%, while in China and Viet Nam (16, 17) a single genotype defined by spoligotyping, the Beijing strain, accounted for 50–80% of the cases sampled. Smaller studies in other high-incidence areas (18, 19) have suggested a surprising lack of recent transmission, with only about 20% of cases belonging to clusters. These findings have raised questions about the correct interpretation of cluster studies and have sparked a growing interest in methodological research on inference in molecular epidemiology.

Several investigators have approached this problem by simulating the process of sampling isolates from a hypothetical distribution of clusters of *M. tuberculosis* isolates and then measuring the impact of this sampling on estimates of the proportion of clustered cases. Glynn et al. used computer simulations to show that recent transmission was increasingly underestimated as the fraction of all cases included in the sample was decreased (20). Murray described an analytical method for estimating the bias incurred by sampling when the underlying cluster distribution was known (21). Although neither of these methods allows estimates of recent transmission to be adjusted without knowledge of the sampling fraction and the distribution of cluster sizes before sampling, they suggest that small studies in areas of high incidence severely underestimate recent transmission.

Another strategy for assessing these molecular methods has involved developing epidemic models as tools for epidemiological inference. Vynnycky et al. (22) modelled the transmission of *M. tuberculosis* strains in the Netherlands over the past century, allowing DNA fingerprint patterns to change over time because of random mutation. By using age-specific rates of primary and reactivation disease, they showed that the clustering of cases on the basis of identical fingerprints underestimated recent transmission among younger cases but tended to overestimate it in the elderly. Murray used a stochastic model of TB transmission to identify social and demographic determinants of cluster distribution and to observe the effect of transmission dynamics on the empirical data obtained from studies in molecular epidemiology (23). It emerged that in this study, the proportion of clustered cases of TB varied with a range of host and population characteristics and that it might not be a direct reflection of the incidence of recent transmission in a specific setting. These findings suggest that studies in the area of molecular epidemiology which attempt to estimate the burden of recently transmitted TB should be interpreted with great caution.

Risk factors for recent transmission

As well as enabling researchers to classify cases as clustered or unique, molecular techniques have made it possible to identify risk factors for clustering and, by extension, for the recent transmission and rapid progression of clinical tuberculosis. In their analysis of data relating to New York city, USA, Alland et al. (10) compared the prevalence of sociodemographic and clinical risk factors in clustered and unique cases: young age, birth in the USA, Hispanic ethnicity, and HIV infection were identified as risk factors for recently transmitted disease. Studies in different geographical settings have confirmed some of these risk factors but have also produced discrepant results: HIV was not a risk factor for clustering among South African gold miners (24), nor did it predict clustering in hospitalized

patients in Rio de Janeiro, Brazil (*25*). Similarly, factors such as homelessness, alcoholism, and intravenous drug use have been identified as risk factors for recent transmission in some areas but are not associated with this outcome in others (*26, 27*).

Since social and demographic risk factors identify people at risk for being exposed to tuberculosis it is hardly surprising that they vary in different communities with differing disease dynamics. Nonetheless, some of these differences may have resulted from the wide variation in the fraction of cases sampled. Just as sampling can bias estimates of the proportions of clustered cases in communities by underestimation, in the same way it can bias estimates of the effects of risk factors for clustering. In a reassessment of the impact of HIV on clustering in New York city, USA, Murray & Alland (*28*) adjusted for potential sampling bias and revised their estimate of the odds ratio from 2.7 to 23.6. Since sampling strategies vary widely between studies, it follows that odds ratios for risk factors for clustering also vary and are more or less accurate, depending on the total number of cases sampled.

Exogenous reinfection

The relative frequency of exogenous reinfection has been a topic of speculation over much of the past century. Styblo (*1*) and others have argued that exogenous reinfection makes an important contribution to the burden of active TB in many communities where the disease is highly prevalent. Styblo cited the experience of Greenland in the 1950s, noting that a vigorous intervention programme led to a decline in clinical TB, not only among the young people in whom a first infection was averted but also among the elderly who had almost certainly been previously infected. Had most disease in the elderly been due to the reactivation of remote tuberculosis, he argued, the incidence in that population would not have been affected by a reduction in transmission. Many years later, Nardell et al. provided further circumstantial evidence for exogenous reinfection (*29*). In a TB outbreak in a shelter for the homeless in Boston, USA, for 7 of 25 cases, linked both by an identical drug-resistance pattern and phage type, there was documentary evidence of previous infection or disease.

