CHAPTER 1

Pulmonary Tuberculosis and *Mycobacterium Tuberculosis*: Modern Molecular Epidemiology and Perspectives

Sylvain Godreuil,[1] Loubna Tazi,[2] and Anne-Laure Bañuls[1]

[1]*Génétique et Evolution des Maladies Infectieuses (UMR CNRS/IRD 2724), IRD de Montpellier, 911 av Agropolis BP 64501, 34394 Montpellier Cedex 5, France*

[2]*Department of Microbiology & Molecular Biology, Brigham Young University, 775 Widtsoe Building, Provo, UT 84602, USA*

1.1 INTRODUCTION

Tuberculosis (TB) is a bacterial infection caused mainly by *Mycobacterium tuberculosis* (MTB). The development of paleopathology and paleoepidemiology in infectious diseases has proven the *very ancient origin* of this disease. TB may have plagued humans at least since the Neolithic times [64,158,186]. This infectious disease was sporadic until the 1700s and became epidemic afterward because of the *industrial revolution*, the *increase in population density*, and *unfavorable living conditions*. Furthermore, human migrations and colonization of countries and continents helped to spread TB, which became an endemic disease. In 1882, *Robert Koch* managed to isolate the tubercle bacillus (called also *Koch bacillus*), the bacterium responsible for TB, and he established TB as an infectious disease. Over the last 100 years, TB has probably killed 100 million people [66]. In the twentieth century, the incidence of this disease began to decline rapidly in developed countries where the sanitation and housing conditions were improved. This scenario was accelerated by the introduction of *BCG vaccine* (Bacillus Calmette Guérin, 1921) and the use of antimicrobials as anti-TB agents, such as *streptomycin* (1943), *isoniazid* (1952), and *rifampin* (1963). However, despite these efforts to eradicate this disease, the incidence of TB increased in the 1980s. The emergence of *multidrug-resistant strains* and the *high incidence of human immunodeficiency virus* (*HIV*) have strongly contributed to this phenomenon. Nevertheless, the success of propagation of this disease remains directly related to the *social and hygiene conditions* of human populations.

TB remains a major public health problem worldwide and the first cause of mortality attributable to a single infectious agent, especially in developing countries where the consequences of this disease remain more serious and the infection risks are higher [54,148] (see Fig. 1.1). According to the World Health Organization (WHO, www.who.int), the estimated number of cases of TB worldwide in 2003 was 8.8 million, 3.9 million of which were sputum positive, and deaths from TB (including TB deaths in people infected with HIV) were 1,747,000. Furthermore, it is estimated that there are currently 2.1 billion people worldwide who are latently infected with the tubercle bacillus and could develop the active form of the disease in the case of reactivation.

The recrudescence of this disease in several countries, the emergence of multidrug-resistant strains and the association of TB with the HIV pandemic show the need to improve research on this pathogen in applied and basic research in order to better understand the transmission of TB and to eradicate this disease. The objective of this chapter is to give an overview of the biology, genetics, and pathogeny of *M. tuberculosis* (MTB), to describe the current molecular methodologies available for identifying the MTB populations responsible for the spread of TB and the outbreaks, and to show the contribution of genetic epidemiology studies in understanding global and local epidemiology of TB.

Encyclopedia of Infectious Diseases: Modern Methodologies, by M. Tibayrenc

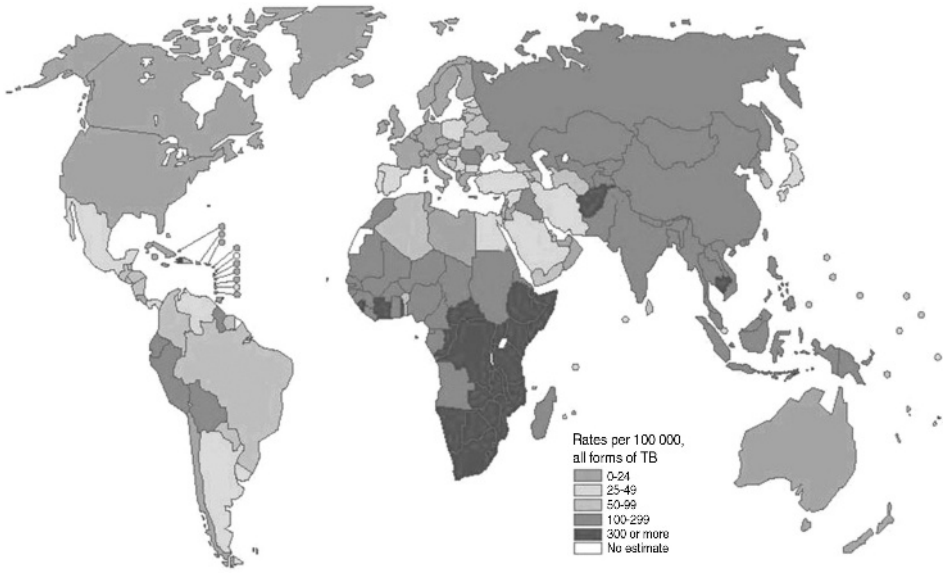

Fig. 1.1. Estimated TB incidence rates, 2003 (WHO website: http://www.who.int/tb/publications/global_report/2005/results/en/index.html)

1.2 GENERAL POINTS ON *MYCOBACTERIUM TUBERCULOSIS* (MTB) AND PULMONARY TUBERCULOSIS (PTB)

1.2.1 Classification and Cellular Characteristics

MTB is a bacterium belonging to the *Mycobacterium* genus, which is the only genus in the Mycobacteriaceae family (Chester, 1897), Actinomycetales order (Buchanan, 1917), and Actinomycetes class (Krasil'nikov, 1949).

The *Mycobacterium* genus, one of the most extensively studied bacterial taxa, was described by Lehman and Neuman in 1896. Its identification is based on the following characteristics: shape of the colonies, growth rate, and biochemical reactivity. To date, 71 species have been described within this genus, and they are subdivided in two main groups based on their growth rates (fast vs. slow) [109,154,163]. The rapidly growing *Mycobacterium* species (species producing visible colonies within 7 days under optimal culture conditions) are mainly common saprotrophs of natural habitats. Only a few of them can be pathogenic for humans or animals (e.g., *Mycobacterium abscessus*, *M. fortuitum*, *M. porcinum*), whereas the majority are nonpathogenic (e.g., *M. smegmatis*, *M. agri*). In contrast, the majority of the slowly growing *Mycobacterium* species are pathogenic for humans and/or animals (e.g., all the species of the MTB complex [MTBC], *M. leprae*, *M. ulcerans*, *M. avium*), and only a few of them are nonpathogenic (e.g., *M. terrae*, *M. gordonae*). This chapter will focus particularly on the *MTB species* belonging to the MTBC. This complex is composed of seven different species, MTB (Koch, 1882), *M. bovis* (Karlsen and Lessel, 1970), *M. africanum* [25], *M. microti* (Reed, 1957), "*M. canettii*" (still not officially recognized on the list of Bacterial Names with Standing in Nomenclature, http://www.bacterio. cict.fr), and recently, *M. caprae* [5] and *M. pinnipedii* [34]. Each member of MTBC is associated with a specific primary host, although infection is known to occur in various alternative hosts. The species responsible for TB in humans and for which no animal reservoirs was found are MTB, *M. africanum*, and *M. canettii*. MTB is the main species of human TB, the other species are less frequent in humans (e.g., *M. africanum* is characterized mainly in Africa, and *M. canettii* was isolated in a few cases of human TB in East Africa). *M. bovis* is principally the agent of bovine TB, but this species can also be pathogenic for humans, with the number of cases related to such infection probably underestimated [6]. Furthermore, an attenuated strain of *M. bovis*, *M. bovis BCG* (Calmette and Guérin, 1921) is used as a *vaccine* for preventing human TB (see below for more details on BCG vaccine). The other species are also isolated specifically in animals, such as *M. microti*, which is the agent of rodent TB, *M. caprae*, which predominantly affects cattle, and *M. pinnipedii*, which has Pinnipeds as natural hosts. These latter species can affect other animal species and humans to a very limited extent [5,34]. The members of the MTBC, as well as all mycobacteria species, are *rod-shaped bacteria* (0.2–0.6 μm wide, 1–10 μm long), nonmotile, nonencapsulated, Gram-positive, aerobes (growing most successfully in tissues with a high oxygen content such as lungs), or facultative anaerobes. They are *facultative intracellular pathogens*, usually infecting mononuclear phagocytes (e.g., macrophages). As deduced from its genome, MTB has the potential to manufacture all of the machinery necessary to synthesize its essential vitamins, amino acids, and enzyme cofactors. *MTB* has an unusual cell wall, with an additional layer beyond the peptidoglycan layer, which is rich in unusual lipids, glycolipids, and polysaccharides. These bacteria can be detected by optical

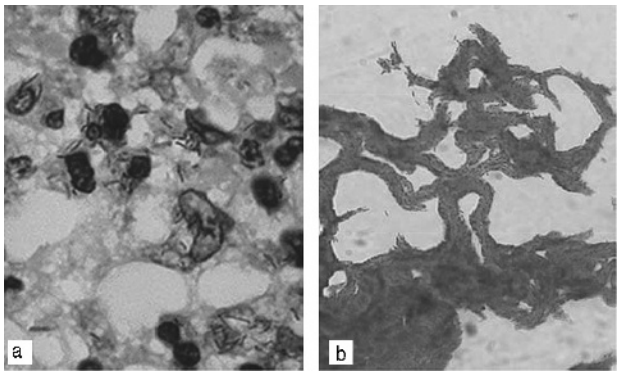

Fig. 1.2. *Mycobacterium tuberculosis* (MTB): *Ziehl–Neelsen* coloration (a) from sputum; (b) from MTB culture in liquid medium (photo taken by S. Godreuil, all rights reserved). See color plates.

microscopy after *Ziehl–Neelsen* (*ZN*) *acid-fast stain* of sputum from a person with active TB (Fig. 1.2). Bacilli appear as thin red rods in the microscopic field, whereas all other materials in the sputum pick up the blue counterstain.

1.2.2 Transmission and Multiplication of MTB (See Also Chapter 5)

TB is considered a disease with an *interhuman transmission*. Tuberculous bacilli are spread out by infected patients coughing, sneezing, or speaking, and they can be inhaled by another individual in close contact. The inhalation of these sprays, called Flugge's droplets—small aerodynamic particles—presents a risk of tuberculous infection. These particles can also remain in the air and play the role of reservoir.

The tubercle bacillus enters the human body mainly via the respiratory tract through the inhalation of the droplets sprayed in the air (Fig. 1.3). These particles are small enough to be able to reach the lower respiratory tract. Indeed, among the infectious particles inhaled, only those with two or three bacilli can reach the bronchic cells, the largest ones are stopped upstream and eliminated [44]. The success of such infection and the development of the pulmonary form of TB depend on four successive stages: bacilli phagocytosis, intracellular multiplication, the stationary stage, and the pulmonary form of TB (see also Chapter 5). These different stages can evolve into different outcomes: spontaneous healing, acute tuberculosis, latent infection, and reactivation or reinfection (see Fig. 1.3).

(i) *Bacilli phagocytosis*: The bacilli that reach the pulmonary alveolus are phagocyted by the mature macrophages. This step, which takes place in the first week following particle inhalation, is the first stage of infection, and it depends on two main factors: the bacillus virulence and the bactericidic activity of the macrophage. In general, the bacteria are destroyed by the alveolar macrophages and the infection is stopped at this stage, otherwise they begin an intracellular cycle of multiplication [119,179].

(ii) *Intracellular multiplication*: This second stage occurs between the 7th and the 21st day. It corresponds to intracellular bacilli multiplication in the macrophage alveoli and is also called the symbiotic stage. Indeed, the bacteria that are not destroyed by the alveolar macrophages will multiply. They are released after cellular lysis, and can thus infect other circulating macrophages and continue their multiplication. At the end of this stage, due to a symbiosis event, a huge number of macrophages and bacilli are concentrated at the level of early pulmonary lesions [44].

(iii) *Stationary stage*: Following the induction of the immune response of the host, particularly cell-mediated immunity

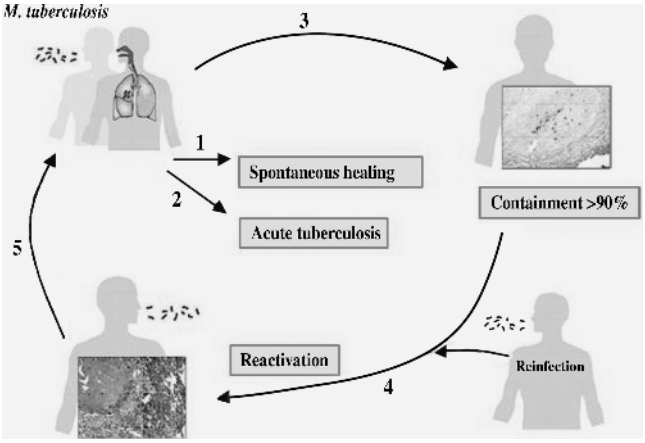

Fig. 1.3. Inspired by Kaufmann and McMichael [96] with permission: *Mycobacterium tuberculosis* (MTB) enters the host within inhaled droplets. Different outcomes are possible. (1) Immediate eradication of MTB by the pulmonary immune system. (2) Infection transforms into active tuberculosis. (3) Infection does not transform into disease because MTB is contained inside granulomas. (4) After a latency phase, MTB can become active after either an endogenous reactivation or an exogenous reinfection or both. (5) At this stage, there is dissemination and transmission of MTB.

(see Chapter 5), bacteria growth becomes stationary [142]. This is the third stage of the infection called *primary infection*. Because of a delayed-type hypersensitivity, the macrophages in which bacilli multiply are destroyed. Bacterial toxins and cellular products are released, and this leads to the formation of solid caseous necrosis [137], where a pseudo-equilibrium settles between inactivated and mature macrophages. At this stage, either the number of infected cells in the caseous center decreases if the released bacilli are phagocyted by the mature macrophages or it increases if the bacilli multiply in the inactivated macrophages. Thus, the progression of the disease depends on which macrophage type prevails [43,44]. At this stage, bacilli may become dormant and never induce TB at all, which is referred to as a *latent infection* that is detected only by a *positive tuberculin skin test*; or the latent organisms can eventually begin to grow, with resultant clinical disease, known as TB reactivation.

(iv) *Pulmonary form of TB* (*PTB*): When the equilibrium between the inactivated and mature macrophages is broken, the infection reaches the last stage, the disease, PTB (see also Chapter 5). This step is characterized by the liquefaction of the caseous center, leading to the formation of a cavity detected by pulmonary radiography. The liquefied material present in this cavity constitutes an excellent growth media for the bacteria, and macrophages do not survive in this environment. *At this stage of the disease, the person becomes contagious by releasing the bacilli into the air*. Furthermore, without treatment, this individual can develop a chronic TB, presumably leading to death.

1.2.3 Clinical and Subclinical TB

The term "TB infection" refers to a positive TB skin test (see below for details) with no evidence of active disease; this state is also called latent infection (see Fig. 1.3). "*TB disease*" *refers to cases that have a positive acid-fast smear or culture for MTB or radiographic and clinical presentation of TB* [117].

The most common clinical manifestation of TB is pulmonary disease; nevertheless, extrapulmonary TB can also occur, but is little or not contagious. Without minimizing the importance of extrapulmonary TB, which currently accounts for 20% of reported cases of TB, we will focus here only on the cases of pulmonary infections. Furthermore, as described above, although some people develop active TB disease after infection, *almost all TB infections are asymptomatic and remain latent* [19].

1.2.3.1 Active disease A patient with PTB presents with the symptoms of chronic or persistent cough and sputum production. If the disease is at an advanced stage, the sputum will contain blood, and the patient will be diagnosed with lack of appetite, weight loss, fever, night sweats, and thoracic pains. Patients with PTB are classified in different categories because a specific treatment is needed for each category. The main categories are as follows:

- New case: TB in a patient who either has never received anti-tuberculous treatment or started a treatment for less than 1 month.
- Relapse: TB already treated and declared cured after sufficient treatment time, which has become active again.
- Chronic TB: A case of relapse from which the microscopic exam of expectoration remains positive after a second complete treatment.
- Primary resistance case: This characterizes the bacilli that are resistant to treatments, although patients have never been treated by anti-tuberculous drugs (see below).
- Multiresistance case: MTB resistant at least to both major anti-tuberculous drugs (isoniazid and rifampin) (see below).

1.2.3.2 Latent infection MTB in a latent state can subsequently reactivate to cause active disease. The latent state of infection is a major obstacle for eradicating TB. In latent TB, the host immune response is capable of controlling the infection but fails to eradicate the pathogen. Latent TB is the product of a complex set of interactions between MTB and the host immune response (for more details, see Chapter 5). Therefore, one-third of the world population is estimated to be infected with the pathogen in the latent stage. The bacilli remain dormant until the host defenses are impaired by a disorder such as HIV infection.

1.2.3.3 MTB and HIV *For many people, TB is the first sign of immune dysfunction associated with HIV infection, and active TB is an AIDS-defining illness*. TB is an ever-increasing concern for people with HIV. In some parts of the world, TB is the leading cause of death of people infected with HIV. Indeed, the risk of developing active TB disease after TB infection, or following an apparent cure of several years, increases considerably for people with a deficient immune system. It was calculated that in case of HIV co-infection, this risk is multiplied 50–300 times [101]. Active TB in HIV-positive patients can result from both reactivation of latent infection and primary disease. HIV increases the chance of reactivating dormant TB infection from 5% to 10% over a person's lifetime to 7% to 10% per year. In patients with low CD4 cell counts, TB arises with atypical pulmonary manifestations and extrapulmonary disease. Indeed, as the level of immunodeficiency increases with advancing HIV disease, atypical pulmonary features predominate [31]. One in 10 people living with HIV will get active TB within 1 year of being diagnosed with HIV. It can occur early in HIV disease when CD4 cell counts are relatively high, in the 300–400 range. In early HIV infection, TB usually infects and affects only the lungs. As CD4 cell counts drop, however, TB is more likely to appear in other organs also. When the immune system responds to TB, it can cause HIV levels to increase, and HIV disease may then progress more quickly. This, in turn, increases the risk of other opportunistic infections. *It is therefore very important for people with HIV to be screened regularly for TB.*

1.2.4 Diagnosis of MTB Species

1.2.4.1 Tests for active disease Tools for the diagnosis of active disease include clinical suspicion, response to treatment, chest radiographs, staining for acid fast bacilli, culture for mycobacteria, and, more recently, nucleic acid amplification assays (for more details, see review in [19]). Briefly, as described above, TB can mimic many forms of disease and must always be considered if no firm diagnosis has been made. The *chest X-ray examination* is traditionally considered as one of the most important tests, but its low specificity can lead to overdiagnosis. To confirm the diagnosis of PTB, respiratory samples (expectorated sputum) are submitted to the bacteriological laboratory for *microscopic examination* and for *mycobacterial culture*. The microscopic examination consists of making a smear of sputum and staining by the *Ziehl–Neelsen* (*ZN*) *method* (see Fig. 1.4). This technique is used in most low-income countries because it is inexpensive and easy to use [144], but its low sensitivity (43–55%) is a major drawback [214]. Cultures increase the sensitivity for diagnosing MTB and allow drug-susceptibility testing and genotyping for epidemiological purposes (see below). Nevertheless, culturing TB is time consuming and the cost is often too high, resulting in reliance solely on microscopy of sputum smear in resource-poor countries [19]. Two types of culture media are commercialized: solid media, which includes egg-based media (Lowenstein-Jensen; see Fig. 1.4), and liquid media (such as BACTEC systems, Becton Dickinson, Sparks, MD, USA). Several studies showed that liquid media can decrease recovery time (2 weeks instead 4–12 weeks) of mycobacteria culture and increases the sensitivity compared to solid media, which remains the reference media for culturing mycobacteria [124,162]. The traditional methods of drug-susceptibility testing relied on culture inoculated with antibiotics and thus can also require several weeks to obtain results.

