Palomino • Leão • Ritacco

TUBERCULOSIS 2007

From Basic Science to Patient Care

www.TuberculosisTextbook.com

Chapter 9: Molecular Epidemiology: Breakthrough Achievements and Future Prospects

Dick van Soolingen, Kristin Kremer, and Peter W.M. Hermans

9.1. Introduction

Our understanding of the transmission of tuberculosis (TB) has been greatly enhanced since the introduction of desoxyribonucleic acid (DNA) fingerprinting techniques for *Mycobacterium tuberculosis* in the early '90s. Historical enigmas have been solved in the last decade and classical dogmas are being evaluated. This review summarizes the most important and recent findings in the molecular epidemiology of TB and discusses essential knowledge still lacking. Furthermore, current developments in the introduction of typing techniques are described, as well as future challenges to improve the usefulness of molecular markers in the epidemiology of TB.

Because the number of publications on the molecular epidemiology of TB has become too large to summarize in detail in a single review, only relatively new findings and subjects currently in the centre of attention are reviewed.

In the '90s, a wide variety of genetic markers for *M. tuberculosis* were identified (Kremer 1999). However, only a minor number of these appeared to offer enough discrimination and reproducibility for wide scale implementation (Table 9-1) (Kremer 1999, Kremer 2005a). In 1993, IS*6110* Restriction Fragment Length Polymorphism (RFLP) typing was adopted as the standard method for routine typing of *M. tuberculosis* (van Embden 1993). In this method, chromosomal DNA is digested with restriction enzyme *Pvu*II. The digested DNA is separated on an agarose gel and, after Southern Blotting, hybridized with a DNA probe. This DNA probe is directed to the IS*6110* insertion sequence and labelled with peroxidase, enabling enhanced chemiluminesence (ECL) detection of IS*6110*-containing restriction fragments (van Soolingen 1994). Another typing method, 'spoligotyping' has been used extensively as a secondary typing method (Bauer 1999, Kamerbeek 1997, Kwara 2003) and as a marker to study the phylogeny of the *M. tuberculosis* complex (Filliol 2002, Filliol 2003, Goyal 1997, Smith 2003). Spoligotyping exploits the polymorphism in the direct-repeat region of *M. tuberculosis* complex strains. This region consists of direct repeats interspersed with unique spacer sequences, and is amplified by Polymerase Chain Reaction (PCR) with primers directed to the repeats. The PCR-product is subsequently hybridized to known spacer sequences which are immobilized on a membrane through reversed-line blotting.

Because one of the primers, and hence the PCR product, is labelled with biotin, ECL detection is achieved after incubation with peroxidase-labelled streptavidin (Kamerbeek 1997). Another DNA typing method frequently used for *M. tuberculosis* is Variable Numbers of Tandem Repeats (VNTR) typing. Typing results of this method are expressed as numerical codes. Each number of the code represents the number of tandem repeats at a particular repeat locus. The number of repeats varies by strain and is determined through PCR amplification of the repeat locus with primers directed to the regions flanking that repeat locus and determination of the PCR-product size. After an extended period of improvement and validation, VNTR typing is now ready to become the next gold standard for typing of *M. tuberculosis* complex isolates (Supply 2006).

Table 9-1: Reproducibility and number of types obtained by using various DNA typing methods for differentiation of 90 *M. tuberculosis* complex strains and 10 non-*M. tuberculosis* complex mycobacterial strains (Kremer 1999, Kremer 2005a)

DNA target	Method used [a]	Reference	Reproducibility (%)[b]	No. of types obtained
IS*6110*	RFLP (*Pvu*II)	(van Soolingen 1994)	100	84
IS*6110*	Mixed-Linker PCR	(Haas 1993)	100	81
IS*6110*	FLiP	(Reisig 2005)	97	81
IS*6110*	IS*6110* inverse PCR	(Otal 1997)	6	nd [c]
IS*6110*	LM-PCR	(Prod'hom 1997)	81	73
IS*6110*/MPTR	IS*6110* ampliprinting	(Plikaytis 1993)	39	nd
IS*6110*/PGRS	DRE-PCR	(Friedman 1995)	58	63
15 loci	VNTR typing	(Supply 2006)	nd	89
12 MIRUs	VNTR typing	(Supply 2001)	100	78
ETRs A-E	VNTR typing	(Frothingham 1998)	97	56
5 QUBs[d]	VNTR typing	(Roring 2004)	87	82
DR locus	Spoligotyping	(Kamerbeek 1997)	94	61
DR locus	2nd gen. spoligotyping	(van der Zanden 2002)	90	61
DR locus	RFLP (*Alu*I)	(van Soolingen 1993)	100	48
PGRS	RFLP (*Alu*I)	(van Soolingen 1993)	100	70

(GTG)5	RFLP (*Hinf*I)	(Wiid 1994)	94	30
Total genome	APPCR	(Palittapongarnpim 1993)	71	71
4 conserved loci	Amadio PCR	(Amadio 2005)	74	13
*EcoR*I/*Mse*I sites	FAFLP typing	(Ahmed 2003)	7	nd
*EcoR*I/*Mse*I sites	FAFLP typing	(Sims 2002)	0	nd
*BamH*I/*Pst*I sites	FAFLP typing	(Kremer 2005a)	0	nd

[a] RFLP; Restriction Fragment Length Polymorphism, FLiP; Fast Ligation Mediated PCR, LM-PCR; Ligation-Mediated PCR, DRE-PCR; Double Repetitive Element PCR, VNTR; Variable Numbers of Tandem Repeats, APPCR; Arbitrarily Primed PCR, FAFLP; Fluorescent Amplified Fragment Length Polymorphism.

[b] Fraction of duplicates showing identical types (31)

[c] nd, not done

[d] Results indicated exclude QUB locus 3232

The disclosure of suitable genetic markers to study the epidemiology of infectious diseases in the last decades has led to the widespread use of a new phrase; 'molecular epidemiology'. In fact, as pointed out by Foxman (Foxman 2001), this phrase is used in many articles on DNA fingerprinting (strain typing) of bacterial isolates, regardless of the inclusion of epidemiological data. Often, the availability of bacterial isolates dictates the design of the study, and not a fundamental, relevant epidemiological question in a given area. In many published studies, microbiologists with an interest in molecular techniques were the main driving forces behind the described research. This was understandable in the initial stage of the implementation of molecular typing techniques, when the main emphasis was on the evaluation of genetic markers. However, now that the value of genetic markers for *M. tuberculosis* has become clear, it is important to involve researchers of different disciplines in the design of any molecular epidemiological study, in order to ensure the validity of the research question, the sample size, the selection of cases and the interpretation of the results.

9.2. Historical context

DNA fingerprinting of *M. tuberculosis* has been applied since the early '90s to study transmission of TB at various scales. The first report on the use of IS*986* RFLP to examine transmission of TB was published in September 1990 (Hermans 1990, McAdam 1990). Nine isolates with identical fingerprint patterns all originated from an outbreak of TB among individuals who were all treated by the same

physician, specialized in the treatment of arthritis patients. This finding led to the understanding that DNA polymorphism even in the genetically conserved *M. tuberculosis* complex isolates could be applied as a strain-specific marker. In the years thereafter, the disclosure of many other genetic makers for *M. tuberculosis* complex would follow.

Many investigators have tried to evaluate the reliability of strain typing by comparing the clustering of *M. tuberculosis* isolates based on DNA fingerprints with the findings on the respective TB patients in contact tracing. However, this was highly cumbersome, as contact tracing by interviews in itself is not at all capable of finding even a quarter of the epidemiological links between sources and follow-up cases. Thus, contact tracing cannot serve as a gold standard to evaluate DNA fingerprint results. In contrast, DNA fingerprinting seems to be a much more sensitive tool to visualize epidemiological links between cases than conventional contact tracing.