The development of molecular fingerprinting has silenced the debate about the existence of exogenous reinfection. Several groups have shown that reinfection with a second distinct strain can occur in both immunocompromised and immunocompetent individuals (*30, 31*). More recently, researchers have used molecular epidemiological techniques to quantify the frequency with which exogenous reinfection occurs and to identify the context in which it may contribute substantially to the dynamics of TB. For example, van Rie et al. enrolled consecutive TB cases over a six-year period from a South African community in which there was a high incidence of the disease (*32*). Of the 698 cases identified, 16 had recurrent disease after completing curative therapy and 12 of the latter were infected with a strain that was different from that isolated during their first episode of disease. Similar studies conducted in areas of lower prevalence showed that in low-risk areas, reinfection accounted for fewer of the second episodes of disease (*33, 34*).

While these studies suggest that exogenous reinfection occurs more commonly than previously believed, they do not indicate how many cases are attributable to recently acquired infection in people previously infected with TB. Most cases of exogenous reinfection are expected to occur in people who have been infected in the past but in whom clinical TB did not result from the initial infection. *M. tuberculosis* isolates are not available from two distinct episodes of disease in these people, who therefore cannot be counted by the simple typing of incident cases. One way of assessing the burden of reinfection would be to estimate the frequency of clustered cases in a cohort of people found to be infected when previous tuberculin skin testing was carried out. However, this study design would require long-term follow-up of an infected cohort, and this could represent a considerable expenditure of time and resources.

Phenotypic differences between strains

The advent of molecular typing has also allowed researchers to describe strain-specific variation in clinical phenotypes of *M. tuberculosis* such as virulence, growth characteristics, immunogenicity, and transmissibility. Although phenotypic differences among clinical isolates have long been recognized, it was not previously possible to determine whether they were stably associated with specific lineages of *M. tuberculosis* circulating in the population. Several of the strains identified in outbreaks have been associated with large clusters that are widely dispersed both geographically and temporally, raising the possibility that they are either more transmissible or more likely to cause disease once transmitted than are other strains. One such lineage is W-related strains, of which over 500 cases have been reported in New York, USA, since 1991 (*35*). DNA fingerprints from this strain closely resemble those of a large family of related lineages representing a significant proportion of the isolates that have been typed throughout Asia, Eastern Europe, and the Russian Federation. The identification of this highly successful strain has led to laboratory-based efforts to identify bacterial factors that may distinguish the lineage from other, less widely disseminated clones. Zhang et al., for example, found that W-related strains had an enhanced capacity to replicate in human macrophages and have suggested that this function might be associated with the organism's success (*36*).

Despite these findings, recent attempts to fully characterize specific clinical strains suggest that the assessment of "notorious" *M. tuberculosis* isolates can be problematic. Host factors clearly affect the behaviour of strains and can be difficult to disentangle from bacterial traits. Laboratory studies measuring the growth rates of such strains have to compare them with those of uniform reference strains that have been shown to vary substantially in different environments. Finally, it can be challenging to use epidemiological patterns in order to differentiate microbial virulence from other bacterial factors such as immunogenicity and/or transmissibility. These problems have been exemplified by strain CDC1551, an *M. tuberculosis* isolate that was identified as being responsible for a microepidemic in a rural area of the USA. In this largely susceptible community, 80% of contacts of the index case gave positive results to the tuberculin skin test, leading investigators to propose that the organism was more virulent than other strains (*37*). This hypothesis was supported by the early observation that CDC1551 grew to very high levels of bacilli in the lungs of infected animal models. Subsequent studies, however, found that other reference laboratory strains could also be induced to attain these levels and, indeed, that CDC1551 grew more slowly than these strains later in the

course of an infection (*38*). Nonetheless, in other experiments it was found that CDC1551 differed from standard laboratory strains in inducing a more vigorous cytokine-mediated immune response in human monocytes (*39*) and producing smaller tubercles in the lungs of infected rabbits (*40*). These results raise the possibility that the high conversion rate among contacts might reflect the heightened immunogenicity of this strain, which would increase the sensitivity of the tuberculin skin test. This experience has made clear the importance of distinguishing between virulence, transmissibility, and immunogenicity in future studies of strain-specific phenotypes.

