Review Article
Indian J Med Res 120, October 2004, pp 233-247

Molecular epidemiology of tuberculosis

Sujatha Narayanan

Department of Immunology, Tuberculosis Research Centre (ICMR), Chennai, India

Received March 11, 2003

Molecular epidemiology (ME), a blend of molecular biology and epidemiology, is very useful to study the spread of tubercle bacilli in mini epidemics, outbreaks, to analyse the transmission dynamics of tuberculosis (TB) and to determine the risk factors for TB transmission in a community. ME has a great role in distinguishing between exogenous reinfection and endogenous reactivation. In the laboratory, molecular epidemiology can be used to identify cross contamination. Many new DNA typing methods have been introduced after the initial introduction of restriction fragment length polymorphism (RFLP) in 1993. An internationally accepted, standardized protocol for RFLP typing of the *Mycobacterium tuberculosis* complex using IS6110 was published in 1993 and is still used today. Most of the newer DNA typing methods are PCR based and microarray based methods are also available. This will enable individual strains of *M. tuberculosis* or clonal groups to be identified by specific phenotypic traits. ME will continue to be a useful tool in future to measure the impact of any public health intervention strategy for control of tuberculosis in the community.

Key words DNA typing - molecular epidemiology - RFLP - transmission dynamics - tuberculosis

Molecular epidemiology, the study of distribution and determinants of disease occurrence in human populations using molecular techniques, is a blend of molecular biology and epidemiology. Epidemiologic investigations that incorporated DNA fingerprinting of the isolates of *Mycobacterium tuberculosis* have been used to provide novel information about the spread of tubercle bacilli in miniepidemics and outbreaks, to analyse the transmission dynamics of tuberculosis (TB) and to distinguish exogenous reinfection from endogenous reactivation. In addition, ME is also being used to identify the source of laboratory contamination, to determine the risk factors for TB transmission in a community, and to track the geographic distribution and spread of clones of *M. tuberculosis* of public health importance.

Fingerprinting of *M. tuberculosis* exploits restriction fragment length polymorphism (RFLP) of chromosomal DNA. The amplified illustration of the procedure is shown schematically (Fig.). Variation in the array of fragments generated by specific restriction endonucleases are called RFLPs. However, restriction enzyme digestion generates many bands in the gel which make comparison of many gels nearly impossible. To simplify analysis it is possible to perform Southern blotting of electrophoretically - separate DNA followed by hybridisation with probes to determine the presence and size of fragments containing specific genomic DNA restriction fragments.

Repetitive elements called insertion sequences (IS) are present in various sites and variable copy numbers