In the beginning, strain typing was mainly used to study outbreaks of TB and institutional transmission. Soon thereafter, in multiple population-based studies, the rate of recent transmission and risk factors for transmission were determined (Diel 2002, Small 1994, van Soolingen 1999). Active transmission of TB in low-prevalence settings appeared to be associated for a large part to particular risk groups such as drug abusers, homeless people, and certain immigrant groups (Diel 2002, Small 1994, van Soolingen 1999). Transmission of drug resistant bacteria could be compared to that of drug-susceptible strains (van Doorn 2006, van Soolingen 2000). These findings are discussed in Section 9.5.

With DNA fingerprinting, laboratory cross-contaminations were identified to occur at a considerable rate of 3-5 % of the positive cultures in low-prevalence settings, even though less than 10 % of the inoculated cultures were found positive in these areas (de Boer 2002, Small 1993). It is still not clear what the magnitude of this problem is in high-throughput laboratories in high-prevalence settings. Also, nosocomial infections by bacille Calmette-Guérin (BCG) have been disclosed by DNA fingerprinting and this contrasts the previous assumption that all *M. bovis* BCG infections are (late) complications of vaccination (Vos 2003b, Vos 2003a). Chemotherapeutics for the treatment of cancer patients were prepared in the same, non-disinfected biosafety cabinets that were used earlier to prepare BCG suspensions to treat bladder carcinoma patients. In this way, BCG bacteria were directly inoculated into cancer patients, in some cases with dramatic consequences.

More recently, hypotheses on the infectiousness of individual patients have also been tested (see below). Another important finding in molecular epidemiology is that exogenous re-infections after curative treatment play a much larger role than

previously anticipated (Das 1995, Sonnenberg 2001, van Rie 1999a). In the light of the description of exogenous re-infections it is interesting to read the recent observations on the detection of mixed infections (see Section 9.7). Can a part of the exogenous re-infections be explained by the initial presence of more than one strain in diagnosed TB patients?

Although *M. tuberculosis* may be one of the most widespread infectious agents in humans, not much is known about the evolution of this bacterium and whether there is an ongoing selection towards better adapted strains under the pressure of the measures introduced against TB in the last century. Because of the introduction of genetic markers for *M. tuberculosis,* the phylogeny of this bacterium can be studied in detail and the changes in the population structure can be disclosed. This has led to the recognition of a wide variety of genotype families worldwide (Bhanu 2002, Douglas 2003, Kremer 1999, Niobe-Eyangoh 2004, van Soolingen 1995, Victor 2004). In particular, the international database of spoligotyping patterns has been used most extensively for this purpose (Brudey 2006, Filliol 2002, Filliol 2003, Sola 2001).

Although it has become clear that the phylogeny of *M. tuberculosis* differs significantly in several geographic areas, not much is known about the dynamics of the population structure and the reasons for the genetic conservation observed among *M. tuberculosis* isolates in high-prevalence areas. If particular genotypes of *M. tuberculosis* are selected, how fast does a shift towards more adapted variants occur? Are we influencing the spread of particular genotypes of *M. tuberculosis*? Best studied in this respect is the Beijing genotype family of *M. tuberculosis*. There are indications that there is indeed a dramatic and relatively fast change in the composition of the worldwide population of *M. tuberculosis* (see Section 9.6). If the current observations hold true, we may be facing a recurrent TB epidemic caused by bacteria with a higher level of evolutionary development. However, more research is needed to draw better conclusions.

9.3. Infectiousness of tuberculosis patients

In most low-incidence settings, the majority of TB transmissions are limited to one or two persons. However, especially in high risk groups, such as the homeless and drug abusers in urbanized areas, ongoing transmission may take place for years and DNA fingerprint clusters sometimes grow over a hundred cases (unpublished observations in the Netherlands). In these clusters, primary, secondary, and tertiary sources can usually not be distinguished. This makes it difficult to know how many cases are derived from individual sources. DNA fingerprinting, however, has dis-

closed new information on the infectiousness of individual patients. For instance in San Francisco, 6 % of the TB cases in a two-year period seemed to have derived from a single source (Small 1994). In the Netherlands, a large outbreak in the small city of Harlingen was traced back to a single case diagnosed with a large doctor's delay (Kiers 1996, Kiers 1997).

It is only partly known what determines the transmissibility of TB. It is known that large patient- and/or doctor-originated delays play a significant role in the magnitude of transmission. Furthermore, a more extensive pulmonary process and a bad coughing hygiene clearly contribute to disease transmission. However, the bacteriological factor has not yet been established very well. It is, for instance, still not clear whether *M. tuberculosis* strains associated with large clusters on the basis of DNA fingerprinting are transmitted more easily than non-clustered strains. Is large-scale transmission only facilitated by risk factors, or do the bacterium's characteristics also contribute to a more efficient transmission and breakdown to disease?

Although there is a correlation between the smear status of a source case and the rate of transmission, smear-negative patients can also transmit TB. In San Francisco smear-negative, but culture-positive cases were found to be responsible for 17 % of the cases (Behr 1999). This indicates that smear-negative pulmonary TB suspects should be considered infectious.

9.4. DNA fingerprinting, contact investigation and source case finding

Case finding and treatment are the most important measures to inhibit the spread of TB in a community. In low-prevalence settings, where contact tracing has been routinely used for decades, a lot is known on how transmission of TB takes place. Prolonged exposure to an infectious source enhances the chance of transmission. Hence, direct and close contact with a TB patient is a main cause of infection in low-prevalence settings. However, in high-prevalence areas the transmission routes are less clear. What is the chance of acquiring an infection from an intimate contact in comparison to the chance of contracting TB from a casual contact in an environment with a high risk of infection? A recent study in South Africa (Verver 2004) pointed out that only 46 % of 313 TB patients had a matching fingerprint with an isolate of another member of the household they were living in. The proportion of transmission in the community that took place in the household was found to be only 19 %. This suggests that in this area, and presumably also in other high-incidence settings, TB transmission mainly occurs outside the household.

In settings in Western countries where the incidence of TB has become very low, the role of contact investigation remains highly important. In each area, the risk factors for the transmission of TB may differ. Factors such as being homeless, a drug abuser, living in urban areas, and low age have commonly been found to increase the risk of transmission (Borgdorff 1999, Borgdorff 2001, Diel 2002, Small 1994, van Soolingen 1999).

Usually, contact investigation is performed on the basis of the stone-in-the-pond principle and uses the Mantoux skin test (Veen 1990, Veen 1992) as an indicator of infection. Depending on the number of contacts found positive in the first ring of close contacts, the contact investigation is extended to the next ring of less intimate contacts. If again the ratio of positive contacts in that ring is high, the number is extended to the next circle of contacts. In many molecular epidemiological studies, it has been found that only a minority of the epidemiological links between TB cases disclosed by DNA fingerprinting, are also found by conventional contact tracing on the basis of interviews (Diel 2002, Lambregts-van-Weezenbeek 2003, Sebek 2000, Small 1994, van Deutekom 2004). This suggests that a large part of the TB transmission takes place through casual contacts in public places, such as bars, discothèques, public transportation, or other crowded settings. These contacts will generally not be found by interviews. Furthermore, in low incidence areas, where the skills of physicians to recognize TB adequately are waning, sources of transmission often spread the disease for extended periods and typing of isolated bacteria can help to find the source of an outbreak.

In the Netherlands, nationwide DNA fingerprinting of *M. tuberculosis* has supported contact investigations since 1993 (Lambregts-van-Weezenbeek 2003, Sebek 2000, van Soolingen 1999). All *M. tuberculosis* cultures are subjected to standardized IS*6110* RFLP typing, and clustered cases are systematically reported to the regional TB services involved (cluster feedback). In an evaluation of six years of routine DNA fingerprint surveillance, it was found that among 2,206 clustered cases, 462 (21 %) of the epidemiological links between patients were expected on the basis of contact tracing information. After cluster feedback, an additional 540 (24 %) epidemiological links were established. Epidemiological links based on documented exposure increased by 35 % (Lambregts-van-Weezenbeek 2003) (Figure 9-1).