These methods remain the gold standard for diagnosis, but the development of DNA probes and the polymerase chain reaction (PCR) assays now provide more sensitive and rapid diagnosis for species identification as well as for analyzing drug susceptibility. Currently, two main methods approved by the Food and Drug Administration (FDA) are available: a PCR-based test targeting a specific portion of the 16S ribosomal RNA gene (Roche) and a transcription-mediated amplification of 16S ribosomal gene transcripts with product detection performed via chemiluminescence (GenProbe). Furthermore, Kaul [97] as well as Brodie and Schluger [19] detailed in their reviews all the latest techniques based on nucleic acid amplification. These diagnostic tests have considerably decreased the diagnosis recovery time and increased the sensitivity for smear-positive and smear-negative specimens. These techniques are the most promising methodologies for diagnosing the 15–20% of adults with TB having negative sputum culture and among children, for whom the proportion of culture-negative cases is much higher. At present, the greatest problem concerning these techniques is the cost; consequently, they are not affordable for resource-poor countries [97].

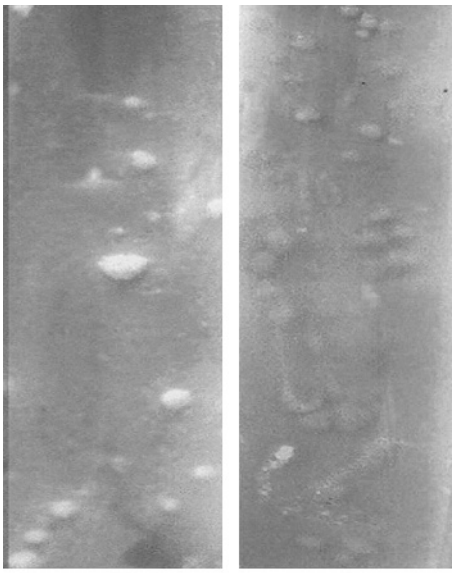

Fig. 1.4. Colonies of *Mycobacterium* on Lowenstein–Jensen. (a) *M. tuberculosis* colonies are irregular, rough, eugonic, beige in color. (b) Atypical mycobacteria colonies are small, smooth, and pigmented (photo taken by S. Godreuil, all rights reserved). See color plates.

During the past few years, there has also been great progress in exploring drug susceptibility in MTB. In their review, Brodie and Schluger [19] described the latest methodologies allowing for a rapid detection of the drug-resistant mutations from smear-positive respiratory specimens or from culture specimens with their limits and advantages: line probe assays, molecular beacons, phage amplification, and luciferase reporter phages.

1.2.4.2 Tests for latent infection It has been demonstrated that almost all TB infections are asymptomatic and remain latent, with a rate of *reactivation* in active disease in approximately 5–10% of infected individuals. These people are a reservoir for the disease and a major barrier to the ultimate control and elimination of TB. Until very recently, a skin reaction Mantoux test or tuberculin skin test or PPD (purified protein derivative) skin test was the only available test to detect latent disease or to confirm the cases of active disease with negative sputum smear or culture. Nevertheless, this test presents various problems such as relatively poor sensitivity and specificity. Recently, a new generation of tests has been developed: QuantiFERON-TB and QuantiFERON-TB Gold (QFN-Gold) tests (Cellestis Limited, St. Kilda, Australia) and the T SPOT-TB test (Oxford Immunotec, Oxford, UK). These tests are based on the detection in serum of either the release of IFN-γ (QuantiFERON) or detection of the T cells themselves (T SPOT-TB) (for more details see review of Brodie and Schluger [19]). These tests seem to improve specificity and sensitivity. At present, in most high-burden, resource-poor countries, latent infections are neither diagnosed nor treated. However, for the TB control and for

stopping the disease progression, it would be imperative to treat the latent infections.

1.2.4.3 Diagnosis in smear-negative PTB and the special case of HIV patients Patients with smear-negative PTB have been found to be less infectious and have a lower mortality rate, but a significant proportion (50–71%) progressed to active disease warranting treatment [31].

Furthermore, several studies showed that with the increase in the HIV/TB coinfection, there has been a disproportionate increase in the reported rate of smear-negative disease (see review in [31] for details). Indeed, different data reports suggest that smear-negative disease is actually more common among HIV-infected patients. Colebunders and Bastian [31] hypothesize that smear examinations have proven less sensitive, as the level of immunodeficiency has been increasing with advanced HIV disease. This is a crucial problem in developing countries where microscopic examination (cheap, simple, and rapid) is the basis of TB diagnosis. Nevertheless, several studies showed that also in HIV-uninfected populations, a non-negligible rate of smear-negative PTB can be observed (see [31] for details). This is normally associated with low bacilli burdens and minimal pulmonary lesions. Furthermore, this is especially more common among children and elderly patients.

As described above, in HIV patients, atypical pulmonary features predominate and chest radiography changes may be atypical or attributable to other infections. Furthermore, the tuberculin skin testing is confounded, especially in developing countries presenting a high rate of HIV/MTB co-infection, with the high coverage of BCG vaccination, with asymptomatic TB infection, with the presence of nontuberculous mycobacteria, and with anergy due to HIV or malnutrition.

Culture and PCR remain the most sensitive techniques, as they can produce a positive result for specimens containing as few as 10 bacilli. This is of a great interest, as HIV-positive patients generally produce sputum with low bacilli loads [31]. Nevertheless, these diagnostics are still financially inaccessible for resource-poor countries.

For these reasons, some authors have proposed various management algorithms to optimize the number of patients correctly treated for smear-negative sputum and thus a major part of HIV-positive patients. In their review, Colebunders and Bastian [31] detail the different parameters included in these management algorithms. They are based on the following combined features: clinical symptoms, response to antibiotic trials, smear investigations, and chest radiography.

1.2.5 Treatment, Drug Resistance, and Control

1.2.5.1 Treatment According to the current recommendations, *effective TB drug therapy requires at least two effective drugs* [157]. Sahbazian and Weis [157] detail how this axiom has emerged to limit drug resistance, which is, along with lack of patient observance, the most important factor of chemotherapy failure. These authors also review available drugs and their toxicity [157]. Of the drugs approved by the Food and Drug Administration (FDA), isoniazid, rifampin, ethambutol, and pyrazinamide are considered first-line anti-TB drugs. Rifapentine and rifabutin can also be considered as first-line drugs under special conditions. The others (see review of Sahbazian and Weis [157]) are categorized as second-line drugs, which are used when the first-line drugs are unsuitable because of drug intolerance or infection with drug-resistant TB. The WHO's Stop TB Department, with the help of the International Union Against TB and Lung Disease and experts worldwide, has published guidelines for a standardized and efficient treatment of TB (http://www.who.int/tb/en/index.html). They give practical guidance for national TB programs and for the medical profession in the effective management of TB. Different targets of TB treatment are reviewed, such as principles of treatment, with an update of the guidelines, care in the context of HIV/TB infection, multidrug-resistant TB (MDR-TB), and chronic diseases. Guidelines for high-income and low-incidence countries, even though they follow the same principles, include recommendations that may not be appropriate for most high-incidence countries where resources for TB control are often limited. The most cost-effective public health measure for the control of TB is the identification and cure of infectious TB cases, that is, patients with smear-positive PTB. Nevertheless, national TB programs provide guidelines for identification and cure of all patients with TB. These guidelines cover the treatment of patients, both adults and children, with smear-positive PTB, smear-negative PTB, and extrapulmonary tuberculosis. *It is important to note that TB treatments require long-term drug administration, which is logistically difficult and generally results in uncontrolled disease burden in developing countries*.

1.2.5.2 Vaccination *Vaccines are desperately needed* because of several factors such as duration and cost of existing treatment, cost of diagnosis, rate of drug resistance, difficult access to cure for poor-resource populations, and the high rate of latent infections. At present, the only available vaccine is the *BCG* (*M. bovis* bacillus Calmette–Guérin), which is a deletional mutant of *M. bovis* that arose spontaneously during subculture on beef-bile-potato medium [135]. In their review, Rook et al. [155] expose the limits of BCG and demonstrate that while it helps to protect against childhood forms of TB, it provides variable protection in adults and it has a minimal impact on disease control in developing countries where the vaccine is most needed. According to Hampton [84] and Ginsberg [73], hundreds of new TB vaccine candidates are under study, including subunit vaccines, consisting of immunogenic mycobacterial components; DNA vaccines; live, attenuated mycobacteria; and live, attenuated nonmycobacterial vectors, such as *Salmonella* or vaccinia virus [73]. Rook et al. [155] highlight that with the aim of developing a successful vaccine, it is crucial to refer to the immunopathogenesis of TB and to consider the immune response, which can differ depending on target populations

and the individual immune status (other infections, nutritional status, etc.).

1.2.5.3 DOTS strategy WHO and the International Union Against TB and Lung Disease (IUATLD) have adopted *directly observed therapy short course* (*DOTS*) as the main strategy for TB control. DOTS consists of political and administrative commitment; case detection by sputum microscopy; standardized short-course chemotherapy given under direct observation by a health professional; adequate supply of good-quality drugs; systematic monitoring; and accountability for every patient diagnosed. Frieden and Munsiff [65] review the principle, scientific basis, and experience with implementation of DOTS. According to a WHO report [213], the number of countries having adopted and launched the DOTS strategy has increased considerably since 1995, and in 2003, it had been implemented in 182 of 211 countries, covering 77% of the world's population. In 132 countries, including most of the industrialized world, DOTS is available to more than 90% of their populations. DOTS programs concur to decrease mortality rates, which are often drastically lower than in non-DOTS programs. This is truly one of the great public health success stories of the past decade. According to WHO data and details given by Frieden and Munsiff [65], DOTS has saved more than 1 million lives in the last 10 years and could save millions of lives over the next 10 years. Nevertheless, there are a number of obstacles to DOTS expansion, four of which were identified to be of overriding importance by WHO: shortages of trained staff, lack of political commitment, weak laboratory services, and inadequate management of MDR-TB and TB in people infected with HIV.

1.3 GENETICS OF MTB, MOLECULAR TOOLS, AND POPULATION STRUCTURE

1.3.1 Genome and Genetic Diversity of MTB

The genome of MTB is *haploid*, as are all bacteria genomes, and is composed of 4,411,529 base pairs (bp). It contains approximately 4000 genes and presents a rich composition in GC content (65%) [30]. This genome is characterized by the presence of numerous repeated sequences. No plasmid was detected in this species. In 1997, Sreevatsan et al. [177] studied 26 structural genes or loci, and they observed *very low levels of genetic variation*. From these results, they concluded that the genetic diversity of the species is localized, especially in transposable elements and in genes involved in host–pathogen interactions, particularly those related to host immunological responses. This last point was refuted by Musser et al. [130] after their study conducted on 24 genes coding for targets of the host immune systems. Of the 24 genes, 19 were monomorphic and the last five appeared slightly polymorphic (only six polymorphic nucleotide sites on all five genes). On the contrary, *the transposable elements show high levels of genetic polymorphism*, and they are widely used for studying the genetic variability in the MTB species [17,95,173] (see below for more details). Nevertheless, the only way to detect the real genetic diversity of an organism is the whole sequencing of several genomes from different clinical isolates. For MTB, the complete genome sequences of three strains, but also of one *M. bovis* strain, are now available (www.tigr.org) [63,70]. In addition, a sequencing project for *M. bovis* BCG is ongoing (http://www.sanger.ac.uk/Projects/Microbes). The comparison of the complete sequences of the two strains (H37Rv, which is the classical reference strain, and one recent MTB strain CDC1551) confirms a much higher degree of polymorphism than previously thought [63,82]. These latter studies made it possible to identify large-sequence polymorphisms (LSPs) and single-nucleotide polymorphisms (SNPs), whereas the molecular basis of variability in virulence and transmissibility remains undefined. Tsolaki et al. [199] have developed a complementary approach to comparative genomics involving the analysis of unsequenced genomes by DNA microarray. Although this approach is limited in the identification of relatively LSPs, it allows the comparison of a large number of genomes and thus provides information on the diversity and frequency of polymorphisms among different strains from a single population. These authors postulate that because rates of SNPs are low in this species, large sequence differences that are detectable by microarray are likely to be an important source of genetic variation. They identified 68 different LSPs (representing 186,137 bp, or 4.2% of the entire genome) that are present in H37Rv but absent from several clinical isolates. A total of 224 genes (5.5%), including genes in all major functional categories, were found to be partially or completely deleted. Deletions are not distributed randomly throughout the genome but instead tend to be aggregated. They observed that the identified deletions were evidently unessential to the development of the disease, as they were found in active clinical cases. In contrast, their frequency spectrum suggested that most polymorphisms are weakly deleterious to the pathogen. These results raise numerous opportunities to advance in the study of drug resistance, virulence, and host–pathogen interactions.

1.3.2 Genetic Tools for Molecular Epidemiology

Because it is still not possible to sequence the whole genome of MTB populations to conduct molecular epidemiology studies, in the last decade, a large number of different molecular methods based on DNA fingerprints have been developed. Several molecular techniques are available to explore the genetic diversity of MTB populations and are useful for epidemiological surveillance and understanding of TB transmission. We will detail here only the three main techniques classically used in molecular epidemiology studies. All three of these molecular tools are based on the study of transposable and repetitive elements of the MTB genome: IS*6110* (Insertion Sequence *6110*) based restriction fragment length polymorphism (RFLP) genotyping, spoligotyping, and MIRU-VNTR

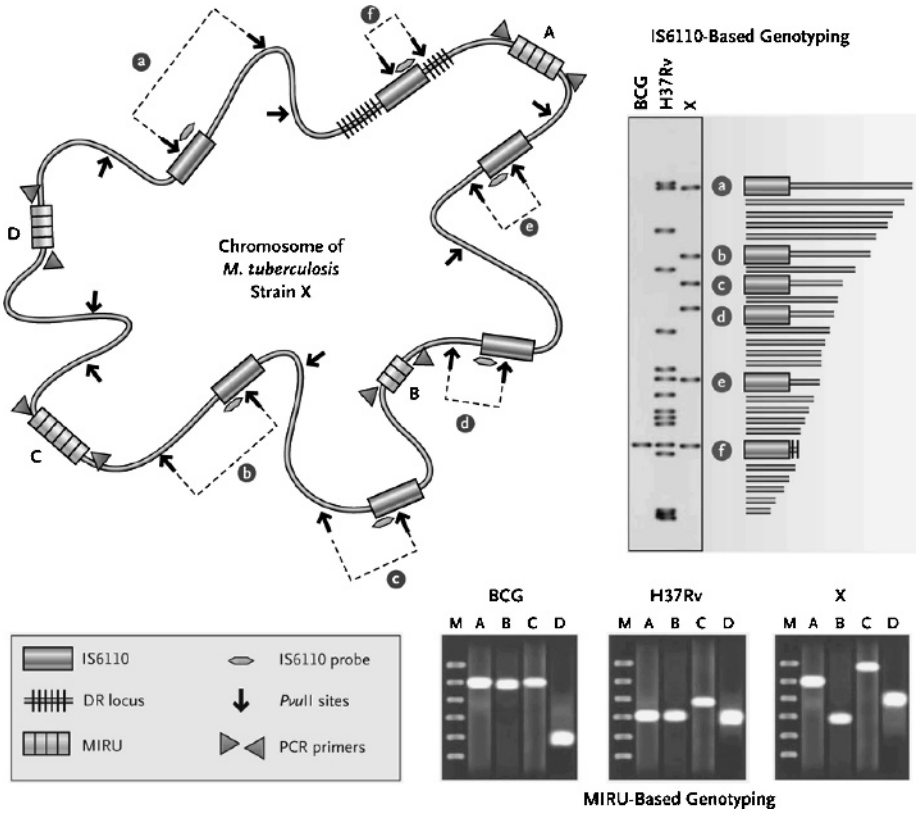

Fig. 1.5. Chromosome of *Mycobacterium tuberculosis* (MTB) hypothetical strain X and genotyping of *M. bovis* bacille Calmette–Guérin (BCG), the MTB laboratory strain H37Rv, and strain X on the basis of IS*6110* insertion sequences and mycobacterial interspersed repetitive units (MIRUs). The top left-hand panel shows the chromosome of hypothetical strain X, as shown by the arrows. The top right-hand panel shows the results of IS*6110*-based genotyping. Mycobacterial DNA is digested with the restriction enzyme *Pvu*II. The IS*6110* probe hybridizes to IS*6110* DNA to the right of the *Pvu*II site in IS*6110*. The size of each hybridizing fragment depends on the distance from this site to the next *Pvu*II site in adjacent DNA (fragments a through f), as reflected by gel electrophoresis of the DNA fragments of BCG, H37Rv, and X. The horizontal lines to the right of the electrophoretic strip indicate the extent of the distribution of fragments in the gel, including *Pvu*II fragments that contain no IS*6110*. The three bottom panels show the results of MIRU-based genotyping. MIRUs contain repeat units, and MIRU analysis involves the use of polymerase chain reaction (PCR) amplification and gel electrophoresis to categorize the number and size of repeats in 12 independent loci, each of which has a unique repeated sequence. The sizes of molecular-weight markers (M) and PCR products for the loci A, B, C, and D in BCG, H37Rv, and X are shown. The specific sizes of the various MIRUs in each strain result in a distinctive fingerprint for the strain (from [10], with permission).