Relative transmissibility of drug-sensitive and drug-resistant *M. tuberculosis*

Sepkowitz summarized the literature predating the emergence of molecular epidemiology on the risk of infection in households and among other contacts of TB cases (*41*). Although the studies identified a number of host factors associated with an increased risk of TB transmission, few of them considered whether there were variations in infectiousness between specific strains. Animal studies had suggested that strains characterized by isoniazid resistance mutations grew less vigorously in guinea-pigs than drug-sensitive strains (*42*). Subsequent work, however, showed that resistance to isoniazid was encoded by a number of different point mutations, raising the possibility that the behaviour of different drug-resistant strains might be quite heterogeneous (*43*).

Epidemiological studies designed to address this question have compared the number of tuberculin skin test positives and/or cases of clinical TB in household contacts exposed to drug-sensitive and drug-resistant source cases. Using this method in 1985, Snider et al. (*44*) found no difference in infectiousness between the two sources. A similar study by Teixeira et al. (*45*) reported that the prevalence of TB infection and progression to active disease were comparable in these groups.

Other molecular studies have taken a different approach to estimating the relative transmissibility of strains, comparing the sizes of clusters of drug-sensitive and drug-resistant isolates. One study conducted in the Netherlands (*46*) showed that isoniazid resistance was negatively associated with clustering while a second (*47*) found that being infected with multidrug-resistant TB was associated with a decreased likelihood of being in a cluster. Conflicting results were reported by Alland et al., who found that drug resistance was a strong predictor of clustering among TB cases in New York city, USA, some of which were attributable to the family of W strains mentioned above (*10*). Many studies have documented the widespread transmission of drug-resistant organisms and noted that these disseminated drug-resistant strains belong disproportionately to the family of W strains (*48–50*). These results may again raise the question of whether molecular cluster studies can be used to infer the risk of infection or disease, in this case among those exposed to drug-sensitive or drug-resistant strains. There are many possible reasons why clusters of TB cases involving resistance to one or more drugs may be smaller than clusters of drug-sensitive cases. People with drug-resistant TB may have poorer access to health care and be less likely to be sampled than people with sensitive strains of the organism. They may also have fewer susceptible contacts than people with drug-sensitive strains, either because they have fewer social interactions in general or because the people with whom they make contact are more likely to have been infected with TB in the past. Finally, since specific drug-resistant and drug-sensitive strains probably differ at a variety of other genetic loci, observed differences in transmissibility and virulence may be related to strain differences that are independent of drug-resistance phenotypes. These potentially confounding factors make such studies difficult to interpret and confirm the need for careful thought and innovative approaches to the design of epidemiological studies that use molecular methods. ■

Résumé

Epidémiologie moléculaire de la tuberculose : acquisitions récentes

Ces dix dernières années, des méthodes moléculaires permettant le typage des souches de *Mycobacterium tuberculosis* sont apparues. Elles ont permis aux chercheurs d'étudier certains aspects importants mais non encore élucidés de l'épidémiologie de la tuberculose. Par exemple, certaines micro-épidémies passées inaperçues ont été mises en évidence et il a été démontré que la contribution relative des cas récents à la charge de la tuberculose dans nombre de contextes était largement supérieure à ce que l'on pensait. Ces résultats ont conduit à renforcer la lutte antituberculeuse. D'autres travaux ont démontré l'existence de réinfections exogènes dans les zones de forte incidence et en ont décrit la fréquence. Récemment, de nombreux travaux ont porté sur la variation phénotypique entre souches et ont évalué la transmissibilité, la virulence et l'immunogénicité relatives de différentes lignées de bacilles tuberculeux. Le présent article résume les acquisitions récentes dans le domaine de l'épidémiologie de la tuberculose grâce aux techniques de typage moléculaire et examine les répercussions de ces techniques sur la conception et l'analyse des études épidémiologiques.