Routine molecular typing also appears highly useful for evaluating the performance of TB control in a given area. In the Netherlands, each regional TB service quarterly receives an overview of the growth of the active-transmission clusters of patients to visualize in which populations ongoing transmission occurs and at what rate. In this way, municipal health services are able to deduce how much active

transmission is ongoing in their region. Sometimes this leads to new measures, such as active screening of particular risk groups.

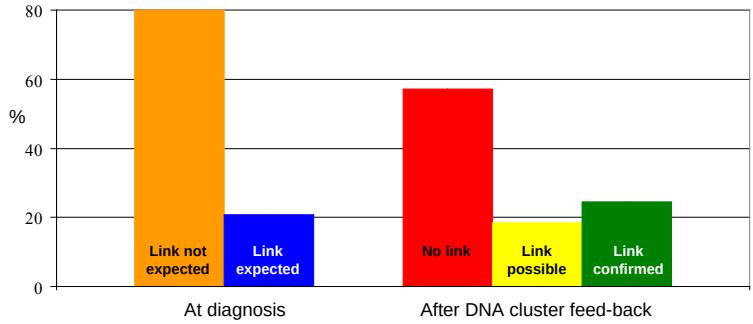

Figure 9-1: Epidemiological linkage at diagnosis and after cluster-feedback, the Netherlands 1994-2004. The bars indicate the percentage of cases with a certain level of epidemiological linkage. Source: KNCV/RIVM DNA fingerprint surveillance project.

One of the significant disadvantages of IS*6110* RFLP typing is that it requires extended culture incubation periods to obtain sufficient quantities of DNA. In the Netherlands, the typing results become available for contact tracing, on average, two months after the diagnosis of TB in a patient. At that time point, the contact investigation has usually already been finalized and not many TB services decide at that stage to re-open the contact investigations, even if the typing results provide new clues. However, the DNA fingerprint analysis clearly helps to evaluate the contact tracing process, and has therefore become an indispensable tool in TB control in the Netherlands. It is expected that the yield of molecular typing in resolving epidemiological links between patients will sharply increase when faster fingerprinting methods are implemented in the near future. In any case, nationwide molecular epidemiological analysis contributes significantly to the evaluation of contact tracing and the performance of a TB control program. It clearly indicates the rate of recent transmission and to what extent, and in which populations and areas it occurs. Figure 9-2, available at http://www.tuberculosistextbook.com/pdf/Figure 9-2.pdf, summarizes the surveillance of active transmission of TB in the Netherlands, 1997-2005.

9.5. Transmission of drug resistant tuberculosis

In a recent paper by Zignol *et al*. (Zignol 2006), the global incidence of multidrug-resistant TB (MDR-TB) was described. The estimates of the World Health Organization (WHO) on the global rate of MDR-TB have been updated from 272,906 MDR-TB cases in the year 2000 to 424,203 in the year 2004 because of the inclusion of countries that had previously not been surveyed. Zignol *et al*. underline the importance of expanding appropriate diagnostic and treatment services for MDR-TB patients, especially in countries with the highest burden of MDR-TB such as China, India, and the Russian Federation. Recently, the WHO also expressed its concern about the occurrence of extensively drug resistant (XDR) strains; *M. tuberculosis* isolates resistant to at least isoniazid (INH), rifampicin (RIF), to one of the fluoroquinolones, and to one of the injectable anti-tuberculosis drugs (Anonymous 2006). These alarming observations trigger the question; are resistant strains as transmissible as susceptible ones?

In as early as the ’50s, Mitchison observed that a large part of the INH resistant *M. tuberculosis* isolates revealed a lower degree of virulence in a guinea pig model (Mitchison 1954). For decades, it remained unclear whether resistant strains caused less transmission of TB than susceptible ones. This is important with respect to hygienic measures to prevent transmission from patients infected by MDR strains. Furthermore, for models predicting the development of the future TB epidemic, it is important to know if and how resistance interferes with transmission of TB. If resistant strains would be able to spread as efficiently as, or even better than susceptible ones, the global rates of anti-tuberculosis drug resistance would rise steadily. Indeed, transmission of highly resistant strains has been reported in, for example, New York (Bifani 1996) and South Africa (van Rie 1999b, Gandhi 2006). However, observations of transmissibility of particular (multidrug) resistant strains should not be generalized to resistance in general. In a review by Cohen *et al*., describing the effect of drug resistance on the fitness of *M. tuberculosis*, it was concluded that the fitness estimates of drug-resistant *M. tuberculosis* strains are quite heterogeneous and that this confusion makes it difficult to predict the influence of resistance on the trend of the TB epidemic (Cohen 2003). Indeed, various bacterial characteristics may influence the interference of resistance in transmissibility, including the drug susceptibility profile, the combination of mutations underlying drug resistance, presumably the genotype family the *M. tuberculosis* bacteria represent, and possibly bacterial DNA repair mechanisms. In addition, non-bacterial factors may influence the interference of resistance and transmissibility, such as the immune status of the humans exposed, and the treatment regimen ap-

plied. Because the above-mentioned factors have not been studied much, no meaningful conclusions can be drawn on the influence of the development of resistance on the worldwide TB epidemic. Yet, because of the contribution of DNA fingerprinting studies, some pieces of the puzzle have been unravelled in the last decade (van Doorn 2006, van Soolingen 2000).

In a recent study in the Netherlands, in which 8,332 patients from the period 1993-2002 were included, the drug susceptibility profiles and transmissibility of the respective isolates were studied with the aid of DNA fingerprinting (van Doorn 2006). In total, 592 isolates were resistant to INH, of which 323 carried a mutation at amino acid position 315 (Δ315) of the catalase-peroxidase gene (*katG*). The remaining INH resistant strains had other mechanisms underlying INH resistance. As predicted by Mitchison (Mitchison 1954), in general INH resistant strains were less transmissible (i.e. less frequently present in DNA fingerprint clusters) than susceptible ones. However, strains with the Δ315 were as frequently part of active transmission as susceptible ones. Moreover, the INH resistant strains with the Δ315 had a higher level of INH resistance and were associated with multidrug resistance (van Doorn 2006, van Soolingen 2000). This suggests that the type of genetic mutation underlying INH resistance is an important factor in the fitness of the bacterium. Thus, particular strains may be the cause of MDR-TB transmission in both high and low-incidence settings, even though INH resistant strains in general are less fit than susceptible ones. In South Africa, most of the childhood contacts of adults with MDR-TB were more likely to be infected from these than other (drug susceptible) TB sources (Schaaf 2000). It would be highly interesting to know the mutations underlying resistance in these cases.

In the Netherlands, transmission of MDR-TB is usually limited to incidental single person-to-person transmission. However, in the period 2003/2004 a single MDR-TB case infected nine other persons, of which two developed active disease. The respective MDR-TB strain had a mutation at amino acid position 315 of *katG* and exceptional mutations underlying RIF resistance (unpublished observations). It is not clear whether this type of resistant variant influences the epidemiology of TB in low and high-incidence areas. Therefore, further, more detailed and representative investigations into the basis of resistance in combination with the behaviour of the bacterium are needed.