(mycobacterial interspersed repetitive units-variable number tandem repeats) (see below for details of each technique). Figures 1.5 and 1.6 illustrate these three techniques and display the different genetic elements in the MTB genome.

1.3.2.1 IS6110-based RFLP genotyping (See Fig. 1.5)

Until recently, this technique was *the gold standard* approach for genotyping MTB isolates. IS*6110* is an insertion sequence that was identified in the MTBC by Thierry et al. [191]. Through a RFLP analysis, these insertion sequences have been used as epidemiological tools since 1991 [139]. They vary in copy number and may have different integration sites in different strains. From a technical point of view, extracted DNA from a bacterial culture is digested with the restriction endonuclease *Pvu-II*. DNA fragments are then separated according to their molecular weight by gel electrophoresis. The gel is then transferred and hybridized by a specific probe of IS*6110* elements, resulting in easily readable band patterns. The three strains presented in Figure 1.5 (BCG, H37Rv, and X) differ in the number of bands corresponding to the number of IS*6110* copies in the genome and the location of the bands. The protocol of this technique is well standardized, providing results that are comparable between laboratories, and large databases are available (http://www.caontb.rivm.nl/) [81]. Nevertheless, this technique presents several disadvantages. First, it requires culture of MTB and a large amount of DNA. Second, this genotyping method is very time-consuming, labor-intensive, and technically demanding. Third, it has

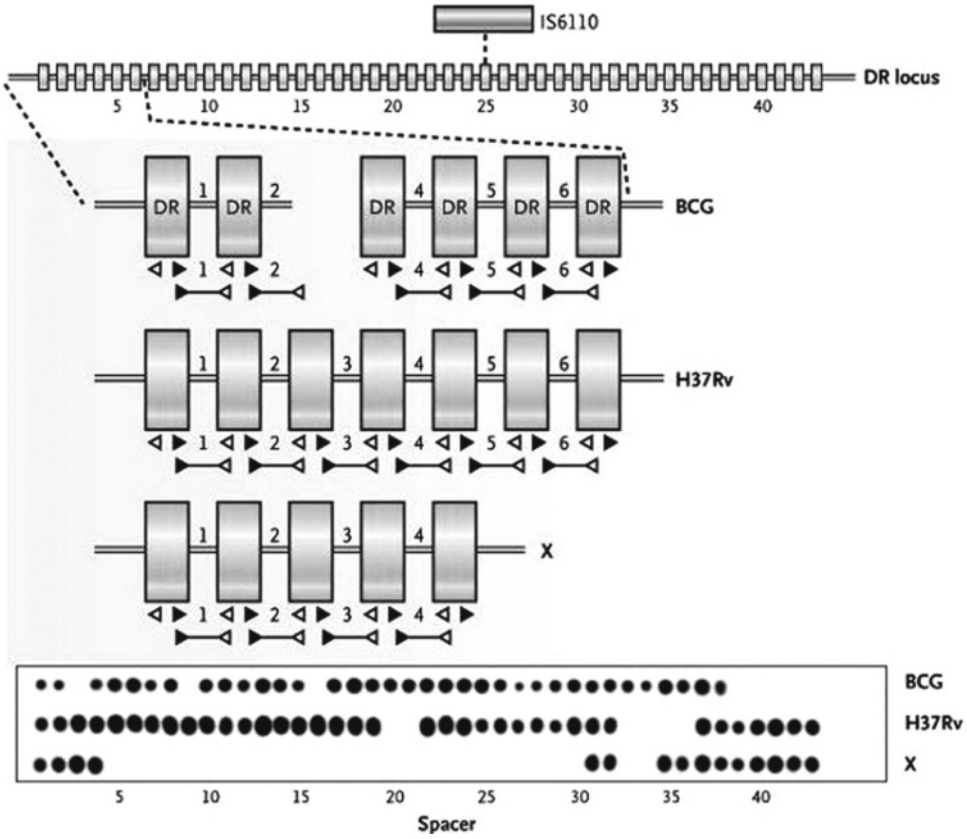

Fig. 1.6. Spoligotyping. The direct-repeat (DR) locus is a chromosomal region that contains 10–50 copies of a 36-bp direct repeat, separated by spacer DNA with various sequences, each of which is 37–41 bp. A copy of IS*6110* is inserted within a 36-bp direct repeat in the middle of the DR locus in most strains. *Mycobacterium tuberculosis* strains have the same overall arrangement of spacers but differ in terms of the presence or absence of specific spacers. Spacer oligonucleotide typing (spoligotyping) involves polymerase chain reaction (PCR) amplification of the DR locus, followed by hybridization of the labeled PCR products to a membrane that contains covalently bound oligonucleotides corresponding to each of 43 spacers. Individual strains have positive or negative signals for each spacer. The top section shows the 43 direct repeats (rectangles) and spacers (horizontal lines) used in spoligotyping. The middle section shows the products of PCR amplification of spacers 1 through 6 of *M. bovis* bacilli Calmette–Guérin (BCG), *M. tuberculosis* strain H37Rv, and *M. tuberculosis* hypothetical strain X, with the use of primers (white and black arrowheads) at each end of the DR locus. The bottom section shows the spoligotypes of the three strains (from [10], with permission).

relatively poor discriminatory power for isolates with fewer than five copies of IS*6110*.

1.3.2.2 Spoligotyping (spacer oligonucleotide typing) (See Fig. 1.6) This method is based on polymorphism of the chromosomal DR (direct repeat) locus. The DR elements, identified by Hermans et al. [88], contain multiple, well-conserved 36-bp DRs interspersed with nonrepetitive spacer sequences (34–41 bp long). Strains vary in the number of DRs and in the presence or absence of particular spacers. Indeed, the spacer oligonucleotide typing (spoligotyping) method described by Kamerbeek et al. [94] detects the presence or absence of spacers of known sequence in an isolate in two steps. PCR is used to amplify the spacers between the DRs. The reverse primer used in the PCR is biotin-labeled, so that all reverse strands synthesized are labeled. Individual spacers are then detected by hybridization of the biotin-labeled PCR product to a membrane on which 43 oligonucleotides derived from spacers of *M. bovis* BCG and MTB H37Rv have been covalently linked (see example in Fig. 1.6): 29 oligonucleotides are from spacers common for BCG and H37Rv, six are from spacers specific to *M. bovis*, and eight are from spacers specific to H37Rv.

Contrary to the IS*6110* genotyping method, spoligotyping is a technique based on polymerase chain reaction (PCR). The method is simple, rapid, and robust, and only small amounts of DNA are needed. It can be done on clinical samples or on strains shortly after inoculation into liquid culture [94]. The results can be represented as a binary code (0 corresponding to absence, 1 to presence) and can be expressed in a digital

format [39], which makes it easy to compare the data between laboratories and with data deposited in the international spoligotyping database SpolDB3 housed at the Pasteur Institute in Guadeloupe [60]. This database is available online at http://www.pasteur-guadeloupe.fr/tb/spoldb3/, although its recent version, SpolDB4, is not yet available. This technique is useful not only to identify the species of the MTBC responsible for the infection, but also to characterize the MTB family at an intraspecific level [205]. The disadvantage of this method is its paucity of discrimination, resulting in the need for another method for resolvent genotyping [169,215]. Second-generation spoligotyping that is more resolvent was recently developed and can detect the presence of the 43 traditional spacers, as well as 51 novel spacers [202].

1.3.2.3 MIRU-VNTR (Mycobacterial interspersed repetitive units—variable number tandem repeat) (See Fig. 1.5) Recently, a new technique was elaborated based on specific repetitive elements of *Mycobacterium tuberculosis* (see Fig. 1.5). Indeed, in 1997 and 1998, novel intergenic repetitive units dispersed throughout the mycobacterial chromosome have been identified and called mycobacterial interspersed repetitive units (MIRUs) by Supply et al. and variable number of tandem repeats (VNTRs) by Frothingham et al. [68,182]. These structures are composed of 40–100-bp repetitive sequences organized in direct tandem repeats that are scattered in several locations throughout the chromosome of MTB H37Rv [183]. The total number of MIRUs or VNTRs is estimated to be about 40–50 per genome (41 loci are present in MTB H37Rv [183]). These structures are comparable to minisatellites observed in higher eukaryotes [118]. Some of these MIRU-VNTR loci have been tested and have shown their usefulness for molecular epidemiology studies [68,118]. The sequencing of MIRUs loci identified 12 of them, displaying variations in tandem repeat copy numbers as well as sequence variations between repeats [183]. Mazars et al. [118] published a PCR-based typing method by using these 12 loci for molecular epidemiology studies, and this technique has already shown its potential to discriminate between MTB strains in different studies (see Fig. 1.7).

1.3.3 How Should the Most Appropriate Molecular Marker be Chosen?

Genetic typing is the means by which the microbiologist is able to discriminate and catalogue microbial nucleic acid molecules. As said above, the diversity among nucleic acid molecules provides the basic information for all fields. Nevertheless, currently, full-genome sequences for multiple isolates are rare, implying that genetic typing is still done by methods that are inherently suboptimal [201]. The first fundamental characteristic to define a genetic marker as a good molecular tool is the portability of methodology between laboratories. Van Belkum et al. [201] noted that communication of data can be obstructed because of a general lack of standardized genetic typing procedures and thus, except for primary DNA sequences, typing data frequently suffer from

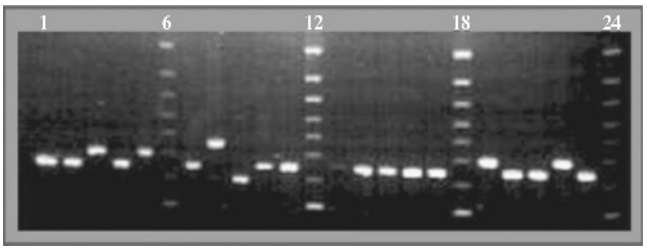

Fig. 1.7. Patterns of MIRU-VNTR of *Mycobacterium tuberculosis*. Lanes 1–5, 7–11, 13–17, and 19–23: Patient's pulmonary isolates from Montpellier Hospital, France. Lanes 6, 12, 18, and 24: Molecular size markers (Fraisse, unpublished data, all rights reserved).

limited interlaboratory reproducibility. Molecular marker standardization is undoubtedly essential for all research fields, whether medical, epidemiological, or genetics. Optimal typability, a high degree of reproducibility, adequate stability, and unprecedented resolving power must characterize the gold standard typing technique. In addition, the procedures should not be too expansive or complex and should be easily accessible. Furthermore, setting up large databases is undoubtedly an advantage for international epidemiological surveillance and for free and easy exchanges between laboratories.

Another important point raised by Van Belkum et al. [201] is that the technique should be chosen with care to provide answers to a specific question. For example, currently, techniques based on nucleic acid polymorphism are more frequently applied and better appreciated than the phenotypic methods in taxonomy, epidemiology, and evolutionary studies. Furthermore, the choice of the optimal molecular markers in accordance with the scope of the study also depends on the space and time scales in which the data were collected or explored. Tibayrenc [192] defined three different time and space scales: (i) days to months, the hospital or village, referred to as short-term epidemiology; (ii) months to years, country- or continent-wide, up to the entire geographical range of the species, referred to as long-term epidemiology; and (iii) millions of years, country- or continent-wide, up to the entire geographical range of the species, such as in phylogenetic studies. Here, Tibayrenc [192] pointed out a central notion, the *speed of evolution* (*molecular clock*) of a given marker, which conditions its power of resolution. Fast markers allow conducting short-term epidemiology studies, while slower markers are more appropriate for long-term epidemiology, and slow markers such as ribosomal RNA genes are more appropriate for phylogenetic studies. Nevertheless, the resolution power of each marker is a function of the organism and the species under study. For MTB, it appears that the gene typing as used by Sreevatsan et al. [177] and Musser et al. [130] could be useful for phylogenetic studies, considering of course polymorphic genes. While the three techniques are based on repetitive chromosomal elements, IS*6110*, spoligotyping, and MIRU-VNTR are better adapted for molecular epidemiology, there is not a single best marker. From the various comparisons of these three markers [86,98,99,115,118], from a resolution power point of view, IS*6110*-based RFLP typing

and MIRU-VNTR appeared very appropriate for short-term epidemiology studies, whereas spoligotyping is more suitable for long-term epidemiology studies. Nevertheless, several studies showed clearly that using multiple methods for molecular epidemiology is necessary [37,132,150,181]. Several authors recommend spoligotyping associated with MIRU-VNTR for molecular epidemiology studies [9,36,40]. This is also developed within the framework of the US national genotyping program to characterize all initial isolates of MTB [40]. One of the limitations is the cost of these techniques, which prevents their routine use in low-income countries.

1.3.4 Population Structure of MTB and Epidemiological Consequences

1.3.4.1 Theoretical and technical assessment

Reproduction is the process by which living creatures transmit their genes to produce another generation of living creatures. This is a common phenomenon to all living organisms but it has a great impact in particular for the population structure of microorganisms, as reproduction strategies are diverse, with a variety of sexual and asexual processes expressed. In bacteria, on a theoretical basis, four different types of population structures have been proposed by Smith et al. [167]; the two extremes being the clonal model on one hand and the sexual model on the other hand. *Clonal or asexual propagation* refers to populations in which the offspring are genetically identical to their parent [195] and thus genetic exchanges are rare or absent (e.g., *Salmonella* [167]). In bacteria, the *sexual model* refers to organisms in which genetic exchanges are very frequent (e.g., *Neisseria gonorrhoeae*; [138]). It is worth noting that in bacteria, the sexual model does not correspond to true sexual reproduction but to frequent exchanges of genetic information (genetic recombination) occurring by the classical bacterial processes such as *transformation*, *conjugation*, *and transduction*. Nevertheless, between these two extremes, Smith et al. [167] described two other intermediate models: cryptic speciation and epidemic clonality. In the case of cryptic speciation, the species under study is subdivided into two or more biological species, each being sexual (e.g., *Rhizobium meliloti*; [167]), but no genetic exchanges occur between the different species. Epidemic clonality is characterized by sudden clonal expansion of a relatively short-lived type occasionally observed for a species that otherwise replicates in a sexual model (e.g., *Neisseria meningitidis*; [167]). Other evolutionary mechanisms, such as migration, selection, and genetic drift, also play a role in the genetic structure of populations, but reproduction is the basic biological process influencing the population structure. Identifying the reproduction system is all the more essential, as it governs the allelic and genetic distribution in natural populations and conditions the stability of genotypes in space and time [193]. Therefore, this has important consequences from an epidemiological and medical point of view (strain typing, pathogenicity, vector specificity, and susceptibility to drugs and vaccines), and hence on the epidemiological and medical relevance of microorganism genotypes. For pathogenic microorganisms, the clarification of population structure provides unique insights into crucial public health issues, such as the appearance and persistence of variants escaping immunity or the emergence of resistance to antibiotics [129,168,176]. Consequently, it appears incontestable that knowledge of the reproductive system is essential to exploiting molecular epidemiological data fully and correctly.

Population genetics is the scientific discipline that studies genetic diversity and its distribution in natural populations and all the biological events influencing the population structure such as the reproduction system. Two kinds of tests are used in order to infer the population structure in samples being investigated. Tibayrenc [193] detailed the theoretical basis of these studies for microorganisms. Briefly, these tests were based on the two main consequences of sexual reproduction: *segregation* of alleles at given loci (reassortment of different alleles at a given locus) and *recombination* of genotypes (reassortment of genotypes at different loci). Segregation tests are related to *Hardy–Weinberg equilibrium* and imperatively require a diploid level of the organism and an identification of alleles. Therefore, these tests are not applicable to bacteria nor to MTB, which has a haploid genome. Recombination tests are related to *linkage disequilibrium* (nonrandom association of genotypes occurring at different independent loci) and contrary to segregation tests, they can be used irrespective of the ploidy level of the organism under study and even without identifying individual alleles and loci [194]. The only requirement for these tests is to use molecular markers that show a sufficient level of polymorphism and make it possible to perform a multilocus analysis (because loci must be independent) (see [193] for details). The MIRU-VNTR technique compared to IS*6110* and spoligotyping techniques shows the necessary properties to be used for population genetics studies: it is a multilocus marker and the loci are distributed independently along the bacterial genome. In contrast, IS*6110*-based RFLP cannot be used to analyze linkage disequilibrium, as they do not reveal the variability of independent genetic loci. Furthermore, spoligotyping cannot be assumed to be independent from IS*6110*-based RFLP, as this locus is a hot spot for IS*6110* insertions, and changes within this region are often caused by IS*6110*-associated events [57,62,81,88,106]. Furthermore, the DNA sequences of multiple *housekeeping genes* can also be used to infer the population structure and the phylogenetic history of bacterial species. Nevertheless, as polymorphic genes should be selected to conduct these studies, it is worth noting that the choice is limited in MTB because there is an extremely limited amount of unselected nucleotide sequence variation in structural genes and housekeeping genes in this bacteria [63,130,177].

Another discipline, molecular phylogenetics, also appears vital for understanding evolutionary molecular biology and molecular epidemiology. This discipline is devoted to understanding the hierarchical structure of biological diversity through genetic data. One important outgrowth of the phylogenetic revolution is the recognition that phylogenetic trees

provide an important and appropriate context to address questions in a variety of disciplines such as molecular epidemiology and evolutionary biology. More and more, the phylogenetic approach is used to explore the population structure and to infer the system of reproduction of various organisms [58,83]. These analyses contribute complementary information beyond population genetics studies, such as genetic structuring in a population, identification of a genetically individualized entity, for example, cryptic species and epidemics. Furthermore, the congruence or incongruence of different gene phylogenies also provides substantial insight into the population structure. Indeed, congruence of several independent genes is evidence of a lack of genetic exchange, whereas phylogeny incongruence reflects frequent genetic exchange.