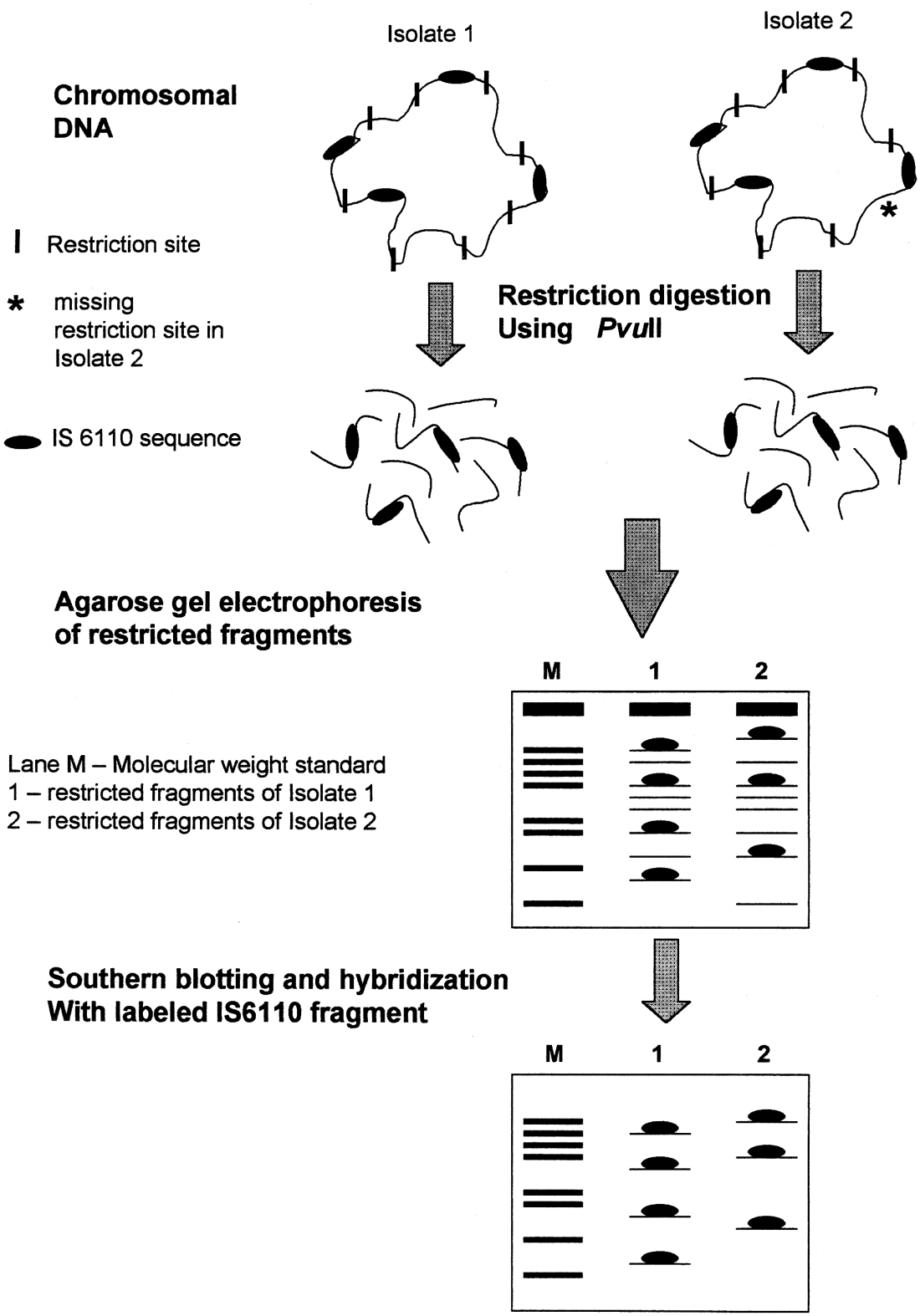

Fig. Restriction fragment length polymorphism (RFLP) can distinguish two isolates of *Mycobacterium tuberculosis*. The chromosomal DNA from 2 clinical isolates of *M. tuberculosis* were digested with restriction enzyme. Pvu II. The resulting DNA fragments were run on agarose gel electrophoresis along with molecular weight marker. The DNA fragments were transferred from the agarose gel to nylon membrane by southern blotting and hybridized with non radioactively labeled IS 6110 repeat element.

in the genomic DNA. These ISs serve as probes, allowing comparison of the number and size of fragments containing an IS. The most commonly used insertion sequence or repetitive element is IS6110 which is found throughout the *M. tuberculosis* complex. It was originally hypothesized that IS6110 insertions occur randomly[1] but that was not true in the sequenced H37Rv strain of *M. tuberculosis*[2,3].

Internal and external molecular weight standards introduced adjacent to the specimen tracks facilitate accurate computer-assisted analysis of IS6110 RFLP patterns. The RFLP band patterns of strains may be compared visually or scanned optically by a computerized reading system and matched to a reference library of strain profiles[2,4]. When used in conjunction with standardized international databases and computer-assisted analysis, this approach allows comparisons of strains between different laboratories in widely separated geographical regions. Two computerized systems, Gel compare version 4.2 program (Applied Maths Inc. Gent Belgium) and (Bio Image whole Band Analyser, version 3.3 Millipore, Ann Arbor MI USA) have been developed specifically for the analysis of RFLP patterns of *M. tuberculosis*. While these systems are suitable for the study of large numbers of isolates, they are expensive and not widely available.

DNA typing methods

An internationally accepted, standardized protocol for RFLP typing of the *M. tuberculosis* complex using IS6110 was published in 1990 and is still used today[5]. Between 0-25 copies of IS6110 are found in almost all strains of *M. tuberculosis* complex[6,7] and is not known to be present in other organisms. IS6110 elements differ in their position and number and this variability is exploited to distinguish between strains.

Though IS6110 RFLP typing is the Gold standard for typing strains of *M. tuberculosis*, it has several disadvantages. It is a slow, cumbersome, labour intensive and technically demanding technique requiring relatively large amounts (*i.e.*, 2 µg) of high quality DNA from each strain of *M. tuberculosis*, an amount that can only be extracted from a large number of bacteria grown from clinical material. The culture of *M. tuberculosis* takes 4-8 wk. Also, this method has poor discriminatory power for isolates with less than 6 copies of IS6110 (<6 bands in the RFLP pattern). To avert the poor discriminatory power of this probe, supplementation of the technique with other probes has been adopted. Various repetitive DNA elements that contribute to strain variation have been discovered in *M. tuberculosis*[8-10]. Polymorphic GC repeat sequence (PGRS), and major polymorphic tandem repeat (MPTR) have a broad host range besides *M. tuberculosis* complex. Among the various repetitive sequences only IS6110 and IS1081 are insertion sequences and the others are short sequences with no known function or phenotype.

The DR region in *M. tuberculosis* complex strains is composed of multiple direct variant repeat sequence (DVRS) each of which is composed of a 36-bp DR and a non repetitive spacer sequence of similar size. It has been shown that there is extensive polymorphism in the DR region by the variable presence of DVRS and this polymorphism is used in the epidemiology of tuberculosis. The DR locus is presently the only well-studied single locus in the genome of *M. tuberculosis* showing considerable strain-to-strain polymorphism. The nature of polymorphism has been used to genotypically classify clinical isolates by DR-RFLP to define epidemiological relationships[11-14].