9.6. Resistance and the Beijing genotype

Another important factor that may determine the transmissibility of resistant strains is the genetic background of the bacterium. Based on several genetic markers, various *M. tuberculosis* genotype families have been identified, such as the Beijing family (van Soolingen 1995), the Haarlem family (Kremer 1999), Family 11 (Victor 2004), the Manila family (Douglas 2003), the Delhi family (Bhanu 2002), the Cameroon family (Niobe-Eyangoh 2004), the Latin American Mediterranean (LAM) family, the Central Asian clade, and the East African Indian clade (Brudey 2006, Filliol 2002, Filliol 2003, Sola 2001). It is important to study genotypic and phenotypic characteristics of the genotype families that fuel the worldwide TB epidemic. Up until now, the Beijing genotype has been studied most extensively. The Beijing genotype was first described in 1995 (van Soolingen 1995), and strains belonging to this genotype family appeared to be genetically highly conserved, which suggests that the spread of these strains started relatively recently. Moreover, in several areas, Beijing genotype strains are more frequently isolated from young patients than from older patients (Anh 2000, Borgdorff 2003, Glynn 2006). If, in high incidence areas, active transmission of TB is associated with lower age of the patients, as it is in low incidence settings (van Soolingen 1999), this suggests that Beijing genotype strains are emerging. The fact that Beijing strains have more often been found recently where population-based molecular epidemiological studies have been ongoing for several years points in that direction (Borgdorff 2003, Glynn 2006). Furthermore, the Beijing strains are associated with drug resistance in some areas (Glynn 2002, Glynn 2006). Thus, strains of the Beijing family may have a genetic background that favours their transmission, despite their drug resistance.In 2006, a large worldwide survey was published on the spread of the Beijing genotype of *M. tuberculosis* and its association with drug resistance (Glynn 2006). In this study, which included 29,259 patients from 35 countries, the overall prevalence of Beijing strains was 9.9 %, and the proportion of TB due to the Beijing genotype ranged from 0 % to over 72.5 % per area. The Beijing genotype was endemic in East Asia and parts of the USA. In Cuba, the former Soviet Union, Vietnam, South Africa, and in parts of Western Europe this genotype was epidemic and associated with drug resistance (Glynn 2006).

Previously, in New York outbreaks of MDR-TB were also caused by one of the evolutionary branches of the Beijing genotype family; the W strains (Bifani 1996, Kurepina 1998). The W strains, however, are a relatively minor branch on the evolutionary tree of the Beijing genotype family.

It is to be determined to what extent the worldwide prevalence of MDR-TB is influenced by the success of particular genotype families of *M. tuberculosis* in absolute terms, such as the Beijing strains. It is at least striking that in many areas with a high rate of MDR-TB, the Beijing strains are also highly prevalent (Glynn 2006, Kruuner 2001, Pfyffer 2001, World Health Organization 2004, Zignol 2006). It has yet to be determined whether there is a causal correlation between these observations.

It remains unclear whether transmission of highly resistant strains in high incidence settings are exceptions to the rule that resistance in general costs fitness of the bacterium, or that particular genotypes of *M. tuberculosis* have developed efficient ways to become resistant to anti-tuberculosis drugs and maintain or even increase their ability to spread in a community. In the latter case, these genotypes will spread in the coming years and will influence the development of the worldwide TB epidemic.

9.7. Genetic heterogeneity of *M. tuberculosis* and multiple infections

When talking about multiple *M. tuberculosis* sub-populations in sputum of TB patients, two phenomena are often confused, although they should be clearly distinguished:

- multiple strain populations derived from a single ancestral strain displaying genetic drift
- multiple infections by more than one strain.

In the case of multiple (or mixed) infections, the presence of more than one *M. tuberculosis* strain is demonstrated on one occasion of culturing from clinical material. This should not be confused with re-infection, usually after curative treatment, as this refers to a new episode of the disease caused by another strain. In South Africa, where the prevalence of TB is very high, the contribution of re-infection to new episodes of TB after curative treatment is considerable, and has been estimated at 75 % (van Rie 1999a, Verver 2005).

Numerous observations in the molecular epidemiology of TB have pointed out that bacteria are subject to evolutionary change. Sometimes minor rearrangements of IS*6110* RFLP profiles are noticed in epidemiologically related- and serial patient isolates. The rate of change of IS*6110* RFLP patterns in such isolates has been studied by several investigators (de Boer 1999, Niemann 1999, Niemann 2000, Yeh

1998). However, also within clinical *M. tuberculosis* isolates, sub-populations of bacteria with minor genomic differences co-exist (de Boer 2000, Shamputa 2004, Shamputa 2006). For example, low-intensity bands in IS*6110* RFLP profiles are a reliable indication of a sub-population of bacteria with, for example, a one-band difference in IS*6110* RFLP. Preparation of single colony cultures and subsequent IS*6110* RFLP typing of isolates with such low-intensity bands showed the co-existence of separate sub-populations of bacteria, either with or without a normal-intensity band at the position where the low-intensity band occurred in the original clinical isolate (de Boer 2000).

Several recent papers describe the finding of multiple *M. tuberculosis* populations in sputum specimens of TB patients (Richardson 2002, Shamputa 2004, Shamputa 2006, van Rie 2005, Warren 2004). These findings point out that multiple infection of *M. tuberculosis* may be more prevalent than previously assumed. In the study by Warren *et al.*, a PCR technique was used to specifically identify *M. tuberculosis* bacteria of the Beijing genotype family and other evolutionary lineages in sputum specimens of patients from South Africa (Warren 2004). These authors concluded that at least 19 % of the patients included were infected by both Beijing and non-Beijing strains. Multiple infections were more frequently observed in re-treatment cases than in new cases. The same group also explored IS*6110* RFLP typing to detect multiple strain infections; a minor part of the IS*6110* RFLP patterns exhibited background patterns suggestive of mixed infections (Richardson 2002). This was confirmed in three (2.3 %) of the cases. In addition, another interesting approach was followed to study the occurrence of multiple infections in TB patients; by investigating *M. tuberculosis* strain diversity in autopsy material in South Africa (Plessis, 2001). In two out of 12 patients, pulmonary infection by two strains was demonstrated. The question remains about how this relates to the practical bacteriology: if this study had been performed at the time of diagnosis of TB in these patients, would one or two strains have been isolated from the sputum? Is the presence of multiple strains in autopsy material related to time-spaced infections, and do they represent re-infections? Is it possible that *M. tuberculosis* bacilli from a first infection are present in the body in a dormant state, and that a super-infection can lead to disease caused by the second infection without reactivation of the dormant bacteria? Therefore, although it is now clear that mixed infections do occur in TB patients, more research is needed to understand this phenomenon.

In the Netherlands, where from 1993 to 2006 about 15,000 *M. tuberculosis* isolates (of which 60 % were derived from patients from high-incidence regions) were subjected to IS*6110* RFLP analysis, double IS*6110* RFLP patterns were observed on only one occasion (de Boer 2000). During an episode of laboratory cross-

contamination in a peripheral laboratory in the Netherlands, clearly the RFLP pattern of the control strain was present as a background pattern in several isolates originating from that laboratory (Van Duin 1998). No other double IS*6110* RFLP patterns with different intensities were observed in any of the typing results. However, the sensitivity of IS*6110* RFLP typing to detect multiple infections is limited; at least 10 % of the DNA of a tested strain needs to be from another strain to be able to see this as a low-intensity, background pattern (de Boer 2000). Thus, multiple infections probably occur more often.