1.3.4.2 Population structure of MTB Mycobacteria, like other bacteria, may have the potential to exchange DNA. Indeed, experimental transduction has been performed in MTB [85], and natural conjugation has been demonstrated for *M. smegmatis* [143]. Nontuberculous mycobacteria can acquire antibiotic resistance genes from other species [80,140]. Furthermore, simultaneous infection of patients by two different strains was evidenced in high-incidence areas [16,27,218]. All these data suggest that MTB could be able to exchange DNA in natural populations. Nevertheless, authors have hypothesized for several years that this species has a clonal population structure. This statement was based on the preponderance of certain genotypes and on the low level of genetic polymorphism and not on a rigorous population genetics analysis. Contrary to a widespread idea, the restricted gene sequence diversity and empirical observation of some predominant genotypes in various epidemiological studies provide no indication of its population structure, as they are compatible with distinct population structures with variable levels of recombination [58,59,167,176]. Furthermore, until the year 2000, no marker presented the necessary properties (i.e., a polymorphic marker based on several independent loci) to conduct population genetic studies (see above). The development of the MIRU-VNTR technique and the sequencing of several MTB strains finally provided appropriate methodologies for studying population structure and thus MTB's mode of reproduction. Consequently, few studies based on these markers supported the conclusion that *MTB is a clonal organism, with no evidence of lateral gene transfer* [7,184,187]. Two studies conducted in South African and Moroccan populations tested linkage disequilibrium by means of MIRU-VNTR techniques [6,175]. A third study was mainly based on a phylogenetic analysis of polymorphic gene sequences of a sample of 316 UK clinical isolates [178]. Despite the strong linkage disequilibrium observed in these populations and consequently the relevant identification of the typically clonal evolutionary model, *the occurrence and significance of genetic exchanges within natural populations of this species remain to be demonstrated*.

Recently, a study based on phylogenetic and sequence analysis was published by Gutierrez et al. [83] in order to unravel the evolutionary success of MTB. Members of the MTBC suggested representing the clonal progeny of a single successful ancestor, resulting from a recent evolutionary bottleneck that occurred 20,000–35,000 years ago [177]. Gutierrez et al. [83] identified the progenitor of MTBC, which includes *M. canettii* (already suggested by Brosch et al. [20]), a rare tubercle bacillus with an unusual smooth colony phenotype [199], and other smooth tubercle bacilli from Djibouti. These authors proposed to call this group of strains *M. prototuberculosis* species. From a population structure point of view, the interesting element in this paper is the observation of a mosaic structure of some genes and an incongruence of gene phylogenies. Both results suggested that DNA recombination is frequent in this population [83]. In contrast, using the same analysis, they detected no evidence of recombination among the MTBC strains, consistent with the previously reported clonal population structure. Furthermore, results supported that despite its present clonal and highly conserved structure, MTBC is actually a composite assembly of genetic sequences resulting from multiple remote horizontal gene transfer events. Therefore, the authors proposed several potential explanations for the apparent absence of recombination among the MTBC strains after the bottleneck [83]: (i) the MTBC strains could have lost the capacity of horizontal gene transfer, (ii) horizontal gene transfer events are too rare among tubercle bacilli to have occurred since the MTBC bottleneck, and (iii) the MTBC ecological niche differs from that of *M. prototuberculosis* and offers no opportunity for recombination events. Thus, further progress in the understanding of evolutionary biology of MTBC and MTB still requires deciphering why MTB is no longer able to exchange genetic information in natural populations and whether the different species, families, and populations belonging to MTBC present all the same population structure.

In summary, because of the strong linkage disequilibrium and the phylogenetic studies developed in several populations, it appears that MTB follows a typical clonal model. This implies that MTB genotypes can be considered as epidemiologically discrete units of research, which Tibayrenc [192] calls discrete typing units (DTUs), and thus can be used as markers for applied studies (epidemiological tracking, vaccine and drug design, clinical studies). From these clonal characteristics, these DTUs or MTB clones can be specifically identified by appropriate genetic markers or "tags" [192]. Nevertheless, the description of a MTB progenitor and of the high frequency of genetic exchanges in this ancestral lineage does not allow excluding the possibility of genetic exchanges in MTB.

1.4 USE OF MOLECULAR EPIDEMIOLOGY FOR UNDERSTANDING TUBERCULOSIS TRANSMISSION AND PATHOGENESIS

The primary goals of TB control at the community and individual levels are to identify the bacteria responsible for infection and to treat infected people. Nevertheless, it is essential to control and fight the disease by tracking the strains identified

as the source of infection and thus discriminating strains. From this crucial need *molecular epidemiology* was born, which has become a major field of research in MTB in the last 20 years. This scientific domain corresponds to the interpretation of molecular data through the conventional epidemiologic studies. Thus, this domain involves several disciplines, encompassing medicine, molecular biology, epidemiology, and biostatistics. Molecular epidemiology is now largely recognized as a science that makes it possible to understand the transmission, pathogenesis, and etiology of human disease [40,188]. This discipline provides tools for clinicians, microbiologists, and epidemiologists for investigating infections. Indeed, molecular epidemiology is essential to studying the spread of MTB in *epidemics* and *outbreaks*, to analyzing the transmission dynamics, and to determining the risk factors for TB transmission in a community. It plays a great role in distinguishing between *exogenous infection* or *reinfection* and *endogenous reactivation*. In the laboratory, it can also be used to identify cross-contamination. In addition, molecular tools have provided markers able to identify specific gene mutations corresponding to various drug resistances [121,147,189]. Genotyping determines whether the development of drug resistance in a TB patient during treatment is caused by the same strain or another strain by exogenous reinfection. In regard to virulence and pathogenesis studies, molecular epidemiology has already proven to be relevant to attaining insight into the strain's capacity to be pathogenic or drug resistant. Finally, this discipline becomes fundamental to developing strategies for treatment and prevention of diseases. Therefore, it is worth noting that for interpretation of molecular epidemiology results, it is important to consider not only all the clinical, biological, and epidemiological data recorded from tuberculous patients (requiring interview and biological analysis) but also phenotypical, biological, and epidemiological data concerning MTB isolates (requiring an antibiogram, genotyping, and culture).

Within this framework, this section is a review of various epidemiological issues for which molecular epidemiology can improve the understanding of MTB transmission and pathogenesis.

1.4.1 MTB Families and Worldwide Distribution

The world has entered an era of "*diseases without borders*," with 1 million people crossing borders daily, too often carrying with them diseases that were once geographically isolated. By virtue of its worldwide distribution, TB, like HIV, is classified in this category. Lazcano-Ponce et al. [104], assert that the *framework is essential for collaboration on alerting the world to epidemics and responding to public health emergencies*. This is necessary to guarantee a high level of security against the dissemination of communicable diseases in an ever more globalized world. Thus, global molecular epidemiology studies of MTB appear as fundamental as local ones in order to develop strategy to control and fight TB.

As described in the previous section of this chapter, genotyping allows tracking of MTB strains at local as well as global levels. Genetic data allow identifying and following the spread of a particular genotype worldwide. For greater convenience, MTB species have been subdivided into *families*, also called clades in the literature, corresponding to specific genotypes or *clusters* (a cluster corresponds to a particular genotype shared by two or more MTB isolates) or groups of genotypes (corresponding to the DTUs described by Tibayrenc [192]; see above). These families or clades appeared from the mid-1990s with the worldwide development and technological progresses of molecular epidemiology studies. The major families, or those that have been studied more thoroughly, bear a specific name. As an example, we can describe in detail the case of the best-known family, *the Beijing family*, which was first described by Van Soolingen et al. in 1993 [208]. These researchers identified this family by analyzing the population structure of MTB strains from the Republic of China. The vast majority of strains under study belonged to a genetically closely related group that originated from the province of Beijing; therefore, they designated this group the Beijing family. They observed that strains of this family were also found to dominate in neighboring countries such as Mongolia, South Korea, and Thailand, whereas a low prevalence of such strains was observed in countries on other continents. From these data, they suggested that strains of the Beijing family recently expanded from a single ancestor that had a selective advantage.

To date, the most recent global study has been conducted by Filliol et al. [61] on a data set of 13,008 isolates from more than 90 countries. This study, based on the spoligotyping technique, updated the data published by Soini et al. and Sola et al. [170–172]. All the results were integrated into the SpolDB3.0 database. They identified 813 different spoligotypes shared by 2 or more isolates, which contained 11,708 isolates, whereas 1300 spoligotypes were orphans. They evidenced seven major MTB families, the Beijing family, the EAI family (East African-Indian), the CAS family (Central Asian), the T group of families, the Haarlem family, the X family, and the LAM family (Latin American and Mediterranean). The Beijing type was predominant (see Fig. 1.8), followed by the Haarlem type, then by the X types, which are highly prevalent in the United Kingdom and the United States. Nevertheless, Filliol et al. [61] underlined major differences in MTB populations between the subcontinents under study. The global observation was able to define that most MTBs are confined to specific geographic locations [40,56,61]. Nevertheless, these worldwide studies and the numerous molecular studies already published showed that some families are widely dispersed both geographically and temporally, suggesting that they are more transmissible, or more pathogenic than other strains [40,61]. Daley [40] described the Beijing family as detected in high proportions among the strains in several countries (Fig. 1.8) and as associated with large outbreaks, febrile response, treatment failure, relapse, and drug resistance. But to date, it is not clear why the Beijing family strains are so widely disseminated [12]. Daley [40] suggested different hypotheses such as a selective advantage of

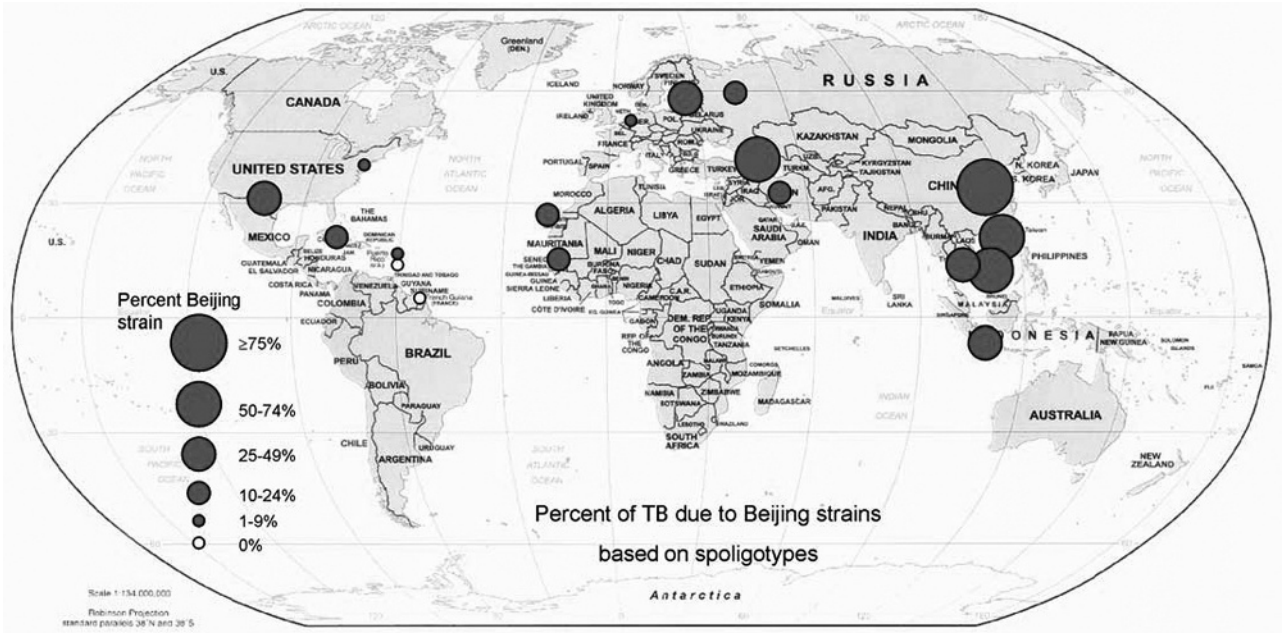

Fig. 1.8. Percentage of tuberculosis due to Beijing strains. Data from studies based on spoligotyping (from Glynn et al. [77]).

these ubiquitous families, a better ability to establish infection, a more rapid progression from infection to disease [10,112], and a longer time to spread. Further research is needed in order to determine why some families are more widespread than others.

1.4.2 MTB in Developing Versus Developed Countries

As we have seen above, *tuberculosis has disseminated globally, but it is not distributed equally throughout the world, with developing countries having by far the highest burden*. Indeed, more than 90% of TB cases occur in developing countries (see Fig. 1.1). The areas most hard hit by this disease are Africa, Southeast Asia, and Eastern Europe. Sub-Saharan Africa has the highest incidence (290 per 100,000 population) with more than 1.5 million cases of TB. The most populous countries of Asia have the highest numbers of TB cases: India, China, Bangladesh, and Pakistan together account for more than half of the global burden. Case numbers have declined more or less steadily in Western and Central Europe, in North and South America, and in the Middle East. These data evidence once again the inequality between developing and industrialized countries in our modern society. As Lazcano-Ponce et al. [104] explained in their paper, investment and investigation in health also involve inequalities at the global level, and this includes insufficient north–south transfer of funds, technology, and expertise in the health field, including the specific area of communicable diseases. Furthermore, although lower-resource countries have by far the highest burden of TB, we can regret that molecular epidemiology studies have not yet been conducted in many of these countries.

Nevertheless, global studies showed that TB transmission varies greatly depending not only on the country but also on the country's development level. The distribution of each MTB family and the number of orphans change geographically (see Filliol et al. [61] for details), for example, the number of orphan types (or singletons) ranged from a low of 8% (North America) to a high of 21% (Middle East and Central Asia); the Beijing family ranged from 2% in South America to 3–5% in Central America, Europe, Africa, and the Middle East and Central Asia, 13% in Oceania, 16% in North America, and as high as 45% in East Asia. Daley [40] noted that considering the ability of the Beijing family to become multidrug resistant, its high prevalence in certain regions of the world is an important issue for effective TB control.

Besides the MTB family distribution, genetic diversity is also an important indicator of TB endemicity and transmission, as well as the efficiency of TB control [10]. MTB's genetic diversity differed greatly in developing versus developed countries [87]. Indeed, low genetic diversity within MTB populations is typical of a high TB incidence or of an epidemic pattern and suggests inadequate TB control. This situation is encountered more particularly in developing countries with a high TB incidence. Some studies evidenced the slight genetic polymorphism in different regions such as Honduras, Ethiopia, Tunisia, different countries of West Asia, and the Southern Africa region [87,145,184,209]. A low genetic polymorphism was also observed in the case of localized epidemics characterized by a rapid spread of particular strains in specific areas (called hotspots). These phenomena can be observed in developing countries as well as in developed countries. A number of examples can be cited: the

particular case of the spread of multidrug-resistant isolates belonging to the W-Beijing family in a Russian prison [197] and the multiple occurrence of MTB epidemics in New York City (see review by Paolo and Nosanchuk [141]).

In contrast, in regions with low TB incidence, mostly in developed countries, genetic diversity is higher, as in Denmark and the Netherlands where most isolates show unique DNA fingerprint patterns [87,216]. This is also true when considering a more restricted geographic area such as cities. Indeed, a retrospective study (2002, 2003) concerning isolates from Montpellier, France, showed, by means of MIRU-VNTR and spoligotype techniques, a high level of genetic polymorphism and weak clustering (Fraisse et al. unpublished data). These results suggest efficient control of TB, which prevents the spread of MTB strains in populations and a higher rate of reactivation compared to recent transmission.

However, unexpected results were obtained for several countries with high TB incidence. For example, studies conducted on samples isolated in Morocco and Burkina Faso showed higher values of genetic diversity (Tazi et al. [187]; Godreuil et al., unpublished data). Further studies are necessary in order to understand the epidemiological significance of these data.

1.4.3 Clinical and Epidemiological Relevance of Molecular Epidemiology at the Local Level

Routine public health investigations do not allow deciphering the chain of transmission, and the source of infection and the characteristics of strains responsible for infection are unknown. *Molecular epidemiology by genotyping with resolvent markers can fill this gap concerning the chain of TB transmission*. Indeed, isolates from patients who were infected by a common source or belonged to the same chain of transmission have identical or closely related genotypes (considering genotyping with the most resolvent markers); in other words, clustering is assumed to reflect recent transmission within a population [150,164]. In contrast, MTB isolates from patients with epidemiologically unrelated TB present a broad variability of genotypes. For example, it has been estimated on the basis of clustering of DNA fingerprint patterns, that half of TB cases in a South African mine hospital were caused by ongoing transmission [78].

At the local level, it is therefore important to link molecular epidemiology and classical epidemiological tools in order to identify contacts of patients outside the home and workplace and in the locations where they spend time. The patient's environment could thus be screened for TB infection and disease and contacts treated if necessary. For example, Torrea et al. [196] identified several chains of transmission (familial or geographical cases) in French Polynesia using a detailed molecular study. Nevertheless, Daley [40] explained on the basis of several studies [164,203] that a relatively small proportion of TB cases presenting identical genotypes were named as a contact by the source case [41]. This may be attributable to unsuspected transmission not easily detected by conventional contact tracing investigations. Indeed, transmission can occur through only short and casual contact, difficult to pinpoint [207].

Furthermore, *molecular fingerprinting can be used at a local level to establish or rule out the existence of an emerging outbreak*. The investigation of outbreaks remains central to the control of TB. For example, Diel et al. [48] described an ongoing outbreak in the Federal State of Hamburg, Germany, by a molecular epidemiology study. They identified various infectious chains of contact that, starting in a bar that played the role of a turntable, moved out rapidly into several areas such as housing for homeless men and alcoholics and a tank-cleaning firm. This study can be considered as a model because it combines detailed clinical and epidemiological data and phenotypical and molecular studies. Nowadays, numerous outbreaks continue to be identified in various public areas such as hospitals, schools, bars, prisons, nursing homes, and homeless shelters in developed and developing countries [38,52,90,92,105,116,120,153,159,178]. These studies were also able to identify risk factors for TB transmission. It is worth noting that *MDR and HIV are often incriminated in the emergence of outbreaks* [67,100]. It has been observed that MDR strains are less responsive to standard therapy, and patients remain infectious for longer periods of time. Breathnach et al. [18] noted that the outbreaks linked to HIV are globally described in hospitals where AIDS patients are cared for together, and increasingly involve MDR strains (see below for more details on HIV and MDR linked to TB).

At a nosocomial level, in addition to detection of outbreaks, genotyping of isolates from patients is also useful for identifying cross-contamination and mixed infection, as well as for differentiating reactivation from reinfection. Barnes et al. [10] evaluated that 3% of patients whose cultures are positive for MTB in clinical laboratories do not have TB. Cross-contamination is suspected because these patients present with negative acid-fast smears and clinical findings. Comparing the isolate genotype with those circulating in the laboratory makes it possible to identify cross-contamination and thus stop unnecessary anti-tuberculous medication. Genotyping can also evidence cases of multiple infections. The occurrence of mixed infection is now widely accepted, whereas until recently it has been assumed that patients could be infected only with a single MTB strain, and infection with one strain is thought to confer immunity to MTB superinfections [161]. Several molecular investigations showed either simultaneous infection with multiple MTB strains [16,161] or multiple infections caused by an exogenous reinfection [27,161]. Furthermore, these mixed-strain infections can involve drug-sensitive and MDR strains [8,190].