Resumen

Epidemiología molecular de la tuberculosis: logros y retos para nuestros actuales conocimientos

A lo largo de los últimos 10 años se ha empezado a disponer de métodos moleculares para tipificar las cepas de *Mycobacterium tuberculosis.* Dichos métodos han permitido a los investigadores estudiar algunos temas importantes que no se habían resuelto en materia de epidemiología de la tuberculosis. Así, por ejemplo, se han detectado algunas microepidemias que nadie sospechaba, y se

ha demostrado que la contribución relativa de la enfermedad recién adquirida a la carga de tuberculosis en muchos entornos es mucho mayor de lo que se creía. Estos resultados han permitido reforzar la lucha antituberculosa. Otras investigaciones han revelado la existencia, y descrito su frecuencia, de casos de reinfección exógena en zonas de alta incidencia. Muchos trabajos recientes, centrándose en las diferencias fenotípicas entre cepas, han evaluado la transmisibilidad, virulencia e inmunogenicidad relativas de las diferentes cepas del microorganismo. Resumimos aquí los últimos logros conseguidos gracias a las técnicas de determinación de las huellas de ADN en el campo de la epidemiología de la tuberculosis, y analizamos las repercusiones de esta tecnología para el diseño y análisis de los estudios epidemiológicos.

References

1. Styblo K. *Epidemiology of tuberculosis*. Jena: Gustav Fischer Verlag; 1984.
2. van Soolingen D, Hermans PW, de Haas PE, Soll DR, van Embden JD. 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. *Journal of Clinical Microbiology* 1991;29:2578-86.
3. Cave MD, Eisenach KD, McDermott PF, Bates JH, Crawford JT. IS*6110*: conservation of sequence in the *Mycobacterium tuberculosis* complex and its utilization in DNA fingerprinting. *Molecular and Cellular Probes* 1991;5:73-80.
4. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *Journal of Clinical Microbiology* 1993;31:406-9.
5. Daley CL, Small PM, Schecter GF, Schoolnik GK, McAdam RA, Jacobs WR. An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *New England Journal of Medicine* 1992;326:231-5.
6. Godfrey-Faussett P, Mortimer PR, Jenkins PA, Stoker NG. Outbreak of tuberculosis in a church. *Chest* 1995;107:447-52.
7. Kline SE, Hedemark LL, Davies SF. Outbreak of tuberculosis among regular patrons of a neighborhood bar. *New England Journal of Medicine* 1995;333:222-7.
8. Godfrey-Faussett P, Mortimer PR, Jenkins PA, Stoker NG. Evidence of transmission of tuberculosis by DNA fingerprinting. *BMJ* 1992;305:221-3.
9. Genewein A, Telenti A, Bernasconi C, Mordasini C, Weiss S, Maurer AM, et al. Molecular approach to identifying route of transmission of tuberculosis in the community. *Lancet* 1993;342:841-4.
10. Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, et al. Transmission of tuberculosis in New York City. An analysis by DNA fingerprinting and conventional epidemiologic methods. *New England Journal of Medicine* 1994;330:1710-6.
11. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, et al. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods in the 1990s. *New England Journal of Medicine* 1994;330:1703-9.
12. Borgdorff MW, Nagelkerke N, van Soolingen D, de Haas PE, Veen J, van Embden JD, et al. Analysis of tuberculosis transmission between nationalities in the Netherlands in the period 1993-1995 using DNA fingerprinting. *American Journal of Epidemiology* 1998;147:187-95.
13. Pfyffer GE, Strassle A, Rose N, Wirth R, Brandli O, Shang H, et al. Transmission of tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprinting. *European Respiratory Journal* 1998;11:804-8.
14. Wilkinson D, Pillay M, Crump J, Lombard C, Davies GR, Sturm AW, et al. Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in rural Africa. *Tropical Medicine and International Health* 1997;2:747-53.