In the study by Shamputa *et al*., the clonality of 97 *M. tuberculosis* isolates was analyzed by first preparing a limited number (mostly 10) of single colony cultures and analyzing them by IS*6110* RFLP typing, spoligotyping, and VNTR analysis (Shamputa 2004). Different subpopulations of bacteria, including the ones representing evolutionary drift, were found in eight (8.2 %) of the isolates, while the frequency of confirmed mixed infections by different strains was 2.1 % (Shamputa 2004). In this study, it was found that the predominant strains and the primary isolates always had concordant drug susceptibility profiles, which suggests that the practical implications for the treatment of the respective cases were limited. However, in the study by Van Rie *et al.,* it was reported that re-infection and mixed infection do cause changes in drug susceptibility patterns of *M. tuberculosis* isolates and that treatment with second-line drugs may lead to re-emergence of drug-susceptible strains in patients with mixed infections (van Rie 2005). If mixed infections are common in high prevalence settings, this may be of concern for the clinician, as pointed out by Behr (Behr 2004); it may be that drug-resistant bacteria are not detected and cause a relapse after an apparent 'curative' treatment. With the current knowledge, such a case would probably be classified as exogenous re-infection, because no representative studies have been undertaken to combine investigations on mixed infections during the first episode of the disease and the presentation of relapses after treatment in the same patients.

Although the study by Shamputa *et al.* (Shamputa 2004) is so far the most extensive study on this subject published so far, one has to realize that the analysis of 10 colonies of a primary isolate is a very limited number. The chance of detecting a mixed infection is limited by the ratio of the strain variants in the isolates and the coincidence of picking the right colonies. When the ratio of a mixture is 1:1, 5 colonies need to be analyzed to identify both strains with a 95 % confidence interval. However, if the ratio of the mixture is 1:10, 29 colonies should be analyzed to detect a mixture with the same reliability. The ratio of mixed infections may be much less balanced in clinical samples; particular strains may predominate over other strains with a ratio of 1:100, 1:1,000, or even less.

It is also not clear whether individuals suffering from mixed infections (or re-infections) constitute a human population hypersensitive to *M. tuberculosis* infections with regard to their immunological and/or genetic background. More studies focusing on the immunological aspects and genetic predispositions possibly associated with re-infections would be highly interesting.

So far, only anecdotal observations on mixed infections have been reported. However, the current observations of mixed and re-infections in any case merit more representative studies to determine the magnitude of this problem. To critically evaluate the results and to check for possible laboratory cross-contamination, at least two culture-positive clinical samples should be analyzed.

9.8. The new standard genetic marker: VNTR typing

IS*6110* RFLP typing (van Embden 1993) has gained recognition as the gold standard in the molecular epidemiology of TB since 1993. However, this method is technically demanding and labor intensive, requires weeks of incubation for culturing of the isolates to obtain sufficient quantities of DNA, and suffers from problems of interpretability and portability of the complex banding patterns. In addition, it provides insufficient discrimination among isolates with a low number of IS*6110* copies (< 6); a problem that is only partly overcome by using additional typing methods, such as spoligotyping (Cowan 2005). Variable Number of Tandem Repeats (VNTR) typing is increasingly used to solve these problems (Frothingham 1998, Le Fleche 2002, Roring 2002, Skuce 2002, Smittipat 2000, Supply 1997, Supply 2000). This method is based on PCR amplification of multiple repeat loci, using primers specific for the flanking regions of each locus and on the determination of the sizes of the PCR products. The sizes of the amplicons reflect the number of tandem repeats present at the respective loci. Sizing can be done using a capillary system (Allix 2004, Kwara 2003, Supply 2001), gel electrophoresis (Mazars 2001), or non-denaturing high performance liquid chromatography (Evans 2004).

VNTR typing is considerably faster than IS*6110* RFLP typing, as it is applicable to crude low-concentration DNA extracts from early mycobacterial cultures. Furthermore, it has been adapted to high throughput format (Allix 2004, Kwara 2003, Supply 2001). Moreover, the results are expressed as numerical codes and are therefore easy to compare and exchange.

Currently, VNTR typing is often based on 12 Mycobacterial Interspersed Repetitive Units (MIRU) loci (Mazars 2001, Supply 2000) and has been integrated in TB control systems on a national scale in, for example, the USA (Cowan 2005). Based

on pilot studies with limited numbers of isolates, the discriminatory power of this 12 loci VNTR set approached that of IS*6110* RFLP typing to discriminate epidemiologically unrelated cases (Mazars 2001, Supply 2001), while VNTR types were stable among isolates from epidemiologically linked cases (Hawkey 2003, Kwara 2003, Savine 2002). A recent population-based study indicated that the use of this 12-loci method as a first-line screening in combination with spoligotyping provides adequate discrimination in most cases for large-scale, prospective genotyping of *M. tuberculosis* in the United States. However, IS*6110* fingerprinting is still required as an additional method to type the clustered isolates in a number of cases, when contact investigation, demographic or epidemiological data do not provide independent clues on the existence or the absence of links between patients (Cowan 2005).

Alternative sets of VNTR loci have been suggested to further improve the discrimination of unrelated isolates, as compared to that provided by this 12-loci system (Kam 2006, Kremer 2005b, Le Fleche 2002, Roring 2002, Roring 2004, Skuce 2002, Smittipat 2005, Surikova 2005). However, the collections of isolates studied were restricted to small samples of local origin and/or included only *M. bovis*, or representatives of only one or two of the defined *M. tuberculosis* lineages. The overall technical robustness and the clonal stability of the individual VNTR loci in the sets tested were not assessed. Furthermore, none of these studies were based on non-selected, population-based samples, and contact tracing data was not available, making it impossible to establish the predictive value of the various VNTR sets for studying ongoing *M. tuberculosis* transmission at a population-based level.

Recently, in an international collaboration, the resolution, stability and technical applicability of 29 VNTR loci was compared (Supply 2006). This study comprised the initial 12 loci and most of the other loci disclosed so far. The typing results of 824 *M. tuberculosis* isolates, including worldwide representatives of the main *M. tuberculosis* lineages, as well as multiple groups of epidemiologically linked or clonal isolates, revealed the 24 most optimal VNTR loci. Locus designations and PCR primer sequences for the 24-loci VNTR typing method are available in Table 9-2 at http://www.tuberculosistextbook.com/pdf/Table 9-2.pdf (Supply 2006).

Based on redundancy analysis, a highly discriminatory subset of 15 loci was selected for first-line epidemiological investigations. The use of these 15 VNTR loci was proposed as the new international standard for typing of *M. tuberculosis* complex isolates (Supply 2006). Extension to the use of 24 loci is especially useful in studying the phylogeny of strains.

As experienced after the standardization of IS*6110* RFLP typing in 1993 (van Embden 1993), it is expected that the international consensus on VNTR typing will facilitate the comparison of molecular epidemiological data from different geographical regions. The establishment of international VNTR databases and the meta-analysis of worldwide typing results will facilitate further study of the population structure of *M. tuberculosis*.

Many institutes in the world have large databases containing high numbers of IS*6110* RFLP patterns of *M. tuberculosis* isolates from extended periods, and are considering a switch from IS*6110* RFLP to VNTR typing. Because the epidemiology of TB demands the consideration of contacts with sources separated in time by years, the switch to VNTR typing cannot be done without any overlap of the use of the two typing methods. In order to trace transmission patterns retrospectively, it would be best to re-type all *M. tuberculosis* isolates from a number of years by VNTR typing. If this is too costly or time demanding, it could be considered to limit re-typing activities to strains from a more limited retrospective period; for instance three years. In addition, the typing of only one isolate from each IS*6110* RFLP cluster could reduce the re-typing workload significantly. If resistance issues play a role in the concerned setting, the re-typing could be restricted to resistant *M. tuberculosis* isolates. Alternatively, it could be considered to define an age limit for the re-typing activities, because active transmission mainly takes place through younger individuals (at least in low prevalence settings where this has been studied extensively) (Borgdorff 1999, van Soolingen 1999). However, these alternative approaches of re-typing, which do not include all isolates, will conceal a part of the VNTR polymorphism among the circulating isolates.