Molecular fingerprinting appears to be useful for differentiating (*i*) *a reactivation of latent infection from a recent infection and* (*ii*) *a relapse with the previous MTB strain from an exogenous reinfection by a new strain*. As described in Section 1.2 of this chapter, the first episode of an active case of TB can be caused by either a recent transmission of MTB strains or a reactivation of latent infection. Isolates that have the same molecular fingerprint are presumed to be part of a cluster of recent transmission, with one or more people in the cluster having transmitted infection

to the others [75,212]. For example, a recent report described a recent infection of MTB in northern Malawi, in which 72% of the strains were clustered [76]. Reactivation of latent infection occurs in about 10% of infected individuals, leading to active and contagious tuberculosis [89,113,114]. It has been demonstrated that reactivation of latent infection contributes substantially to the incidence of adult TB, especially in more developed countries where disease prevalence is fairly low. As examples, (i) Geng et al. [71] obtained data that suggested that in the United States among foreign-born people, TB is largely caused by reactivation of latent infection, whereas among US-born individuals, many cases result from recent transmission and (ii) Lillebaeck et al. [110] presented molecular evidence of *reactivation of MTB 33 years after primary infection*.

Concerning the cases of a second TB episode, in their review Chiang and Riley [28] detailed the debate that has existed for decades concerning reactivation and reinfection because reinfection was considered to be an uncommon cause of TB. Classically, before genotyping development, relapses were associated with reactivation of MTB infection. *Several molecular studies showed clearly that reinfection causes a significant proportion of recurrent TB episodes* [24,45,206]. This proportion seems to vary as a function of the area, the endemicity, and the biological status of patients [45]. Nevertheless, Lambert et al. [102] reported that the importance of reinfection remains unclear because only very few studies are adequately designed for that particular research objective and/or report a sufficient number of observations. They consider that only the study published by Sonnenberg et al. [174] provides an exact estimate of the incidence of recurrence due to reinfection, indicating its importance in HIV-infected patients in an environment with an unusually high TB incidence.

1.4.4 Use of Genotyping to Study the Impact of HIV/AIDS and Drug Resistance on Pathogenesis and Transmission

As described above, TB incidence is linked to poverty and poor living conditions, in some cases to civil conflicts and wars, to deteriorating health services, and to lack of drug availability. Besides these social factors, two major problems regarding the control of TB are emerging: coinfection with HIV and resistance of MTB to the currently used regimen of tuberculostatics. We can approach the problem of HIV and drug resistance together, as it has been demonstrated that they are strongly associated [11,51,55,125,127,165]. Agerton et al. [1] also described outbreaks of MDR-TB involving hundreds of cases, many of whom were infected with HIV, with high mortality rates [13,33,67,93,200].

DNA fingerprinting can determine that in many of these outbreaks, the susceptibility of HIV-positive patients to tuberculosis infections and the accelerated breakdown to disease often result in more rapid transmission of the infection [42,50,55,149,153,165]. Furthermore, the study of MTB isolates obtained from AIDS patients by fingerprinting showed that reinfection and relapse both occur in HIV-infected patients, as the susceptibility to superinfections will most likely be related to the immune status of the patient. By combining classic and molecular epidemiology, Sonnenberg et al. [174] showed that HIV-1 infection is a risk factor for recurrence, as HIV-1 is strongly associated with disease caused by reinfection but not with relapse.

From an evolutionary point of view, since the development of detailed fingerprinting of MTB strains, the genetic divergence of strains circulating in HIV-positive and HIV-negative patients has been debated. As stated by Ahmed and Hasnain [3], it has been speculated that HIV/AIDS patients constitute an ecological niche for MTB, where less virulent strains multiply freely without the selection pressure provided by an optimal immune response. One study has been conducted where significantly different genotypes were observed for HIV-associated tubercle bacilli as compared to bacilli recovered from HIV-uninfected patients [2]. In contrast, Yang et al. [217] obtained results that suggested an equal risk of infection with a defined MTB clone for HIV-seropositive and HIV-seronegative individuals. These latter results were also confirmed by a recent study on a MTB population from Burkina-Faso (Godreuil et al., unpublished data).

With drug-resistant strains, genotyping determines whether the treatment failure and the development of drug resistance are caused by the same strain (inadequate treatment) or a new strain (reinfection during treatment). Several studies have reported that the development of drug resistance may be associated with either the same or different strains [78,79,165,206,211]. The relative contribution of these different mechanisms to treatment failure and/or the development of drug resistance seems to vary according to the populations studied. Nevertheless, Sonnenberg et al. [174] demonstrated that even in a setting with high rates of TB transmission and HIV-1 infection, the dominant mechanism of drug resistance while on treatment was acquisition rather than transmission. Thus, despite reinfection being a possible mechanism of treatment failure and the development of drug resistance, it appears uncommon in comparison with the number of patients who had acquired drug resistance with the same strain.

Furthermore, molecular studies identify whether drug-resistant strains are significant risk factors for secondary cases and thus for outbreaks. Daley [40] found that several molecular epidemiological studies have reported that patients who have drug-resistant strains were less likely to cluster, suggesting that drug-resistant strains might be less prone to being transmitted or to causing active disease [69,78,210]. Burgos and Pym [21] have also recently reported that isoniazid-resistant strains confer a significantly lower number of secondary cases than drug-susceptible strains. Daley [40] concluded that these findings support the hypothesis that drug-resistant strains are less likely to cause disease than drug-susceptible strains. Nevertheless, different environmental or biological conditions counterbalance this weak potential for being transmitted and to causing active disease. First, Post et al. [146] estimated that 8–35% of patients with MDR-TB have persistently active disease that is refractory to a multidrug

regimen [156,160,180,185] and thus are a constant source of transmission of MDR-TB [67,151,160,206]. Second, Daley [40] reported that there are some populations in which drug resistance is neither detected nor treated effectively and where the longer-duration regimens might offset the bacterium's diminished capacity to cause secondary cases [21]. Third, Daley [40] explained that in areas that have high prevalence rates of HIV, the increased host susceptibility, even to strains that have diminished virulence, may also offset bacterial difference [128]. Fourth, a review of the literature concerning drug-resistant TB and especially MDR-TB showed that only a few clones are mainly responsible of MDR-TB outbreaks and thus would have a higher virulence and a higher capacity to be transmitted and to cause disease. The most frequently cited example is the Beijing/W type, which is described worldwide and is involved in numerous MDR-TB outbreaks [74,103,126,128,206].

From an experimental point of view, it has also been demonstrated by several authors that drug-resistant strains are less virulent and present a decrease in pathogenicity in comparison to drug-sensitive strains [29,122,152]. Indeed, Meacci et al. [121] exposed that drug-resistant bacteria are believed to grow more slowly than susceptible bacteria, as mutations conferring resistance reduce their overall fitness, a phenomenon known as cost of resistance [107]. Nevertheless, as described above, the emergence of MDR MTB strains is alarming and is a worldwide health care problem, thus contradicting most experimental data. Several authors, however, have demonstrated in other bacteria that fit variants are quickly selected in a drug-resistant bacterial population, in which compensatory mutations eliminate the biological cost of resistance [4,15,108,175]. Meacci et al. [121] demonstrated by following up a tuberculous patient with active disease for more than 12 years that phenotypic and genotypic changes occurred in the drug resistance of MTB isolates. First, molecular typing showed a single parental strain that infected the patient and persisted throughout the disease. Second, molecular analysis of the drug-resistance-related genes revealed that discrete subpopulations evolved over time from the parental strain by acquiring and accumulating resistance-conferring mutations to isoniazid, rifampin, and streptomycin. Overall, authors noted that during a chronic infection, several subpopulations may coexist in the same patient with different drug susceptibility profiles [121]. This was also observed by Post et al. [146] in a population of 13 HIV-negative patients with MDR-TB that was refractory to chemotherapy given for 12 months. Meacci et al. [121] described the emergence of a successful MDR-TB strain during the genetic and phenotypic changes, resulting from progressive accumulation of genetic alterations, possibly conferring a selective advantage for bacterial survival. Low compliance with therapy may have elicited the selection of resistant strains, which also persisted after stopping treatment. These evolutionary changes could partly explain the numerous outbreaks of peculiar drug-resistant strains recovering an increased potential for being transmitted and causing disease.

1.5 URGENT NEEDS FOR TB CONTROL, LIMITATIONS, AND NEW ISSUES FOR MOLECULAR EPIDEMIOLOGY

This section aims to define the urgent needs for improving TB control and to detail the limitations of modern molecular epidemiology studies. Indeed, molecular epidemiology approaches still present drawbacks that need to be resolved in order to advance the knowledge on TB transmission and enable better public health control strategies. This section will also include the description of molecular technologies that promise to improve molecular epidemiology studies. All the molecular methods described here are not particularly recent, but they are not used routinely for MTB and seem promising for MTB molecular epidemiology.

1.5.1 Urgent Needs for TB Control and Molecular Epidemiology

The development of DNA fingerprinting and molecular epidemiology has pushed forward our understanding of MTB transmission dynamics. Nevertheless, *the TB problem is far from being solved, especially in developing countries*. There are *urgent needs* for control of the disease and thus it is essential to progress in *applied research*. Indeed, new vaccines [49,131], new drugs [23,134], and new diagnostics and advances in TB management [19,136] are urgently needed. Furthermore, we believe that it is no longer necessary to justify that *basic research*, including evolutionary and population genetics, experimental evolution, immunology, and cellular biology, is indispensable in order to progress in applied research. As we described above, molecular epidemiology is a scientific domain that can make the connection between applied and basic research. As demonstrated in this chapter, *molecular epidemiology studies may be useful in public health control and in management of clinical situations*. Nevertheless, today genotyping is exploited only in a few TB control programs and is usually done within the framework of retrospective studies. Very few studies have been conducted prospectively, and these mainly in developed countries [14,26,46,47,78,111,204]. Furthermore, it is essential first to extend these studies worldwide, particularly in developing countries. This requires, of course, financial and governmental support, an efficient and disinterested worldwide commitment, and technological improvement to develop molecular tools that are usable in developed as well as in developing countries. Second, the rapid exploitation of molecular data, *in real time*, is essential in order to control TB at global and local levels.

At the local level, it is crucial to rapidly and efficiently identify the source of contamination, the cases of cross-contamination, and the drug sensitivity of strains in order to select the best-adapted treatment and to rapidly propose prevention and treatment to patients' relatives when needed.

At a global level, the rapid international communication and global infectious disease surveillance and management are fundamental in order to identify an international outbreak

and prevent pandemia. This is relevant today not only for TB control but also for all the emerging infectious diseases such as avian flu in 2005, or Asian severe acute respiratory syndrome (SARS) in 2003, or mad cow disease in the 1990s. Constant concerted efforts at an international level spanning several decades can help solve the global TB problem, HIV/MTB co-infection, and drug resistance in both developed and developing countries. Global commitment and engagement of all interest groups will also be necessary to achieve this goal. To develop an efficient program of TB surveillance, interactive national and international databases are also required recovering all the epidemiological, molecular, and biological data of each isolate. As described in Section 1.3 of this chapter, a few databases already exist, but they require greater development (either they are not interactive or they are not updated regularly or they are too restricted). These databases should be accessible online and interactive in order to provide access to all clinicians and researchers so they can compare their data. They could be constructed on the model of nucleic acid or protein databases such as GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) or EBI (http://www.ebi.ac.uk/), which allow researchers to submit, consult, and analyze sequence data. To succeed, several conditions are needed: (i) the definition of standardized methodologies for data exploitation, (ii) the development of a completely disinterested structure that is updated frequently and that will avoid publication pressure and scientific competition, and (iii) the validation of data in order to prevent errors. This type of database may be a powerful public health tool to follow the evolution of TB from a drug resistance or epidemiological point of view at an international level. Several publications have shown the usefulness of databases: Drobniewski et al. [52], Filliol et al. [60], Zozio et al. [219], and Niobe-Eyangoh et al. [133].

1.5.2 Limitations of Modern Molecular Tools

Even today, three main caveats restrict the routine use of DNA fingerprinting: the cost, the complexity of the techniques, and the length of time needed to obtain results. Indeed, the high cost of most of these molecular techniques and the sophisticated equipment and skilled personnel required have precluded their implementation on a routine basis, especially in low-income countries. IS*6110*-based RFLP, spoligotyping, and automated MIRU-VNTR require sophisticated material and specifically trained personnel. Furthermore, globally the time between sputum collection and data interpretation is too long: (i) IS*6110*, spoligotyping, and classical drug resistance identification require mycobacteria culture lasting several weeks (see Section 1.2). In addition, only a limited number of strains can be rapidly identified at the same time for IS*6110*-based RFLP, spoligotyping, and MIRU-VNTR on agarose gel. However, MIRU-VNTR appears as the most appropriate technique to develop standardized data in a short time period with a larger number of samples. Indeed, automated MIRU-VNTR can analyze a high number of samples a day, which makes this technique promising in molecular epidemiology studies in real time, as it does not require cell culture (Supply, personal communication). Nevertheless, as described in Section 1.3 of this chapter, automated MIRU-VNTR requires sophisticated equipment and skilled personnel, and the cost is still high for low-resource countries. On the contrary, MIRU-VNTR on agarose gel is an easy technique with a lower cost, but only a few samples per day can be studied. Thus, at present, there are still no perfect molecular tools. In Chapter 41, Kathleen Victoir emphasizes that "*the creative scientific challenge at present is to develop the best possible tools adapted to resource-poor settings*." The perfect markers should be cheap, rapid, easy to use, and exportable between laboratories.

1.5.3 Promising New Technologies

The availability of whole genome sequences has aided the development of new genomic technologies such as *microarrays* or *genechips* (Fig 1.9). Today, this advanced technology is reserved for researchers in leading laboratories, but these DNA chips may soon invade hospitals and hopefully medical dispensaries. DNA microarrays are small, solid supports, typically glass, filter, or silicon wafer, upon which DNA molecules of known sequences are deposited or synthesized in a predetermined spatial order so that they can be made available as probes in a high-throughput, parallel manner. They can consist of a few hundred to hundreds of thousands of sets. There are three major applications for the DNA microarray technology: identifying the sequence (gene/gene mutation), determining the expression level (abundance) of the genes of one sample, or comparing gene transcription in two or more different cell types. Butcher [22] described the usefulness of microarrays for MTB research and their contribution for enhancing a TB control program. This review showed the broad application of microarrays in understanding MTB physiology, host–pathogen interactions, mechanisms of drug action, in vitro and in vivo gene expression, host responses, comparative genomics, and functional genomics of particular genes. As they can also help identify individuals with similar biological patterns, microarray analysis can assist drug companies in choosing the most appropriate candidates for clinical trials of new drugs. In the future, this emerging technology has the potential to help health care professionals select the most effective drugs, or those with the fewest side effects, for each patient. Butcher [22] stated that microarrays are one of the new functional genomics technologies exploiting genome sequence information that will bring us closer to reaching the scientific and moral imperatives of better vaccines, diagnostics, and new drugs for the control of TB throughout the world. They could help at all steps of the patient's follow-up: disease and strain identification, treatment selection, and observation of therapy efficacy. For the moment, cost is a limiting factor, but the objective of the specialists in biotechnology is to reduce the production cost in order to make this advanced technology routinely accessible.

Another method based on spoligotyping was recently developed by Cowan et al. [35]. The authors transferred

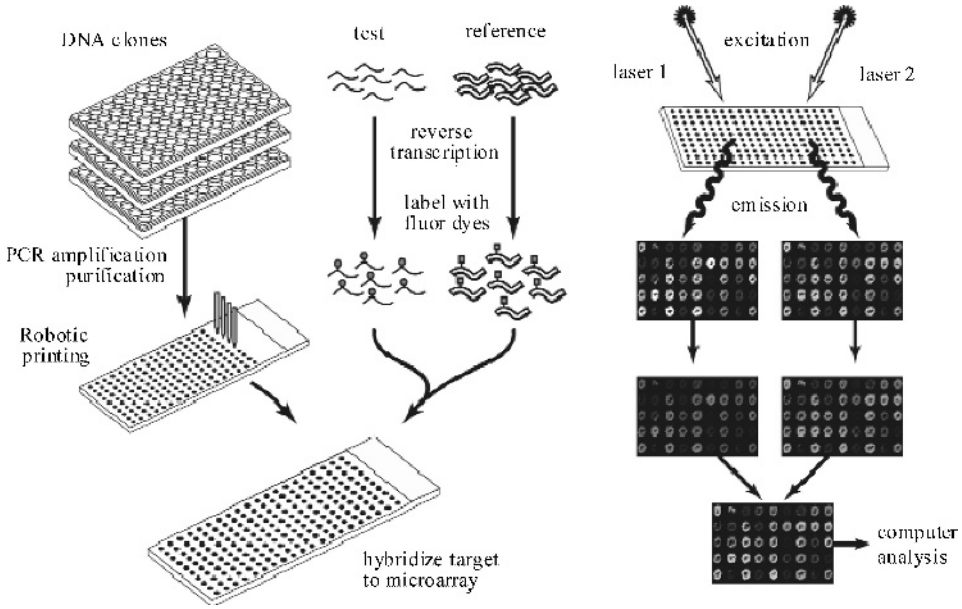

Fig. 1.9. cDNA microarray schema. Templates for genes of interest are obtained and amplified by PCR. Following purification and quality control, aliquots (~5 nl) are printed on coated glass microscope slides using a computer-controlled, high-speed robot. Total RNA from both the test and reference sample is fluorescently labeled using a single round of reverse transcription. The fluorescent targets are pooled and allowed to hybridize under stringent conditions to the clones on the array. Laser excitation of the incorporated targets yields an emission with a characteristic spectra, which is measured using a scanning confocal laser microscope. Monochrome images from the scanner are imported into software in which the images are pseudo-colored and merged. Information on the clones, including gene name, clone identifier, intensity values, intensity ratios, normalization constant, and confidence intervals is attached to each target (from [53], with permission).

spoligotyping from a reverse line-blot hybridization, membrane-based assay to a luminex multianalyte profiling system. This technique may offer many benefits such as a decrease in the turnaround time and the labor involved, a decrease in technical complexity, and greater flexibility (1–96 isolates can be used without increasing the labor time or cost per isolate, and reproducibility is increased) [35]. The authors demonstrate that the luminex system is an attractive alternative for laboratories that perform spoligotyping on a high-throughput scale or for those that frequently require a rapid turnaround time for only a few isolates per run [35]. Nevertheless, as a classical spoligotyping method, another technique with a greater discriminatory power would have to be used for obtaining a maximum of resolution.