15. Warren R, Richardson M, van der Spuy G, Victor T, Sampson S, Beyers N, et al. DNA fingerprinting and molecular epidemiology of tuberculosis: use and interpretation in an epidemic setting. *Electrophoresis* 1999;20:1807-12.
16. Anh DD, Borgdorff MW, Van LN, Lan NT, van Gorkom T, Kremer K. *Mycobacterium tuberculosis* Beijing genotype emerging in Viet Nam. *Emerging Infectious Diseases* 2000;6:302-5.
17. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, et al. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *Journal of Clinical Microbiology* 1995;33:3234-8.
18. Pineda-Garcia L, Ferrera A, Hoffner SE. DNA fingerprinting of *Mycobacterium tuberculosis* strains from patients with pulmonary tuberculosis in Honduras. *Journal of Clinical Microbiology* 1997;35:2393-7.
19. Doroudchi M, Kremer K, Basiri EA, Kadivar MR, Van Soolingen D, Ghaderi AA, et al. IS6110-RFLP and spoligotyping of *Mycobacterium tuberculosis* isolates in Iran. *Scandinavian Journal of Infectious Diseases* 2000;32:663-8.
20. Glynn JR, Vynnycky E, Fine PE. Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. *American Journal of Epidemiology* 1999;149:366-71.
21. Murray M. Sampling bias in the molecular epidemiology of tuberculosis. *Emerging Infectious Diseases* 2002;8:363-9.
22. Vynnycky E, Nagelkerke N, Borgdorff MW, van Soolingen D, van Embden JD. The effect of age and study duration on the relationship between ''clustering'' of DNA fingerprint patterns and the proportion of tuberculosis disease attributable to recent transmission. *Epidemiology and Infection* 2001;126:43-62.
23. Murray M. Determinants of cluster distribution in the molecular epidemiology of tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99:1538-43.
24. Godfrey-Faussett P, Sonnenberg P, Shearer SC, Bruce MC, Mee C, Morris L, et al. Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet* 2000;356:1066-71.
25. Fandinho FC, Kritski AL, Hofer C, Conde H, Ferreira RM, Saad MH, et al. RFLP patterns and risk factors for recent tuberculosis transmission among hospitalized tuberculosis patients in Rio de Janeiro, Brazil. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 2000;94:271-5.
26. Yang ZH, Rendon A, Flores A, Medina R, Ijaz K, Llaca J, et al. A clinic-based molecular epidemiologic study of tuberculosis in Monterrey, Mexico. *International Journal of Tuberculosis and Lung Disease* 2000;5:313-20.
27. March F, Coll P, Guerrero RA, Busquets E, Cayla JA, Prats G, et al. Predictors of tuberculosis transmission in prisons: an analysis using conventional and molecular methods. *AIDS* 2000;14:525-35.
28. Murray MB, Alland DA. Methodological problems in the molecular epidemiology of tuberculosis. *American Journal of Epidemiology* 2002;155:565-71.
29. Nardell E, McInnis B, Thomas B, Weidhaas S. Exogenous reinfection with tuberculosis in a shelter for the homeless. *New England Journal of Medicine* 1986;315:1570-4.
30. Small PM, Shafer RW, Hopewell PC, Singh SP, Murphy MJ, Desmond E, et al. Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. *New England Journal of Medicine* 1993;328:1137-44.
31. Shafer RW, Singh SP, Larkin C, Small PM. Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in an immunocompetent patient. *Tubercle and Lung Disease* 1995;76:575-7.
32. van Rie A, Warren R, Richardson M, Victor TC, Gie RP, Enarson DA, et al. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *New England Journal of Medicine* 1999;341:1174-9.
33. Bandera A, Gori A, Catozzi L, Degli Esposti A, Marchetti G, Molteni C, et al. Molecular epidemiology study of exogenous reinfection in an area with a low incidence of tuberculosis. *Journal of Clinical Microbiology* 2000;39:2213-8.
34. Caminero JA, Pena MJ, Campos-Herrero MI, Rodríguez JC, Afonso O, Martin C, et al. Exogenous reinfection with tuberculosis on a European island with a moderate incidence of disease. *American Journal of Respiratory and Critical Care Medicine*. 2000;163:717-20.
35. Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. *Trends in Microbiology* 2002;10:45-52.