9.9. DNA fingerprinting to monitor eradication of tuberculosis

DNA fingerprinting may also be useful for studying the stage of the TB epidemic and to predict the future developments. In a recent study in the Netherlands, covering the period of 1993-2002, changes in TB transmission were determined using DNA fingerprinting to assess the progress towards TB elimination (Borgdorff 2005). Strains were defined as 'new' if their DNA fingerprint pattern had not been observed in any other patient during the previous two years. Other cases were defined as clustered and attributed to recent transmission. The incidence of TB cases involving new strains was stable among the non-Dutch and declined among Dutch nationals. However, the decline among the Dutch cases was restricted to those aged 65 years and over. It was concluded that the decline of TB in the Netherlands over the past decade is therefore mainly the result of a cohort effect: those with lower

infection prevalence replaced older birth cohorts with high infection prevalence. It is expected that TB will not be eliminated in the Netherlands in the near future, mainly because of the contact with high-burden countries through immigrants and international travel (Borgdorff 2005, Cobelens 2000).

9.10. Future prospects

Although the introduction of molecular markers for *M. tuberculosis* in the early '90s has greatly facilitated our understanding of the epidemiology of TB, even the latest most optimal typing, VNTR typing, will not be completely reliable. In fact, each genetic marker only reveals a minor part of the genomic information of a bacterium. Depending on the marker, different strains will exhibit identical genotyping profiles. Furthermore, in order to be able to follow the chains of transmission in a given area and to subdivide primary, secondary, etc. sources of infection, the turnover of genotyping profiles will never be in range with the pace of transmission.

To distinguish between even genetically related strains, and to be able to follow the spread of offspring of strains in the community, more detailed multiple-marker typing systems need to be developed. In fact, the most accurate typing would be whole genome sequence analysis of *M. tuberculosis* isolates. It is expected that with this information, the exact sequence in the evolutionary development of the offspring of a *M. tuberculosis* bacterium can be identified, without the interference of differences in time of incubation and confusion about the spread from primary and secondary, or even tertiary sources of infection in the same period. The current developments in DNA sequence techniques (Bennett 2005, Margulies 2005) provide possibilities to test these expectations and will provide more accurate predictions on transmission of TB. A largely unrecognized problem that has to be dealt with in due time is the occurrence of multiple (mixed) infections in high incidence settings (Shamputa 2004, Shamputa 2006, van Rie 2005, Warren 2004). This may hamper molecular studies on transmission severely.

Furthermore, the evolution of bacteria does not take place through whole population shifts in the genomic make up, but through mutation and multiplication of initially a single bacterium. By applying the current DNA amplification and sequence techniques, subtle genetic changes among bacterial strains are still difficult to visualize. However, creative future solutions may also deal with this phenomenon.

References

1. Ahmed N, Alam M, Abdul-Majeed A, et al. Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the *Mycobacterium tuberculosis* complex. Infect Genet Evol 2003; 2: 193-9.
2. Allix C, Supply P, Fauville-Dufaux M. Utility of fast mycobacterial interspersed repetitive unit-variable number tandem repeat genotyping in clinical mycobacteriological analysis. Clin Infect Dis 2004; 39: 783-9.
3. Amadio A, Romano MI, Bigi F, et al. Identification and characterization of genomic variations between *Mycobacterium bovis* and *M. tuberculosis* H37Rv. J Clin Microbiol 2005; 43: 2481-4.
4. Anh DD, Borgdorff MW, Van LN, et al. *Mycobacterium tuberculosis* Beijing genotype emerging in Vietnam. Emerg Infect Dis 2000; 6: 302-5.
5. Anonymous. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs--worldwide, 2000-2004. MMWR Morb Mortal Wkly Rep 2006; 55: 301-5.
6. Bauer J, Andersen AB, Kremer K, Miorner H. Usefulness of spoligotyping To discriminate IS*6110* low-copy-number *Mycobacterium tuberculosis* complex strains cultured in Denmark. J Clin Microbiol 1999; 37: 2602-6.
7. Behr MA. Tuberculosis due to multiple strains: a concern for the patient? A concern for tuberculosis control? Am J Respir Crit Care Med 2004; 169: 554-5.
8. Behr MA, Warren SA, Salamon H, et al. Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. Lancet 1999; 353: 444-9.
9. Bennett ST, Barned C, Cox A, Davies L, Brown C. Towards the 1,000 dollars human genome. Pharmacogenomics 2005; 6:373-82.
10. Bhanu NV, van Soolingen D, van Embden JD, Dar L, Pandey RM, Seth P. Predominace of a novel *Mycobacterium tuberculosis* genotype in the Delhi region of India. Tuberculosis Edinburgh, Scotland 2002; 82: 105-12.
11. 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: 452-7.
12. Borgdorff MW, de Haas P, Kremer K, van Soolingen D. *Mycobacterium tubercuiosis* Beijing genotype, the Netherlands. Emerg Infect Dis 2003; 9: 1310-3.
13. Borgdorff MW, Nagelkerke NJ, van Soolingen D, Broekmans JF. Transmission of tuberculosis between people of different ages in The Netherlands: an analysis using DNA fingerprinting. Int J Tuberc Lung Dis 1999; 3: 202-6.
14. Borgdorff MW, Nagelkerke NJD, de Haas PEW, van Soolingen D. Transmission of *Mycobacterium tuberculosis* depending on the age and sex of source cases. Amer J Epidemiol 2001; 154: 934-43.
15. Borgdorff MW, van der Werf M, de Haas PEW, Kremer K, van Soolingen D. Prospects for tuberculosis elimination in The Netherlands: a molecular epidemiologic analysis, 1993 through 2002. Emerg Infect Dis 2005; 11: 597-602.
16. Brudey K, Driscoll JR, Rigouts L, et al. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol 2006; 6: 23.
17. Cobelens FG, van Deutekom H, Draayer-Jansen IW, et al. Risk of infection with *Mycobacterium tuberculosis* in travellers to areas of high tuberculosis endemicity. Lancet 2000; 356: 461-5.

18. Cohen T, Sommers B, Murray M. The effect of drug resistance on the fitness of *Mycobacterium tuberculosis*. Lancet Infect Dis 2003; 3: 13-21.

19. 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: 688-95.

20. Das S, Paramasivan CN, Lowrie DB, Prabhakar R, Narayanan PR. IS*6110* restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, south India. Tuber Lung Dis 1995; 76: 550-4.

21. de Boer AS, Blommerde B, de Haas PEW, et al. False-positive *Mycobacterium tuberculosis* cultures in 44 laboratories in The Netherlands (1993-2000): incidence, risk factors, and consequences. J Clin Microbiol 2002; 40: 4004-9.

22. de Boer AS, Borgdorff MW, de Haas PE, Nagelkerke NJ, van Embden JD, van Soolingen D. Analysis of rate of change of IS*6110* RFLP patterns of *Mycobacterium tuberculosis* based on serial patient isolates. J Infect Dis 1999; 180: 1238-44.

23. de Boer AS, Kremer K, Borgdorff MW, de Haas PEW, Heersma HF, van Soolingen D. Genetic heterogeneity in *Mycobacterium tuberculosis* isolates reflected IS*6110* restriction fragment length polymorphism patterns as low-intensity bands. J Clin Microbiol 2000; 38: 4478-84.

24. Diel R, Schneider S, Meywald-Walter K, Ruf CM, Rusch-Gerdes S, Niemann S. Epidemiology of tuberculosis in Hamburg, Germany: long-term population-based analysis applying classical and molecular epidemiological techniques. J Clin Microbiol 2002; 40: 532-9.

25. Douglas JT, Qian LS, Montoya JC, et al. Characterization of the Manila family of *Mycobacterium tuberculosis*. J Clin Microbiol 2003; 41: 2723-6.

26. Evans JT, Hawkey PM, Smith EG, Boese KA, Warren RE, Hong G. Automated high-throughput mycobacterial interspersed repetitive unit typing of *Mycobacterium tuberculosis* strains by a combination of PCR and nondenaturing high-performance liquid chromatography. J Clin Microbiol 2004; 42: 4175-80.