Concerning genotypic susceptibility testing, the elucidation of the molecular mechanisms responsible for the action of various anti-tuberculous drugs facilitated the development of rapid methods for susceptibility testing. Jalava and Marttila [91] introduced genetic methods and new techniques useful for both resistance genetic studies and rapid molecular diagnostics of resistance for several bacteria including MTB. They described six different molecular techniques, from which two techniques held their attention: PCR single-strand conformation polymorphism (SSCP) and high-density oligonucleotide arrays. SSCP is a rapid screening method for base-pair alterations in PCR-amplified DNA. This method appears cost-effective and presents a short turnaround time, which makes it suitable for use in clinical laboratories. Jalava and Marttila [91] described that high-density oligonucleotide arrays may also offer a powerful solution to genotypic detection of drug-resistant MTB isolates. So far, these microarrays have mainly been used for the detection of rifampin resistance [72,123,198] with promising results. Consequently, Jalava and Marttila [91] argue that the DNA microarray strategy could be expanded to include parallel testing of various genes mediating drug resistance in MTB. Furthermore, an array for the simultaneous testing of isoniazid, rifampin, streptomycin, and fluoroquinolone susceptibilities has already been designed by Gingeras et al. [72] and could be integrated into a TB control program for the rapid diagnosis of drug-resistant TB. Jalava and Marttila [91] also defined the necessary requirements for assessing the suitability of molecular methods for anti-tuberculous susceptibility testing. First, the technique should have a high sensitivity because the amount of MTB cells in sputum varies and can be very low. Secondly, it should be able to detect minor drug-resistant subpopulations in a sample when the majority of the bacilli are susceptible. Two methods could help in this challenge: the invader assay [32] and on-chip ligase detection reaction [123]. The invader assay uses the thermostable flap endonuclease Cleavase VIII, derived from

Archaeoglobus fulgidus, which cleaves a structure formed by the hybridization of two overlapping oligonucleotide probes to a target nucleic acid strand [32]. This method can discriminate single-base differences. On-chip ligase detection reaction is applied to identify approximately 1% of mutant sequences in model samples consisting of mixtures of DNA from wild-type and resistant strains (see [123] for details). These technologies may be useful for clinical research in developed countries, but remain inaccessible for a TB control program on a large scale including low-resource countries.

1.6 CONCLUSION AND PERSPECTIVES

Despite the multitude of investigations launched in various scientific, clinical, and pharmaceutical domains on TB, this disease remains, along with AIDS and malaria, one of the three major killers among infectious diseases. In this chapter, we attempted to demonstrate (i) the contribution of molecular epidemiology in the understanding of transmission and pathogenesis of TB and (ii) the need for routine molecular epidemiology to improve TB surveillance and control programs at global and local levels.

As described by Smith et al. [166], it is clear that new approaches to preventing, diagnosing, and curing tuberculosis are needed, which depend on a better understanding of MTB and the host. They detailed that the National Heart, Lung, and Blood Institute developed recommendations for future TB research [166]. Among these different recommendations, all fundamental for fighting infectious diseases, five directly concern the domain of molecular epidemiology: (i) new resources for characterizing the MTB genome, proteome chips for more specific diagnoses; (ii) prospective studies associated with clinical trials in populations with TB or that are at risk for TB, to advance development of diagnostics and prognosis; (iii) genetic epidemiology studies; (iv) new quantitative and bioinformatics approaches to study the interaction between MTB and the infected host and how this influences the infection process; and (v) coordination between international organizations. This chapter provides evidence that all these points are of public health interest in the fight against TB. We believe that the fourth point is of particular importance, as it is now fully accepted in the scientific area of infectious diseases that the outcome of transmission, infection, and disease are dependent on both the intrinsic characteristics of the microbes and the host. Indeed, as developed by Hide et al. in Chapter 6 on leishmaniasis, integrated analysis of MTB genetics, MTB virulence factors, host immune responses, host genetics, as well as socioeconomic and environmental risk factors are all necessary for a better understanding of the interplay between these different factors and the risk of developing TB. This approach could also provide information on the critical biological pathways involved in the host resistance (latent infection) or susceptibility to TB and therefore help in orienting new therapeutic or vaccine strategies. Indeed, factors determining host resistant/susceptible status are complex and largely not clarified. Moreover, as demonstrated in this chapter, it has been suggested that the outcome of transmission and disease may be MTB strain dependent. This emphasizes the necessity of integrating different approaches to better understand the epidemiological situation's complexity.

ACKNOWLEDGMENTS

We are grateful to the IRD (Institut de Recherche pour le Développement), the CNRS (Centre National de la Recherche Scientifique) and to the "Laboratoire de bactériologie, Hôpital Arnaud-de-Villeneuve," Montpellier France, for financial support. We would like to thank Prof. Philippe Van De Perre, Dr. Jeffrey R. Driscoll and Dr. Stephen V. Gordon, for their critical review.

ABBREVIATIONS

AIDS:	Acquired immune deficiency syndrome
ATP:	Adenosine triphosphate
CAS:	Central Asian
cDNA:	Complementary deoxyribonucleic acid
DNA:	Deoxyribonucleic acid
DOTS:	Directly observed therapy short course
DR:	Direct repeat
DTU:	Discrete typing unit
EAI:	East African-Indian
FDA:	Food and Drug Administration
HIV:	Human immunodeficiency virus
IUATLD:	International Union Against Tuberculosis and Lung Disease
LAM:	Latin American and Mediterranean
LSP:	Large-sequence polymorphism
MDR-TB:	Multidrug-resistant tuberculosis
MIRU:	Mycobacterial interspersed repetitive units
MTB:	*Mycobacterium tuberculosis*
PCR:	Polymerase chain reaction
PPD:	Purified protein derivative
PTB:	Pulmonary tuberculosis
RLFP:	Restriction fragment length polymorphism
SNP:	Single nucleotide polymorphism
SSCP:	Single-strand conformation polymorphism
TB:	Tuberculosis
VNTR:	Variable number tandem repeat
WHO:	World Health Organization
ZN:	Ziehl–Neelsen

GLOSSARY

Allele: A variant of a single gene, inherited at a particular genetic locus; it is a particular sequence of nucleotides.

Allelic frequency: This index is the ratio of the number of a given allele to the total number of alleles in the population under survey.

Bacillus Calmette–Guérin (*BCG*) *vaccine*: A vaccine against tuberculosis that is prepared from a strain of the attenuated (weakened) live bovine tuberculosis bacillus, *Mycobacterium bovis*, that has lost its virulence by special culturing in artificial medium for years. The bacilli have retained sufficient antigenicity to become an effective vaccine for the prevention of human tuberculosis.

Bacteriophage: A virus that infects only bacteria.

Cell-mediated immunity: An immune response that does not involve antibodies but instead involves the activation of macrophages and natural killer cells, the production of antigen-specific cytotoxic T lymphocytes, and the release of various cytokines in response to an antigen.

Clone, clonal, clonality: From a genetic point of view, this term refers to all cases in which the daughter cells are genetically identical to the parental cell, whatever the actual mating system.

Cluster: Refers to a particular genotype shared by two or several MTB isolates.

Conjugation: Bacterial conjugation is the transfer of genetic material between bacteria through cell-to-cell contact.

Cost of resistance: Although mutations that provide resistance to an antibiotic can be considered beneficial, they often come with a physiological cost.

Endemic disease: Disease present or usually prevalent in a population or geographical area at all times.

Epidemiology: This scientific domain corresponds to the study of the distribution and determinants of health-related states and events in populations and the control of health problems.

Etiology: In medicine, the causes of diseases or pathologies.

Fitness: In biology, an individual's ability to propagate its genes.

Genetic drift: This phenomenon is a contributing factor in biological evolution in which traits that do not affect reproductive fitness change in a population over time. Although natural selection causes traits to become more prevalent when they contribute to fitness or eliminates those that harm it, genetic drift is a random process that affects traits that are more neutral.

Haploid: Refers to the ploidy level, that is, the number of copies of the basic number of chromosomes. Haploid cells bear one copy of each chromosome.

Hardy–Weinberg equilibrium: States that under certain conditions after one generation of random mating, the genotype frequencies at a single gene locus will become fixed at a particular equilibrium value. It also specifies that those equilibrium frequencies can be represented as a simple function of the allele frequencies. "Allele frequency" is a term from population genetics that is used in characterizing the genetic diversity of a species population, or equivalently the richness of its gene pool.

Housekeeping gene: A gene that codes for proteins needed all the time for agent survival and multiplication.

Immunosuppression: This immunological status occurs when T and/or B clones of lymphocytes are depleted in size or suppressed in their reactivity, expansion, or differentiation.

Linkage disequilibrium: The nonrandom association of alleles at two or more loci.

Locus: The position of a gene (or other significant sequence) on a chromosome. A locus can be occupied by any of the alleles.

Molecular clock: Refers here to the speed of evolution of a given molecular marker.

Natural selection: A process by which biological populations are altered over time, as a result of the propagation of heritable traits that affect the capacity of individual organisms to survive and reproduce. It is one of several mechanisms that give rise to the evolution of biological species (other mechanisms include genetic drift and gene flow).

Pandemic: Corresponds to a global epidemic and refers to an outbreak of an infectious disease that affects people or animals over an extensive geographical area.

Phagocytosis: This process involves the ingestion and digestion by phagocyte cells of microorganisms, insoluble particles, damaged or dead host cells, cell debris, or activated clotting factors. The principal phagocytes include the neutrophils and monocytes (types of white blood cells).

Phenotype: The observable characteristics of an organism, the expression of gene alleles (genotype) as an observable physical or biochemical trait. It is the result of interaction between the genotype and the environment.

Phylogeny: This scientific domain studies the evolutionary history of a species or group of related species.

Polymerase chain reaction (*PCR*): A technique used to amplify a specific region of DNA. An excess of two amplimers, oligonucleotide primers complementary to two sequences that flank the region to be amplified, are annealed to denatured DNA and subsequently elongated, usually by a heat-stable DNA polymerase from *Thermus aquaticus* (Taq polymerase).

Population genetics: This scientific domain studies the distribution of and change in allele frequencies.

Prevalence: The prevalence of a disease is defined as the ratio of the number of cases of a disease present in a population at

a given time and the number of individuals in the population at that time.

Random mating: This process involves the mating of individuals regardless of any physical, genetic, or social preference. In other words, the mating between two organisms is not influenced by any environmental, hereditary, or social interaction. Hence, potential mates have an equal chance of being selected.

Recombination: In molecular biology, "recombination" generally refers to the molecular process by which alleles at two genes in a linkage group can become separated. In this process, alleles are replaced by different alleles from the same genes, thereby preserving the structure of genes. One mechanism leading to recombination is chromosomal crossover.

Saprotroph: An organism that obtains its nutrients from nonliving organic matter, usually dead and decaying plant or animal matter, by absorbing soluble organic compounds. Because saprotrophs cannot make food for themselves, they are considered as a type of heterotroph (an organism that requires organic substrates to obtain its carbon for growth and development).

Segregation: In biology, this process refers to the separation of homologous chromosomes during mitosis and meiosis.

Symbiosis: An interaction between two organisms living together in more or less intimate association or even the merging of two dissimilar organisms.

Taxonomy: This science refers to the theory and practice of biological classification. This regroups the theories and techniques of naming, describing, and classifying organisms, the study of the relationships of taxa, including positional changes that do not involve changes in the names of taxa.

Transduction: The process in which bacterial DNA is moved from one bacterium to another by a bacterial virus (a bacteriophage, commonly called a phage).

Transformation: In bacteria, "transformation" refers to a genetic change brought about by taking up and recombining DNA, and "competence" refers to the state of being able to take up DNA.

Tuberculin skin test: Tuberculin (also called Mantoux test, currently named Purified Protein Derivative PPD) is an antigen used to aid in the diagnosis of tuberculosis infection. A standard dose of Tuberculin is injected intradermally (into the skin) and read 48–72 h later. A person who has been exposed to the bacteria is expected to mount an immune response in the skin containing the bacterial proteins.

REFERENCES

1. Agerton T, Valway S, Gore B, et al. Transmission of a highly drug-resistant strain (strain W1) of *Mycobacterium tuberculosis*. Community outbreak and nosocomial transmission via a contaminated bronchoscope. *JAMA* 1997;**278**(13):1073–7.
2. Ahmed N, Caviedes L, Alam M, et al. Distinctiveness of *Mycobacterium tuberculosis* genotypes from human immunodeficiency virus type 1-seropositive and -seronegative patients in Lima, Peru. *J Clin Microbiol* 2003;**41**(4):1712–6.
3. Ahmed N, Hasnain SE. Genomics of *Mycobacterium tuberculosis*: old threats and new trends. *Indian J Med Res* 2004;**120**(4):207–12.
4. Anderson T, Brian P, Riggle P, Kong R, Champness W. Genetic suppression analysis of non-antibiotic-producing mutants of the Streptomyces coelicolor absA locus. *Microbiology* 1999;**145**(Pt 9):2343–53.
5. Aranaz A, Cousins D, Mateos A, Dominguez L. Elevation of *Mycobacterium tuberculosis* subsp. caprae to species rank as *Mycobacterium caprae* comb. nov., sp. nov. *Int J Syst Evol Microbiol* 2003;**53**(Pt 6):1785–9.
6. Ayele WY, Neill SD, Zinsstag J, Weiss MG, Pavlik I. Bovine tuberculosis: an old disease but a new threat to Africa. *Int J Tuberc Lung Dis* 2004;**8**(8):924–37.
7. Baker L, Brown T, Maiden MC, Drobniewski F. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg Infect Dis* 2004;**10**(9):1568–77.
8. Baldeviano-Vidalon GC, Quispe-Torres N, Bonilla-Asalde C, Gastiaburu-Rodriguez D, Pro-Cuba JE, Llanos-Zavalaga F. Multiple infection with resistant and sensitive *M. tuberculosis* strains during treatment of pulmonary tuberculosis patients. *Int J Tuberc Lung Dis* 2005;**9**(10):1155–60.
9. Banu S, Gordon SV, Palmer S, et al. Genotypic analysis of *Mycobacterium tuberculosis* in Bangladesh and prevalence of the Beijing strain. *J Clin Microbiol* 2004;**42**(2):674–82.
10. Barnes PF, Cave MD. Molecular epidemiology of tuberculosis. *N Engl J Med* 2003;**349**(12):1149–56.
11. Bifani P, Mathema B, Campo M, et al. Molecular identification of streptomycin monoresistant *Mycobacterium tuberculosis* related to multidrug-resistant W strain. *Emerg Infect Dis* 2001;**7**(5): 842–8.
12. Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. *Trends Microbiol* 2002;**10**(1):45–52.
13. Bifani PJ, Plikaytis BB, Kapur V, et al. Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. *JAMA* 1996;**275**(6):452–7.
14. Bishai WR, Graham NM, Harrington S, et al. Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. *JAMA* 1998;**280**(19):1679–84.
15. Bottger EC, Springer B, Pletschette M, Sander P. Fitness of antibiotic-resistant microorganisms and compensatory mutations. *Nat Med* 1998;**4**(12):1343–4.
16. Braden CR, Morlock GP, Woodley CL, et al. Simultaneous infection with multiple strains of *Mycobacterium tuberculosis*. *Clin Infect Dis* 2001;**33**(6):e42–7.
17. Bradford WZ, Koehler J, El-Hajj H, et al. Dissemination of *Mycobacterium tuberculosis* across the San Francisco Bay Area. *J Infect Dis* 1998;**177**(4):1104–7.
18. Breathnach AS, de Ruiter A, Holdsworth GM, et al. An outbreak of multi-drug-resistant tuberculosis in a London teaching hospital. *J Hosp Infect* 1998;**39**(2):111–7.