36. Zhang M, Gong J, Yang Z, Samten B, Cave MD, Barnes PF. Enhanced capacity of a widespread strain of *Mycobacterium tuberculosis* to grow in human macrophages. *Journal of Infectious Diseases* 1999;179:1213-7.

37. Valway SE, Sanchez MP, Shinnick TF, Orme I, Agerton T, Hoy D, et al. An outbreak involving extensive transmission of a virulent strain of *Mycobacterium tuberculosis*. *New England Journal of Medicine* 1998;338:633-9.

38. Kelley CL, Collins FM. Growth of a highly virulent strain of *Mycobacterium tuberculosis* in mice of differing susceptibility to tuberculous challenge. *Tubercle and Lung Disease* 1999;79:367-70.

39. Manca C, Tsenova L, Barry CE, Bergtold A, Freeman S, Haslett PA, et al. *Mycobacterium tuberculosis* CDC1551 induces a more vigorous host response *in vivo* and *in vitro*, but is not more virulent than other clinical isolates. *Journal of Immunology* 1999;162:6740-6.

40. Bishai WR, Dannenberg AM, Parrish N, Ruiz R, Chen P, Zook BC, et al. Virulence of *Mycobacterium tuberculosis* CDC1551 and H37Rv in rabbits evaluated by Lurie's pulmonary tubercle count method. *Infection and Immunity* 1999;67:4931-4.

41. Sepkowitz KA. How contagious is tuberculosis? *Clinical Infectious Diseases* 1996;23:954-62.

42. Middlebrook G, Cohn ML. Some observations on the pathogenicity of isoniazid-resistant variants of the tubercle bacilli. *Science* 1953;118:297-9.

43. Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* 1992;358:591-3.

44. Snider DE, Kelly GD, Cauthen GM, Thompson NJ, Kilburn JO. Infection and disease among contacts of tuberculosis cases with drug-resistant and drug-susceptible bacilli. *American Review of Respiratory Diseases* 1985;13:125-32.

45. Teixeira L, Perkins MD, Johnson JL, Keller R, Palaci M, do Valle Dettoni V, et al. Infection and disease among household contacts of patients with multidrug-resistant tuberculosis. *International Journal of Tuberculosis and Lung Disease* 2001;5:321-8.

46. García-García ML, Ponce de León A, Jiménez-Corona ME, Jiménez-Corona A, Palacios-Martínez M, Balandrano-Campos S, et al. Clinical consequences and transmissibility of drug-resistant tuberculosis in southern Mexico. *Archives of Internal Medicine* 2000;160:630-6.

47. van Soolingen D, Borgdorff MW, de Haas PE, Sebek MM, Veen J, Dessens M, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *Journal of Infectious Diseases* 1999;18:726-36.

48. Bifani PJ, Mathema B, Liu Z, Moghazeh SL, Shopsin B, Tempalski B, Driscol J, et al. Identification of a W variant outbreak of *Mycobacterium tuberculosis* via population-based molecular epidemiology. *JAMA* 1999;282:2321-7.

49. Githui WA, Hawken MP, Juma ES, Godfrey-Faussett P, Swai OB, Kibuga DK, et al. Surveillance of drug-resistant tuberculosis and molecular evaluation of transmission of resistant strains in refugee and non-refugee populations in North-Eastern Kenya. *International Journal of Tuberculosis and Lung Disease* 2000;4:947-55.

50. Friedman CR, Stoeckle MY, Kreiswirth BN, Johnson WD, Manoach SM, Berger J, et al. Transmission of multidrug-resistant tuberculosis in a large urban setting. *American Journal of Respiratory and Critical Care Medicine* 1995;152:355-9.