27. Filliol I, Driscoll JR, van Soolingen D, et al. Global distribution of *Mycobacterium tuberculosis* spoligotypes. Emerg Infect Dis 2002; 8: 1347-9.

28. Filliol I, Driscoll JR, van Soolingen D, et al. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assesed by spoligotyping in an international study. J Clin Microbiol 2003; 41: 1963-70.

29. Foxman B, Riley L. Molecular epidemiology: focus on infection. Am J Epidemiol 2001; 153: 1135-41.

30. Friedman CR, Stoeckle MY, Johnson-WD J, Riley LW. Double-repetitive-element PCR method for subtyping *Mycobacterium tuberculosis* clinical isolates. J Clin Microbiol 1995; 33: 1383-4.

31. Frothingham R, Meeker-O'Connell WA. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. Microbiology 1998; 144: 1189-96.

32. Gandhi NR, Moll A, Sturm AW, et al. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. Lancet 2006; 368:1575-80.

33. Glynn JR, Kremer K, Borgdorff M, Rodrigues MP, van Soolingen D, the European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis. Beijing/W genotype *Mycobacterium tuberculosis* and drug resistance. Emerg Infect Dis 2006; 12: 736-43.

34. 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: 843-9.

35. Goyal M, Saunders NA, van Embden JD, Young DB, Shaw RJ. Differentiation of *Mycobacterium tuberculosis* isolates by spoligotyping and IS*6110* restriction fragment length polymorphism. J Clin Microbiol 1997; 35: 647-51.

36. Haas WH, Butler WR, Woodley CL, Crawford JT. Mixed-linker polymerase chain reaction: a new method for rapid fingerprinting of isolates of the *Mycobacterium tuberculosis* complex. J Clin Microbiol 1993; 31: 1293-8.

37. Hawkey PM, Smith EG, Evans JT, et al. Mycobacterial interspersed repetitive unit typing of *Mycobacterium tuberculosis* compared to IS*6110*-based restriction fragment length polymorphism analysis for investigation of apparently clustered cases of tuberculosis. J Clin Microbiol 2003; 41: 3514-20.

38. Hermans PW, van Soolingen D, Dale JW, et al. Insertion element IS*986* from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. J Clin Microbiol 1990; 28: 2051-8.

39. Kam KM, Yip CW, Tse LW, et al. Optimization of variable number tandem repeat typing set for differentiating *Mycobacterium tuberculosis* strains in the Beijing family. FEMS Microbiol Lett 2006; 256: 258-65.

40. 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: 907-14.

41. Kiers A, Drost AP, van Soolingen D, Veen J. Border-crossing source tracing in tuberculosis via DNA fingerprint technique. [Grensoverschrijdende bronopsporing bij tuberculose door DNA-'fingerprint'-techniek]. Ned Tijdschr Geneeskd 1996; 140: 2290-3.

42. Kiers A, Drost AP, van Soolingen D, Veen J. Use of DNA fingerprinting in international source case finding during a large outbreak of tuberculosis in The Netherlands. Int J Tuberc Lung Dis 1997; 1: 239-45.

43. Kremer K, Arnold C, Cataldi A, et al. Discriminatory power and reproducibility of novel DNA typing methods for *Mycobacterium tuberculosis* complex strains. J Clin Microbiol 2005a; 43: 5628-38.

44. 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 2005b; 43: 314-20.

45. 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: 2607-18.

46. Kruuner A, Hoffner SE, Sillastu H, et al. Spread of drug-resistant pulmonary tuberculosis in Estonia. J Clin Microbiol 2001; 39: 3339-45.

47. Kurepina NE, Sreevatsan S, Plikaytis BB, et al. Characterization of the phylogenic distribution and chromosomal insertion sites of five IS*6110* elements in *Mycobacterium tuberculosis:* non-random integration in the *dnaA-dnaN* region. Tuber Lung Dis 1998; 79: 31-42.

48. Kwara A, Schiro R, Cowan LS, et al. Evaluation of the epidemiologic utility of secondary typing methods for differentiation of *Mycobacterium tuberculosis* isolates. J Clin Microbiol 2003; 41: 2683-5.

49. Lambregts-van-Weezenbeek CS, Sebek MM, van Gerven PJ, et al. Tuberculosis contact investigation and DNA fingerprint surveillance in The Netherlands: 6 years' experience with nation-wide cluster feedback and cluster monitoring. Int J Tuberc Lung Dis 2003; 7: S463-S470.
50. Le Fleche P, Fabre M, Denoeud F, Koeck JL, Vergnaud G. High resolution, on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing. BMC Microbiol 2002; 2: 37.
51. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005; 437: 376-80.
52. 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 U S A 2001; 98: 1901-6.
53. McAdam RA, Hermans PW, van Soolingen D, et al. Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS*3* family. Mol Microbiol 1990; 4: 1607-13.
54. Mitchison DA. Tubercle bacilli resistant to isoniazid; virulence and response to treatment with isoniazid in guinea-pigs. BMJ 1954; 1: 128-30.
55. Niemann S, Richter E, Rusch-Gerdes S. Stability of *Mycobacterium tuberculosis* IS*6110* restriction fragment length polymorphism patterns and spoligotypes determined by analyzing serial isolates from patients with drug-resistant tuberculosis. J Clin Microbiol 1999; 37: 409-12.
56. Niemann S, Rusch-Gerdes S, Richter E, Thielen H, Heykes-Uden H, Diel R. Stability of IS*6110* restriction fragment length polymorphism patterns of *Mycobacterium tuberculosis* strains in actual chains of transmission. J Clin Microbiol 2000; 38: 2563-7.
57. Niobe-Eyangoh SN, Kuaban C, Sorlin P, Thonnon J, Vincent V, Gutierrez MC. Molecular characteristics of strains of the cameroon family, the major group of *Mycobacterium tuberculosis* in a country with a high prevalence of tuberculosis. J Clin Microbiol 2004; 42: 5029-35.
58. Otal I, Samper S, Asensio MP, et al. Use of a PCR method based on IS*6110* polymorphism for typing *Mycobacterium tuberculosis* strains from BACTEC cultures. J Clin Microbiol 1997; 35: 273-7.
59. Palittapongarnpim P, Chomyc S, Fanning A, Kunimoto D. DNA fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates by arbitrarily primed polymerase chain reaction. J Infect Dis 1993; 167: 975-8.
60. Pfyffer GE, Strassle A, van Gorkum T, et al. Multidrug-resistant tuberculosis in prison inmates, Azerbaijan. Emerg Infect Dis 2001; 7: 855-61.
61. Plikaytis BB, Crawford JT, Woodley CL, et al. Rapid, amplification-based fingerprinting of *Mycobacterium tuberculosis*. J Gen Microbiol 1993; 139: 1537-42.
62. Prod'hom G, Guilhot C, Gutierrez MC, Varnerot A, Gicquel B, Vincent V. Rapid discrimination of *Mycobacterium tuberculosis* complex strains by ligation-mediated PCR fingerprint analysis. J Clin Microbiol 1997; 35: 3331-4.
63. Reisig F, Kremer K, Amthor B, van Soolingen D, Haas WH. Fast ligation-mediated PCR, a fast and reliable method for IS*6110*-based typing of *Mycobacterium tuberculosis* complex. J Clin Microbiol 2005; 43: 5622-7.
64. Richardson M, Carroll NM, Engelke E, et al. Multiple *Mycobacterium tuberculosis* strains in early cultures from patients in a high-incidence community setting. J Clin Microbiol 2002; 40: 2750-4.

65. Roring S, Scott A, Brittain D, et al. Development of variable-number tandem repeat typing of *Mycobacterium bovis*: comparison of results with those obtained by using existing exact tandem repeats and spoligotyping. J Clin Microbiol 2002; 40: 2126-33.