19. Brodie D, Schluger NW. The diagnosis of tuberculosis. *Clin Chest Med* 2005;**26**(2):247–71, vi.
20. Brosch R, Gordon SV, Marmiesse M, et al. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA* 2002;**99**(6):3684–9.
21. Burgos MV, Pym AS. Molecular epidemiology of tuberculosis. *Eur Respir J Suppl* 2002;**36**:54s–65s.
22. Butcher PD. Microarrays for *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 2004;**84**(3–4):131–7.
23. Caminero JA. Management of multidrug-resistant tuberculosis and patients in retreatment. *Eur Respir J* 2005;**25**(5):928–36.
24. Caminero JA, Pena MJ, Campos-Herrero MI, et al. Epidemiological evidence of the spread of a *Mycobacterium tuberculosis* strain of the Beijing genotype on Gran Canaria Island. *Am J Respir Crit Care Med* 2001;**164**(7):1165–70.
25. Castets M, Sarrat H. Experimental study of the virulence of *Mycobacterium africanum* (preliminary note). *Bull Soc Med Afr Noire Lang Fr* 1969;**14**(4):693–6.
26. Chan-Yeung M, Tam CM, Wong H, et al. Molecular and conventional epidemiology of tuberculosis in Hong Kong: a population-based prospective study. *J Clin Microbiol* 2003;**41**(6): 2706–8.
27. Chaves F, Dronda F, Alonso-Sanz M, Noriega AR. Evidence of exogenous reinfection and mixed infection with more than one strain of *Mycobacterium tuberculosis* among Spanish HIV-infected inmates. *Aids* 1999;**13**(5):615–20.
28. Chiang CY, Riley LW. Exogenous reinfection in tuberculosis. *Lancet Infect Dis* 2005;**5**(10):629–36.
29. Cohn ML, Kovitz C, Oda U, Middlebrook G. Studies on isoniazid and tubercle bacilli. II. The growth requirements, catalase activities, and pathogenic properties of isoniazid-resistant mutants. *Am Rev Tuberc* 1954;**70**(4):641–64.
30. Cole ST, Brosch R, Parkhill J, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998;**393**(6685):537–44.
31. Colebunders R, Bastian I. A review of the diagnosis and treatment of smear-negative pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2000;**4**(2):97–107.
32. Cooksey RC, Holloway BP, Oldenburg MC, Listenbee S, Miller CW. Evaluation of the invader assay, a linear signal amplification method, for identification of mutations associated with resistance to rifampin and isoniazid in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 2000;**44**(5):1296–301.
33. Coronado VG, Beck-Sague CM, Hutton MD, et al. Transmission of multidrug-resistant *Mycobacterium tuberculosis* among persons with human immunodeficiency virus infection in an urban hospital: epidemiologic and restriction fragment length polymorphism analysis. *J Infect Dis* 1993;**168**(4): 1052–5.
34. Cousins DV, Bastida R, Cataldi A, et al. Tuberculosis in seals caused by a novel member of the *Mycobacterium tuberculosis* complex: *Mycobacterium pinnipedii* sp. nov. *Int J Syst Evol Microbiol* 2003;**53**(Pt 5):1305–14.
35. Cowan LS, Diem L, Brake MC, Crawford JT. Transfer of a *Mycobacterium tuberculosis* genotyping method, spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. *J Clin Microbiol* 2004;**42**(1):474–7.
36. Cowan LS, Diem L, Monson T, et al. Evaluation of a two-step approach for large-scale, prospective genotyping of *Mycobacterium tuberculosis* isolates in the United States. *J Clin Microbiol* 2005; **43**(2):688–95.
37. Cowan LS, Mosher L, Diem L, Massey JP, Crawford JT. Variable-number tandem repeat typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110 by using mycobacterial interspersed repetitive units. *J Clin Microbiol* 2002;**40**(5): 1592–602.
38. Curtis AB, Ridzon R, Novick LF, et al. Analysis of *Mycobacterium tuberculosis* transmission patterns in a homeless shelter outbreak. *Int J Tuberc Lung Dis* 2000;**4**(4):308–13.
39. Dale JW, Brittain D, Cataldi AA, et al. Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: recommendations for standardised nomenclature. *Int J Tuberc Lung Dis* 2001;**5**(3):216–9.
40. Daley CL. Molecular epidemiology: a tool for understanding control of tuberculosis transmission. *Clin Chest Med* 2005;**26**(2):217–31, vi.
41. Daley CL, Kawamura LM. The role of molecular epidemiology in contact investigations: a US perspective. *Int J Tuberc Lung Dis* 2003;**7**(12 Suppl 3):S458–62.
42. Daley CL, Small PM, Schecter GF, et al. 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* 1992;**326**(4):231–5.
43. Dannenberg AM, Jr. Delayed-type hypersensitivity and cell-mediated immunity in the pathogenesis of tuberculosis. *Immunol Today* 1991;**12**(7):228–33.
44. Dannenberg AM, Jr. Immune mechanisms in the pathogenesis of pulmonary tuberculosis. *Rev Infect Dis* 1989;**11**(Suppl 2): S369–78.
45. de Boer AS, Borgdorff MW, Vynnycky E, Sebek MM, Van Soolingen D. Exogenous re-infection as a cause of recurrent tuberculosis in a low-incidence area. *Int J Tuberc Lung Dis* 2003;**7**(2):145–52.
46. De Bruyn G, Adams GJ, Teeter LD, Soini H, Musser JM, Graviss EA. The contribution of ethnicity to *Mycobacterium tuberculosis* strain clustering. *Int J Tuberc Lung Dis* 2001;**5**(7): 633–41.
47. Diel R, Meywald-Walter K, Gottschalk R, Rusch-Gerdes S, Niemann S. Ongoing outbreak of tuberculosis in a low-incidence community: a molecular-epidemiological evaluation. *Int J Tuberc Lung Dis* 2004;**8**(7):855–61.
48. Diel R, Rusch-Gerdes S, Niemann S. Molecular epidemiology of tuberculosis among immigrants in Hamburg, Germany. *J Clin Microbiol* 2004;**42**(7):2952–60.
49. Doherty TM, Andersen P. Vaccines for tuberculosis: novel concepts and recent progress. *Clin Microbiol Rev* 2005;**18**(4): 687–702, table of contents.
50. Dooley SW, Jarvis WR, Martone WJ, Snider DE, Jr. Multidrug-resistant tuberculosis. *Ann Intern Med* 1992;**117**(3):257–9.
51. Dooley SW, Villarino ME, Lawrence M, et al. Nosocomial transmission of tuberculosis in a hospital unit for HIV-infected patients. *JAMA* 1992;**267**(19):2632–4.
52. Drobniewski FA, Gibson A, Ruddy M, Yates MD. Evaluation and utilization as a public health tool of a national molecular

epidemiological tuberculosis outbreak database within the United Kingdom from 1997 to 2001. *J Clin Microbiol* 2003; **41**(5):1861–8.
53. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. *Nat Genet* 1999;**21** (Suppl 1):10–4.
54. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *JAMA* 1999;**282**(7): 677–86.
55. Edlin BR, Tokars JI, Grieco MH, et al. An outbreak of multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome. *N Engl J Med* 1992;**326**(23):1514–21.
56. Ellis BA, Crawford JT, Braden CR, McNabb SJ, Moore M, Kammerer S. Molecular epidemiology of tuberculosis in a sentinel surveillance population. *Emerg Infect Dis* 2002;**8**(11): 1197–209.
57. Fang Z, Morrison N, Watt B, Doig C, Forbes KJ. IS6110 transposition and evolutionary scenario of the direct repeat locus in a group of closely related *Mycobacterium tuberculosis* strains. *J Bacteriol* 1998;**180**(8):2102–9.
58. Feil EJ, Holmes EC, Bessen DE, et al. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc Natl Acad Sci USA* 2001;**98**(1):182–7.
59. Feil EJ, Spratt BG. Recombination and the population structures of bacterial pathogens. *Annu Rev Microbiol* 2001;**55**: 561–90.
60. Filliol I, Driscoll JR, Van Soolingen D, et al. Global distribution of *Mycobacterium tuberculosis* spoligotypes. *Emerg Infect Dis* 2002;**8**(11):1347–9.
61. Filliol I, Driscoll JR, Van Soolingen D, et al. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. *J Clin Microbiol* 2003;**41**(5):1963–70.
62. Filliol I, Sola C, Rastogi N. Detection of a previously unamplified spacer within the DR locus of *Mycobacterium tuberculosis*: epidemiological implications. *J Clin Microbiol* 2000;**38**(3): 1231–4.
63. Fleischmann RD, Alland D, Eisen JA, et al. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J Bacteriol* 2002;**184**(19):5479–90.
64. Formicola V, Milanesi Q, Scarsini C. Evidence of spinal tuberculosis at the beginning of the fourth millennium BC from Arene Candide cave (Liguria, Italy). *Am J Phys Anthropol* 1987;**72**(1):1–6.
65. Frieden TR, Munsiff SS. The DOTS strategy for controlling the global tuberculosis epidemic. *Clin Chest Med* 2005;**26**(2): 197–205, v.
66. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C. Tuberculosis. *Lancet* 2003;**362**(9387):887–99.
67. Frieden TR, Woodley CL, Crawford JT, Lew D, Dooley SM. The molecular epidemiology of tuberculosis in New York City: the importance of nosocomial transmission and laboratory error. *Tuberc Lung Dis* 1996;**77**(5):407–13.
68. Frothingham R, Meeker-O'Connell WA. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology* 1998;**144**(Pt 5):1189–96.
69. Garcia-Garcia M, Palacios-Martinez M, Ponce-de-Leon A, et al. The role of core groups in transmitting *Mycobacterium tuberculosis* in a high prevalence community in Southern Mexico. *Int J Tuberc Lung Dis* 2000;**4**(1):12–7.
70. Garnier T, Eiglmeier K, Camus JC, et al. The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci USA* 2003; **100**(13):7877–82.
71. Geng E, Kreiswirth B, Driver C, et al. Changes in the transmission of tuberculosis in New York City from 1990 to 1999. *N Engl J Med* 2002;**346**(19):1453–8.
72. Gingeras TR, Ghandour G, Wang E, et al. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res* 1998;**8**(5):435–48.
73. Ginsberg AM. What's new in tuberculosis vaccines? *Bull World Health Organ* 2002;**80**(6):483–8.
74. Githui WA. Laboratory methods for diagnosis and detection of drug resistant *Mycobacterium tuberculosis* complex with reference to developing countries: a review. *East Afr Med J* 2002;**79**(5): 242–8.
75. Glynn JR, Bauer J, de Boer AS, et al. Interpreting DNA fingerprint clusters of *Mycobacterium tuberculosis*. European Concerted Action on Molecular Epidemiology and Control of Tuberculosis. *Int J Tuberc Lung Dis* 1999;**3**(12):1055–60.
76. Glynn JR, Crampin AC, Yates MD, et al. The importance of recent infection with *Mycobacterium tuberculosis* in an area with high HIV prevalence: a long-term molecular epidemiological study in Northern Malawi. *J Infect Dis* 2005;**192**(3):480–7.
77. Glynn JR, Whiteley J, Bifani PJ, Kremer K, Van Soolingen D. Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg Infect Dis* 2002;**8**(8): 843–9.
78. Godfrey-Faussett P, Sonnenberg P, Shearer SC, et al. Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet* 2000;**356**(9235): 1066–71.
79. Godfrey-Faussett P, Stoker NG, Scott JA, Pasvol G, Kelly P, Clancy L. DNA fingerprints of *Mycobacterium tuberculosis* do not change during the development of rifampicin resistance. *Tuberc Lung Dis* 1993;**74**(4):240–3.
80. Gormley EP, Davies J. Transfer of plasmid RSF1010 by conjugation from *Escherichia coli* to *Streptomyces lividans* and *Mycobacterium smegmatis*. *J Bacteriol* 1991;**173**(21):6705–8.
81. Groenen PM, Bunschoten AE, Van Soolingen D, van Embden JD. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol Microbiol* 1993;**10**(5):1057–65.
82. Gutacker MM, Smoot JC, Migliaccio CA, et al. Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains. *Genetics* 2002;**162**(4):1533–43.
83. Gutierrez MC, Brisse S, Brosch R, et al. Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog* 2005;**1**(1):e5.
84. Hampton T. TB drug research picks up the pace. *JAMA* 2005;**293**(22):2705–7.

85. Hatfull GF, Jacobs WR, Jr. Molecular Genetics of Mycobacteriophages. American Society for Microbiogy Press, Washington, DC, 2000.
86. Hawkey PM, Smith EG, Evans JT, et al. Mycobacterial interspersed repetitive unit typing of *Mycobacterium tuberculosis* compared to IS6110-based restriction fragment length polymorphism analysis for investigation of apparently clustered cases of tuberculosis. *J Clin Microbiol* 2003;**41**(8):3514–20.
87. Hermans PW, Messadi F, Guebrexabher H, et al. 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* 1995;**171**(6): 1504–13.
88. Hermans PW, Van Soolingen D, Bik EM, de Haas PE, Dale JW, van Embden JD. 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* 1991;**59**(8):2695–705.
89. Hernandez-Pando R, Jeyanathan M, Mengistu G, et al. Persistence of DNA from *Mycobacterium tuberculosis* in superficially normal lung tissue during latent infection. *Lancet* 2000;**356**(9248):2133–8.
90. Itah AY, Udofia SM. Epidemiology and endemicity of pulmonary tuberculosis (PTB) in Southeastern Nigeria. *Southeast Asian J Trop Med Public Health* 2005;**36**(2):317–23.
91. Jalava J, Marttila H. Application of molecular genetic methods in macrolide, lincosamide and streptogramin resistance diagnostics and in detection of drug-resistant *Mycobacterium tuberculosis*. *Apmis* 2004;**112**(11–12):838–55.
92. Jereb JA, Burwen DR, Dooley SW, et al. Nosocomial outbreak of tuberculosis in a renal transplant unit: application of a new technique for restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates. *J Infect Dis* 1993;**168**(5):1219–24.
93. Jereb JA, Klevens RM, Privett TD, et al. Tuberculosis in health care workers at a hospital with an outbreak of multidrug-resistant *Mycobacterium tuberculosis*. *Arch Intern Med* 1995;**155**(8): 854–9.
94. Kamerbeek J, Schouls L, Kolk A, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997;**35**(4):907–14.
95. Kato-Maeda M, Rhee JT, Gingeras TR, et al. Comparing genomes within the species *Mycobacterium tuberculosis*. *Genome Res* 2001;**11**(4):547–54.
96. Kaufmann SH, McMichael AJ. Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. *Nat Med* 2005;**11**(4 Suppl):S33–44.
97. Kaul KL. Molecular detection of *Mycobacterium tuberculosis*: impact on patient care. *Clin Chem* 2001;**47**(8):1553–8.
98. Kremer K, Au BK, Yip PC, et al. Use of variable-number tandem-repeat typing to differentiate *Mycobacterium tuberculosis* Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping. *J Clin Microbiol* 2005;**43**(1):314–20.
99. Kremer K, Van Soolingen D, Frothingham R, et al. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* 1999;**37**(8):2607–18.
100. Kruuner A, Danilovitsh M, Pehme L, Laisaar T, Hoffner SE, Katila ML. Tuberculosis as an occupational hazard for health care workers in Estonia. *Int J Tuberc Lung Dis* 2001;**5**(2): 170–6.
101. Lagrange PW, A. Hermann JL. Physiopathologie et immunité de l'infection tuberculeuse. In: bio M, ed. *Mycobacterium tuberculosis* et mycobacteries atypiques. Elsevier, Paris, 2004, pp. 19–45.
102. Lambert ML, Hasker E, Van Deun A, Roberfroid D, Boelaert M, Van der Stuyft P. Recurrence in tuberculosis: relapse or reinfection? *Lancet Infect Dis* 2003;**3**(5):282–7.
103. Laserson KF, Osorio L, Sheppard JD, et al. Clinical and programmatic mismanagement rather than community outbreak as the cause of chronic, drug-resistant tuberculosis in Buenaventura, Colombia, 1998. *Int J Tuberc Lung Dis* 2000;**4**(7):673–83.
104. Lazcano-Ponce E, Allen B, Gonzalez CC. The contribution of international agencies to the control of communicable diseases. *Arch Med Res* 2005;**36**(6):731–8.
105. Lee AS, Lim IH, Tang LL, Wong SY. High frequency of mutations in the rpoB gene in rifampin-resistant clinical isolates of *Mycobacterium tuberculosis* from Singapore. *J Clin Microbiol* 2005; **43**(4):2026–7.
106. Legrand E, Filliol I, Sola C, Rastogi N. Use of spoligotyping to study the evolution of the direct repeat locus by IS6110 transposition in *Mycobacterium tuberculosis*. *J Clin Microbiol* 2001; **39**(4):1595–9.
107. Levin BR, Lipsitch M, Perrot V, et al. The population genetics of antibiotic resistance. *Clin Infect Dis* 1997;**24**(Suppl 1):S9–16.
108. Levin BR, Perrot V, Walker N. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* 2000;**154**(3):985–97.
109. Levy-Frebault VV, Portaels F. Proposed minimal standards for the genus *Mycobacterium* and for description of new slowly growing *Mycobacterium* species. *Int J Syst Bacteriol* 1992;**42** (2):315–23.
110. Lillebaek T, Andersen AB, Dirksen A, Smith EG, Skovgaard LT, Kok-Jensen A. Persistent high incidence of tuberculosis in immigrants in a low-incidence country. *Emerg Infect Dis* 2002;**8**(7):679–84.
111. Lockman S, Sheppard JD, Braden CR, et al. Molecular and conventional epidemiology of *Mycobacterium tuberculosis* in Botswana: a population-based prospective study of 301 pulmonary tuberculosis patients. *J Clin Microbiol* 2001;**39** (3):1042–7.
112. Lopez B, Aguilar D, Orozco H, et al. A marked difference in pathogenesis and immune response induced by different *Mycobacterium tuberculosis* genotypes. *Clin Exp Immunol* 2003;**133**(1):30–7.
113. Manabe YC, Bishai WR. Latent *Mycobacterium tuberculosis*-persistence, patience, and winning by waiting. *Nat Med* 2000;**6**(12):1327–9.
114. Manabe YC, Dannenberg AM, Jr., Bishai WR. What we can learn from the *Mycobacterium tuberculosis* genome sequencing projects. *Int J Tuberc Lung Dis* 2000;**4**(2 Suppl 1):S18–23.
115. March F, Coll P, Costa R, et al. Usefulness of DR, PGRS, and spoligotyping in the typing of *Mycobacterium tuberculosis*. Comparison with IS6110. *Enferm Infecc Microbiol Clin* 1996;**14**(3):160–6.

116. Mardassi H, Namouchi A, Haltiti R, et al. Tuberculosis due to resistant Haarlem strain, Tunisia. *Emerg Infect Dis* 2005;**11** (6):957–61.
117. Martin G, Lazarus A. Epidemiology and diagnosis of tuberculosis. Recognition of at-risk patients is key to prompt detection. *Postgrad Med* 2000;**108**(2):42–4, 47–50, 53–4.
118. Mazars E, Lesjean S, Banuls AL, et al. High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. *Proc Natl Acad Sci USA* 2001;**98**(4):1901–6.
119. McDonough KA, Kress Y, Bloom BR. Pathogenesis of tuberculosis: interaction of *Mycobacterium tuberculosis* with macrophages. *Infect Immun* 1993;**61**(7):2763–73.
120. McElroy PD, Sterling TR, Driver CR, et al. Use of DNA fingerprinting to investigate a multiyear, multistate tuberculosis outbreak. *Emerg Infect Dis* 2002;**8**(11):1252–6.
121. Meacci F, Orru G, Iona E, et al. Drug resistance evolution of a *Mycobacterium tuberculosis* strain from a noncompliant patient. *J Clin Microbiol* 2005;**43**(7):3114–20.
122. Middlebrook G, Cohn ML. Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. *Science* 1953;**118**(3063):297–9.
123. Mikhailovich V, Lapa S, Gryadunov D, et al. Identification of rifampin-resistant *Mycobacterium tuberculosis* strains by hybridization, PCR, and ligase detection reaction on oligonucleotide microchips. *J Clin Microbiol* 2001;**39**(7):2531–40.
124. Morgan MA, Horstmeier CD, DeYoung DR, Roberts GD. Comparison of a radiometric method (BACTEC) and conventional culture media for recovery of mycobacteria from smear-negative specimens. *J Clin Microbiol* 1983;**18**(2):384–8.
125. Moro ML, Gori A, Errante I, et al. An outbreak of multidrug-resistant tuberculosis involving HIV-infected patients of two hospitals in Milan, Italy. Italian Multidrug-Resistant Tuberculosis Outbreak Study Group. *Aids* 1998;**12**(9): 1095–102.
126. Moss AR, Alland D, Telzak E, et al. A city-wide outbreak of a multiple-drug-resistant strain of *Mycobacterium tuberculosis* in New York. *Int J Tuberc Lung Dis* 1997;**1**(2):115–21.
127. Munsiff SS, Bassoff T, Nivin B, et al. Molecular epidemiology of multidrug-resistant tuberculosis, New York City, 1995–1997. *Emerg Infect Dis* 2002;**8**(11):1230–8.
128. Munsiff SS, Nivin B, Sacajiu G, Mathema B, Bifani P, Kreiswirth BN. Persistence of a highly resistant strain of tuberculosis in New York City during 1990–1999. *J Infect Dis* 2003;**188**(3):356–63.
129. Musser JM. Molecular population genetic analysis of emerged bacterial pathogens: selected insights. *Emerg Infect Dis* 1996;**2**(1):1–17.
130. Musser JM, Amin A, Ramaswamy S. Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure. *Genetics* 2000;**155**(1):7–16.
131. Nagelkerke NJ, de Vlas SJ, Mahendradhata Y, Ottenhoff TH, Borgdorff M. The search for a tuberculosis vaccine: an elusive quest? *Tuberculosis (Edinb)* 2006;**86**(1):41–6.
132. Nguyen LN, Gilbert GL, Marks GB. Molecular epidemiology of tuberculosis and recent developments in understanding the epidemiology of tuberculosis. *Respirology* 2004;**9**(3):313–9.
133. Niobe-Eyangoh SN, Kuaban C, Sorlin P, et al. Genetic biodiversity of *Mycobacterium tuberculosis* complex strains from patients with pulmonary tuberculosis in Cameroon. *J Clin Microbiol* 2003;**41**(6):2547–53.
134. O'Brien RJ, Spigelman M. New drugs for tuberculosis: current status and future prospects. *Clin Chest Med* 2005;**26**(2):327–40, vii.
135. Oettinger T, Jorgensen M, Ladefoged A, Haslov K, Andersen P. Development of the *Mycobacterium bovis* BCG vaccine: review of the historical and biochemical evidence for a genealogical tree. *Tuberc Lung Dis* 1999;**79**(4):243–50.
136. Okeke IN, Klugman KP, Bhutta ZA, et al. Antimicrobial resistance in developing countries. Part II. Strategies for containment. *Lancet Infect Dis* 2005;**5**(9):568–80.
137. Orme IM. The immunopathogenesis of tuberculosis: a new working hypothesis. *Trends Microbiol* 1998;**6**(3):94–7.
138. O'Rourke M, Stevens E. Genetic structure of *Neisseria gonorrhoeae* populations: a non-clonal pathogen. *J Gen Microbiol* 1993;**139**(11):2603–11.
139. Otal I, Martin C, Vincent-Levy-Frebault V, Thierry D, Gicquel B. Restriction fragment length polymorphism analysis using IS6110 as an epidemiological marker in tuberculosis. *J Clin Microbiol* 1991;**29**(6):1252–4.
140. Pang Y, Brown BA, Steingrube VA, Wallace RJ, Jr., Roberts MC. Tetracycline resistance determinants in *Mycobacterium* and *Streptomyces* species. *Antimicrob Agents Chemother* 1994;**38**(6): 1408–12.
141. Paolo WF, Jr., Nosanchuk JD. Tuberculosis in New York city: recent lessons and a look ahead. *Lancet Infect Dis* 2004;**4**(5):287–93.
142. Parrish NM, Dick JD, Bishai WR. Mechanisms of latency in *Mycobacterium tuberculosis*. *Trends Microbiol* 1998;**6**(3):107–12.
143. Parsons LM, Jankowski CS, Derbyshire KM. Conjugal transfer of chromosomal DNA in *Mycobacterium smegmatis*. *Mol Microbiol* 1998;**28**(3):571–82.
144. Perkins MD. New diagnostic tools for tuberculosis. *Int J Tuberc Lung Dis* 2000;**4**(12 Suppl 2):S182–8.
145. Pineda-Garcia L, Ferrera A, Hoffner SE. DNA fingerprinting of *Mycobacterium tuberculosis* strains from patients with pulmonary tuberculosis in Honduras. *J Clin Microbiol* 1997;**35**(9): 2393–7.
146. Post FA, Willcox PA, Mathema B, et al. Genetic polymorphism in *Mycobacterium tuberculosis* isolates from patients with chronic multidrug-resistant tuberculosis. *J Infect Dis* 2004;**190** (1):99–106.
147. Pozzi G, Meloni M, Iona E, et al. rpoB mutations in multidrug-resistant strains of *Mycobacterium tuberculosis* isolated in Italy. *J Clin Microbiol* 1999;**37**(4):1197–9.
148. Raviglione MC, Snider DE, Jr., Kochi A. Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic. *JAMA* 1995;**273**(3):220–6.
149. Resende MR, Villares MC, Ramos Mde C. Transmission of tuberculosis among patients with human immunodeficiency virus at a university hospital in Brazil. *Infect Control Hosp Epidemiol* 2004;**25**(12):1115–7.
150. Rhee JT, Tanaka MM, Behr MA, et al. Use of multiple markers in population-based molecular epidemiologic studies of tuberculosis. *Int J Tuberc Lung Dis* 2000;**4**(12):1111–9.

151. Ridzon R, Kent JH, Valway S, et al. Outbreak of drug-resistant tuberculosis with second-generation transmission in a high school in California. *J Pediatr* 1997;**131**(6):863–8.
152. Riley MV, Maegraith BG. Changes in the metabolism of liver mitochondria of mice infected with rapid acute *Plasmodium berghei* malaria. *Ann Trop Med Parasitol* 1962;**56**:473–82.
153. Ritacco V, Di Lonardo M, Reniero A, et al. Nosocomial spread of human immunodeficiency virus-related multidrug-resistant tuberculosis in Buenos Aires. *J Infect Dis* 1997;**176**(3):637–42.
154. Rogall T, Flohr T, Bottger EC. Differentiation of *Mycobacterium* species by direct sequencing of amplified DNA. *J Gen Microbiol* 1990;**136**(9):1915–20.
155. Rook GA, Dheda K, Zumla A. Immune responses to tuberculosis in developing countries: implications for new vaccines. *Nat Rev Immunol* 2005;**5**(8):661–7.
156. Saenghirunvattana S, Charoenpan P, Vathesatogkit P, Kiatboonsri S, Aeursudkij B. Multidrug-resistant tuberculosis: response to treatment. *J Med Assoc Thai* 1996;**79**(9):601–3.
157. Sahbazian B, Weis SE. Treatment of active tuberculosis: challenges and prospects. *Clin Chest Med* 2005;**26**(2):273–82, vi.
158. Salo WL, Aufderheide AC, Buikstra J, Holcomb TA. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc Natl Acad Sci USA* 1994;**91**(6):2091–4.
159. Saunders NA, Metherell L, Patel S. Investigation of an outbreak of multidrug resistant tuberculosis among renal patients using rpo B gene sequencing and IS6110 inverse PCR. *J Infect* 1997;**35**(2):129–33.
160. Schaaf HS, Botha P, Beyers N, et al. The 5-year outcome of multidrug resistant tuberculosis patients in the Cape Province of South Africa. *Trop Med Int Health* 1996;**1**(5): 718–22.
161. Shamputa IC, Rigouts L, Eyongeta LA, et al. Genotypic and phenotypic heterogeneity among *Mycobacterium tuberculosis* isolates from pulmonary tuberculosis patients. *J Clin Microbiol* 2004;**42**(12):5528–36.
162. Sharp SE, Lemes M, Sierra SG, Poniecka A, Poppiti RJ, Jr. Lowenstein-Jensen media. No longer necessary for mycobacterial isolation. *Am J Clin Pathol* 2000;**113**(6):770–3.
163. Shinnick TM, Good RC. Mycobacterial taxonomy. *Eur J Clin Microbiol Infect Dis* 1994;**13**(11):884–901.
164. Small PM, Hopewell PC, Singh SP, et al. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. *N Engl J Med* 1994;**330**(24):1703–9.
165. Small PM, Moss A. Molecular epidemiology and the new tuberculosis. *Infect Agents Dis* 1993;**2**(3):132–8.
166. Smith I, Nathan C, Peavy HH. NHLBI Working Group: progress and new directions in genetics of tuberculosis. *Am J Respir Crit Care Med* 2005;**172**(12):1491–6.
167. Smith JM, Smith NH, O'Rourke M, Spratt BG. How clonal are bacteria? *Proc Natl Acad Sci USA* 1993;**90**(10):4384–8.
168. Smith SM, Dockrell HM. Role of CD8 T cells in mycobacterial infections. *Immunol Cell Biol* 2000;**78**(4):325–33.
169. Soini H, Pan X, Amin A, Graviss EA, Siddiqui A, Musser JM. Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston, Texas, by spoligotyping. *J Clin Microbiol* 2000;**38**(2):669–76.
170. Soini H, Pan X, Teeter L, Musser JM, Graviss EA. Transmission dynamics and molecular characterization of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110. *J Clin Microbiol* 2001;**39**(1):217–21.
171. Sola C, Devallois A, Horgen L, et al. Tuberculosis in the Caribbean: using spacer oligonucleotide typing to understand strain origin and transmission. *Emerg Infect Dis* 1999;**5** (3):404–14.
172. Sola C, Ferdinand S, Mammina C, Nastasi A, Rastogi N. Genetic diversity of *Mycobacterium tuberculosis* in Sicily based on spoligotyping and variable number of tandem DNA repeats and comparison with a spoligotyping database for population-based analysis. *J Clin Microbiol* 2001;**39**(4):1559–65.
173. Sola C, Horgen L, Devallois A, Rastogi N. Combined numerical analysis based on the molecular description of *Mycobacterium tuberculosis* by four repetitive sequence-based DNA typing systems. *Res Microbiol* 1998;**149**(5):349–60.
174. Sonnenberg P, Murray J, Glynn JR, Shearer S, Kambashi B, Godfrey-Faussett P. HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers. *Lancet* 2001;**358**(9294):1687–93.
175. Spratt BG. Antibiotic resistance: counting the cost. *Curr Biol* 1996;**6**(10):1219–21.
176. Spratt BG, Maiden MC. Bacterial population genetics, evolution and epidemiology. *Philos Trans R Soc Lond B Biol Sci* 1999;**354**(1384):701–10.
177. Sreevatsan S, Pan X, Stockbauer KE, et al. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci USA* 1997;**94**(18):9869–74.
178. Sterling TR, Thompson D, Stanley RL, et al. A multi-state outbreak of tuberculosis among members of a highly mobile social network: implications for tuberculosis elimination. *Int J Tuberc Lung Dis* 2000;**4**(11):1066–73.
179. Sturgill-Koszycki S, Schlesinger PH, Chakraborty P, et al. Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase. *Science* 1994;**263** (5147):678–81.
180. Suarez PG, Watt CJ, Alarcon E, et al. The dynamics of tuberculosis in response to 10 years of intensive control effort in Peru. *J Infect Dis* 2001;**184**(4):473–8.
181. Sun YJ, Lee AS, Ng ST, et al. Characterization of ancestral *Mycobacterium tuberculosis* by multiple genetic markers and proposal of genotyping strategy. *J Clin Microbiol* 2004;**42** (11):5058–64.
182. Supply P, Magdalena J, Himpens S, Locht C. Identification of novel intergenic repetitive units in a mycobacterial two-component system operon. *Mol Microbiol* 1997;**26**(5):991–1003.
183. Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B, Locht C. Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. *Mol Microbiol* 2000;**36**(3):762–71.
184. Supply P, Warren RM, Banuls AL, et al. Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. *Mol Microbiol* 2003;**47**(2):529–38.
185. Tahaoglu K, Kizkin O, Karagoz T, Tor M, Partal M, Sadoglu T. High initial and acquired drug resistance in pulmonary tuberculosis in Turkey. *Tuberc Lung Dis* 1994;**75**(5):324–8.

186. Taylor M. TB outbreak fallout. ER doc who contracted disease sues hospital for lack of warning, precautions. *Mod Healthc* 1999;**29**(19):40–2.
187. Tazi L, El Baghdadi J, Lesjean S, et al. Genetic diversity and population structure of *Mycobacterium tuberculosis* in Casablanca, a Moroccan city with high incidence of tuberculosis. *J Clin Microbiol* 2004;**42**(1):461–6.
188. Tazi L, Kreiswirth B, Carriere C, Tibayrenc M. Molecular epidemiology of *Mycobacterium tuberculosis* and its relevance to the surveillance and control of TB: an e-debate. *Infect Genet Evol* 2002;**2**(2):153–8.
189. Telenti A. Genetics of drug resistance in tuberculosis. *Clin Chest Med* 1997;**18**(1):55–64.
190. Theisen A, Reichel C, Rusch-Gerdes S, et al. Mixed-strain infection with a drug-sensitive and multidrug-resistant strain of *Mycobacterium tuberculosis*. *Lancet* 1995;**345**(8963):1512.
191. Thierry D, Cave MD, Eisenach KD, et al. IS6110, an IS-like element of *Mycobacterium tuberculosis* complex. *Nucleic Acids Res* 1990;**18**(1):188.
192. Tibayrenc M. Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. *Int J Parasitol* 1998;**28**(1):85–104.
193. Tibayrenc M. Population genetics and strain typing of microorganisms: how to detect departures from panmixia without individualizing alleles and loci. *C R Acad Sci III* 1995;**318**(1):135–9.
194. Tibayrenc M. Population genetics of parasitic protozoa and other microorganisms. *Adv Parasitol* 1995;**36**:47–115.
195. Tibayrenc M, Ayala FJ. Towards a population genetics of microorganisms: the clonal theory of parasitic protozoa. *Parasitol Today* 1991;**7**(9):228–32.
196. Torrea G, Levee G, Grimont P, Martin C, Chanteau S, Gicquel B. Chromosomal DNA fingerprinting analysis using the insertion sequence IS6110 and the repetitive element DR as strain-specific markers for epidemiological study of tuberculosis in French Polynesia. *J Clin Microbiol* 1995;**33**(7):1899–904.
197. Toungoussova OS, Mariandyshev A, Bjune G, Sandven P, Caugant DA. Molecular epidemiology and drug resistance of *Mycobacterium tuberculosis* isolates in the Archangel prison in Russia: predominance of the W-Beijing clone family. *Clin Infect Dis* 2003;**37**(5):665–72.
198. Troesch A, Nguyen H, Miyada CG, et al. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. *J Clin Microbiol* 1999;**37**(1):49–55.
199. Tsolaki AG, Hirsh AE, DeRiemer K, et al. Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *Proc Natl Acad Sci USA* 2004;**101**(14):4865–70.
200. Valway SE, Richards SB, Kovacovich J, Greifinger RB, Crawford JT, Dooley SW. Outbreak of multi-drug-resistant tuberculosis in a New York State prison, 1991. *Am J Epidemiol* 1994;**140**(2):113–22.
201. Van Belkum A, Struelens M, de Visser A, Verbrugh H, Tibayrenc M. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. *Clin Microbiol Rev* 2001;**14**(3):547–60.
202. van der Zanden AG, Kremer K, Schouls LM, et al. Improvement of differentiation and interpretability of spoligotyping for *Mycobacterium tuberculosis* complex isolates by introduction of new spacer oligonucleotides. *J Clin Microbiol* 2002;**40**(12):4628–39.
203. van Deutekom H, Gerritsen JJ, Van Soolingen D, van Ameijden EJ, van Embden JD, Coutinho RA. A molecular epidemiological approach to studying the transmission of tuberculosis in Amsterdam. *Clin Infect Dis* 1997;**25**(5):1071–7.
204. van Deutekom H, Hoijng SP, de Haas PE, et al. Clustered tuberculosis cases: do they represent recent transmission and can they be detected earlier? *Am J Respir Crit Care Med* 2004;**169**(7):806–10.
205. van Embden JD, van Gorkom T, Kremer K, Jansen R, van Der Zeijst BA, Schouls LM. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J Bacteriol* 2000;**182**(9):2393–401.
206. van Rie A, Warren R, Richardson M, et al. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *N Engl J Med* 1999;**341**(16):1174–9.
207. Van Soolingen D, Borgdorff MW, de Haas PE, et al. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 1999;**180**(3):726–36.
208. Van Soolingen D, de Haas PE, Hermans PW, Groenen PM, van Embden JD. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1993;**31**(8):1987–95.
209. Van Soolingen D, Hermans PW. Epidemiology of tuberculosis by DNA fingerprinting. *Eur Respir J Suppl* 1995;**20**:649s–56s.
210. Van Soolingen D, Hoogenboezem T, de Haas PE, et al. A novel pathogenic taxon of the *Mycobacterium tuberculosis* complex, Canetti: characterization of an exceptional isolate from Africa. *Int J Syst Bacteriol* 1997;**47**(4):1236–45.
211. Victor TC, Warren R, Beyers N, van Helden PD. Transmission of multidrug-resistant strains of *Mycobacterium tuberculosis* in a high incidence community. *Eur J Clin Microbiol Infect Dis* 1997;**16**(7):548–9.
212. Warren R, Richardson M, van der Spuy G, et al. DNA fingerprinting and molecular epidemiology of tuberculosis: use and interpretation in an epidemic setting. *Electrophoresis* 1999;**20**(8):1807–12.
213. WHO. Global tuberculosis control – surveillance, planning, financing. WHO Report, WHO/HTM/TB/ 2004.331: http://www.who.int/tb/publication/global_report/en/, 2004.
214. Wilkinson D, Pillay M, Crump J, Lombard C, Davies GR, Sturm AW. Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in rural Africa. *Trop Med Int Health* 1997;**2**(8):747–53.
215. Wilson SM, Goss S, Drobniewski F. Evaluation of strategies for molecular fingerprinting for use in the routine work of a *Mycobacterium* reference unit. *J Clin Microbiol* 1998;**36**(11):3385–8.
216. Yang ZH, de Haas PE, Wachmann CH, Van Soolingen D, van Embden JD, Andersen AB. Molecular epidemiology of tuberculosis in Denmark in 1992. *J Clin Microbiol* 1995;**33**(8):2077–81.

217. Yang ZH, Mtoni I, Chonde M, et al. DNA fingerprinting and phenotyping of *Mycobacterium tuberculosis* isolates from human immunodeficiency virus (HIV)-seropositive and HIV-seronegative patients in Tanzania. *J Clin Microbiol* 1995;**33**(5):1064–9.
218. Yeh RW, Hopewell PC, Daley CL. Simultaneous infection with two strains of *Mycobacterium tuberculosis* identified by restriction fragment length polymorphism analysis. *Int J Tuberc Lung Dis* 1999;**3**(6):537–9.
219. Zozio T, Allix C, Gunal S, et al. Genotyping of *Mycobacterium tuberculosis* clinical isolates in two cities of Turkey: description of a new family of genotypes that is phylogeographically specific for Asia Minor. *BMC Microbiol* 2005;**5**:44.