66. Roring S, Scott AN, Glyn-Hewinson R, Neill SD, Skuce RA. Evaluation of variable number tandem repeat (VNTR) loci in molecular typing of *Mycobacterium bovis* isolates from Ireland. Vet Microbiol 2004; 101: 65-73.

67. Savine E, Warren RM, van der Spuy GD, et al. Stability of variable-number tandem repeats of mycobacterial interspersed repetitive units from 12 loci in serial isolates of *Mycobacterium tuberculosis*. J Clin Microbiol 2002; 40: 4561-6.

68. Schaaf HS, van Rie A, Gie RP, et al. Transmission of multidrug-resistant tuberculosis. Pediatr Infect Dis J 2000; 19: 695-9.

69. Sebek M. DNA fingerprinting and contact investigation. Int J Tuberc Lung Dis 2000; 4: S45-S48.

70. Shamputa IC, Jugheli L, Sadradze N, et al. Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia. Respir Res 2006; 7: 99.

71. 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: 5528-36.

72. Sims EJ, Goyal M, Arnold C. Experimental versus in silico fluorescent amplified fragment length polymorphism analysis of *Mycobacterium tuberculosis*: improved typing with an extended fragment range. J Clin Microbiol 2002; 40: 4072-6.

73. Skuce RA, McCorry TP, McCarroll JF, et al. Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. Microbiology Sgm 2002; 148 Part 2: 519-28.

74. 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: 1703-9.

75. Small PM, McClenny NB, Singh SP, Schoolnik GK, Tompkins LS, Mickelsen PA. Molecular strain typing of *Mycobacterium tuberculosis* to confirm cross-contamination in the mycobacteriology laboratory and modification of procedures to minimize occurrence of false-positive cultures. J Clin Microbiol 1993; 31: 1677-82.

76. Smith NH, Dale J, Inwald J, et al. The population structure of *Mycobacterium bovis* in Great Brittain: clonal expansion. Proc Natl Acad Sci U S A 2003; 100: 15271-5.

77. Smittipat N, Billamas P, Palittapongarnpim M, et al. Polymorphism of variable-number tandem repeats at multiple loci in *Mycobacterium tuberculosis*. J Clin Microbiol 2005; 43: 5034-43.

78. Smittipat N, Palittapongarnpim P. Identification of possible loci of variable number of tandem repeats in *Mycobacterium tuberculosis*. Tuber Lung Dis 2000; 80: 69-74.

79. Sola C, Filliol I, Gutierrez MC, Mokrousov I, Vincent V, Rastogi N. Spoligotype database of *Mycobacterium tuberculosis*: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives. Emerg Infect Dis 2001; 7: 390-6.

80. Sonnenberg P, Murray J, Glynn JR, Shearer S, Kambashi B, Godfrey FP. HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers. Lancet 2001; 358: 1687-93.

81. Supply P, Allix C, Lesjean S, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J Clin Microbiol 2006; 44: 4498-510.

82. Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J Clin Microbiol 2001; 39: 3563-71.

83. 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: 991-1003.

84. 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: 762-71.

85. Surikova OV, Voitech DS, Kuzmicheva G, et al. Efficient differentiation of *Mycobacterium tuberculosis* strains of the W-Beijing family from Russia using highly polymorphic VNTR loci. Eur J Epidemiol 2005; 20: 963-74.

86. 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: 4628-39.

87. 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: 806-10.

88. van Doorn HR, de Haas PE, Kremer K, Vandenbroucke-Grauls CM, Borgdorff MW, van Soolingen D. Public health impact of isoniazid-resistant *Mycobacterium tuberculosis* strains with a mutation at amino-acid position 315 of katG: a decade of experience in The Netherlands. Clin Microbiol Infect 2006; 12: 769-75.

89. Van Duin JM, Pijnenburg JE, van Rijswoud CM, de Haas PE, Hendriks WD, van Soolingen D. Investigation of cross contamination in a *Mycobacterium tuberculosis* laboratory using IS*6110* DNA fingerprinting. Int J Tuberc Lung Dis 1998; 2: 425-9.

90. van Embden JD, Cave MD, Crawford JT, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol 1993; 31: 406-9.

91. van Rie A, Victor TC, Richardson M, et al. Reinfection and mixed infection cause changing *Mycobacterium tuberculosis* drug-resistance patterns. Am J Respir Crit Care Med 2005; 172: 636-42.

92. van Rie A, Warren R, Richardson M, et al. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. N Engl J Med 1999a; 341: 1174-9.

93. van Rie A, Warren RM, Beyers N, et al. Transmission of a multidrug-resistant *Mycobacterium tuberculosis* strain resembling "strain W" among noninstitutionalized, human immunodeficiency virus-seronegative patients. J Inf Dis 1999b; 180: 1608-15.

94. 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: 726-36.

95. 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: 1987-95.

96. van Soolingen D, de Haas PE, Hermans PW, van Embden JD. DNA fingerprinting of *Mycobacterium tuberculosis*. Methods Enzymol 1994; 235: 196-205.

97. van Soolingen D, de Haas PEW, van Doorn HR, Kuijper E, Rinder H, Borgdorff MW. Mutations at amino acid position 315 of the *katG* gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of *Mycobacterium tuberculosis* in The Netherlands. J Infec Dis 2000; 182: 1788-90.

98. van Soolingen D, Qian L, de Haas PE, et al. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. J Clin Microbiol 1995; 33: 3234-8.

99. Veen J. Methods of tuberculosis case-finding in The Netherlands. Bull Int Union Tuberc Lung Dis 1990; 65: 67-9.

100. Veen J. Microepidemics of tuberculosis: the stone-in-the-pond principle. Tuber Lung Dis 1992; 73: 73-6.

101. Verver S, Warren RM, Beyers N, et al. Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. Am J Respir Crit Care Med 2005; 171: 1430-5.

102. Verver S, Warren RM, Munch Z, et al. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. Lancet 2004; 363: 212-4.

103. Victor TC, de Haas PE, Jordaan AM, et al. Molecular characteristics and global spread of *Mycobacterium tuberculosis* with a western cape F11 genotype. J Clin Microbiol 2004; 42: 769-72.

104. Vos MC, de Haas PE, Verbrugh HA, et al. Nosocomial *Mycobacterium bovis*-bacille Calmette-Guerin infections due to contamination of chemotherapeutics: case finding and route of transmission. J Infect Dis 2003a; 188: 1332-5.

105. Vos MC, van Deutekom H, de Haas P, van Soolingen D. Fatal *Mycobacterium bovis* bacille Calmette-Guerin infection caused by contamination of chemotherapeutic agents and not by endogenous reactivation: correction of a previous conclusion. Clin Infect Dis 2003b; 37: 738-9.

106. Warren RM, Victor TC, Streicher EM, et al. Patients with active tuberculosis often have different strains in the same sputum specimen. Am J Respir Crit Care Med 2004; 169: 610-4.

107. Wiid IJ, Werely C, Beyers N, Donald P, van Helden PD. Oligonucleotide (GTG)5 as a marker for *Mycobacterium tuberculosis* strain identification. J Clin Microbiol 1994; 32: 1318-21.

108. World Health Organization. Anti-tuberculosis drug resistance in the world: The WHO/IUATLD global project on anti-tuberculosis drug resistance surveillance. Geneva, Switzerland, WHO/CDS/TB/2004. Available from http://www.who.int/tb/publications/who_htm_tb_2004_343/en/.

109. Yeh RW, Ponce dL, Agasino CB, et al. Stability of *Mycobacterium tuberculosis* DNA genotypes. J Infect Dis 1998; 177: 1107-11.

110. Zignol M, Hosseini MS, Wright A, et al. Global incidence of multidrug-resistant tuberculosis. J Infect Dis 2006; 194: 479-85.