Spoligotyping is a polymerase chain reaction (PCR)-based method that interrogates a small DR sequence with 36 bp repeats interspersed with short unique, non repetitive sequences 35-71 bp in length. All these spacer nucleotides between the direct repeats can be amplified simultaneously using one set of primers. The presence or absence of spacers in a given biotinylated strain is determined by hybridization with a set of 43 oligonucleotides derived from spacer sequences of *M. tuberculosis* H37Rv. Although the overall discriminatory power of spoligotyping is lower than that of IS6110 typing[15], it has the specific advantage of higher discrimination of strains with low copy numbers of IS6110[16].

The multiple synthetic spacer nucleotides are covalently bound to a nylon membrane in parallel

lines. Hybridization is performed in a 45-lane blotter by applying PCR products of 2X Sodium chloride Sodium Phosphate Ethylene diamino tetracetic acid (SSPE) in the wells. After washing the membrane, the bound fragments are revealed by chemiluminescence by incubating with horse radish peroxidase labeled streptavidin and the autoradiogram is developed.

The most commonly used secondary markers are the polymorphic guanine/cytosine-rich repetitive sequences (PGRS), a triplet repeat of GTG and the major polymorphic tandem repeat (MPTR). The PGRS typing system uses the polymorphic GC-rich sequence contained in the recombinant plasmid pTBN12 as a probe[17,18]. Two other nucleic acid-based typing systems for *M. tuberculosis* have been described. Pulse field gel electrophoresis (PFGE) allows simplified chromosomal restriction fragment patterns to be generated without using probe hybridization methods. In this method, DNA is cleaved with restriction endonucleases that cut DNA infrequently, creating large fragments of chromosomal DNA[19]. The restriction fragments are then separated using sophisticated and expensive electrophoresis equipment. This method discriminates the strains with low IS6110 copies. There is discrepancy between PFGE and IS6110 in classifying strains with IS6110 high copy numbers[19].

PCR-based methods

PCR-based methods are easier to perform, require relatively smaller amounts of genomic DNA and even can be performed on non viable organisms or directly from clinical specimens relative to RFLP genotyping[20-22].

Many PCR based typing assays have been developed in the recent past based on IS6110 as the target. Ligation mediated PCR[23] mixed linker PCR[25], hemi-nested inverse PCR, IS6110 inverse PCR, IS6110 ampliprinting and double-repetitive (DR) element PCR[24] are among the techniques developed to date. Spoligotyping is a PCR based method which has been described before.

Automated detection of DNA fingerprints was achieved using mixed-linker PCR[26]. Mixed-linker DNA fingerprint analysis was attempted using *M. tuberculosis* isolates spotted onto filter paper and concluded that the results were identical to those obtained from conventional culture material[27]. The other method fast ligation-mediated PCR (Flip) is based on mixed-linker method and has the same discriminating power but *M. tuberculosis* isolates can be typed within 6.5 h. Another method, ligation-mediated PCR (LMPCR) uses the 5' end of the flanking sequence of IS6110 for amplification[25,28]. Hemi nested inverse PCR method targets the insertion sequence IS6110 and the upstream flanking regions[29,30,]. All these methods are based on IS6110 element and hence not useful for typing the isolates with low copy numbers of IS6110.

Exact tandem repeats (ETRs) have also been used for PCR-based strain typing assays[31,32]. ETRs differ from polymorphic repeat sequences by having a variable number of tandem repeats ranging from 53 to 79 bp in length, which vary between strains and between different species of the *M. tuberculosis* complex.

A high resolution typing method based on the variable number of tandem repeats (VNTR) of mycobacterial interspersed repetitive units (MIRUs) has been successfully employed in typing the mycobacterial isolates yielding a resolution power close to IS6110-RFLP. MIRUs are short (40-100 bp) DNA elements often found as tandem repeats and dispersed in intergenic regions in the genome of the *M. tuberculosis* complex[33]. The strains vary in the number of repeats at different loci. Each typed strain is assigned a 12-digit number corresponding to the number of repeats at each MIRU loci, forming the basis of a coding system that facilitates interlaboratory comparisons[34-36]. The technical difficulty of sizing the multiple small PCR fragments is overcome by combining multiplex PCR with a fluorescence-based DNA analyzer[37]. Relative to IS6110 RFLP typing, MIRU VNTR profiling is fast, appropriate for strains regardless of their IS6110 RFLP copy number and permits rapid comparison of global strains using a binary data classification system[33].

Fluorescent amplified fragment length polymorphism (FAFLP) typing is a whole genome

approach that involves digesting genomic DNA with two restriction enzymes (*EcoRI* and *Msc I*). The restriction fragments are linked to the adaptors using a DNA ligase. Only particular restriction fragments are visualized after PCR amplification because the primer for the EcoRI adaptor sites contains the selection bases ATC or G labeled with fluorescent dyes and then amplifying the resulting fragments with different fluorescent dye-labeled primers[37]. This method is useful for discriminating low copy number strains.

Kremer *et al*[38] compared 5 different methods of RFLP typing which employed IS6110, IS1081, PGRS, the DR and the GTSS repeat as probes. Of the PCR- based methods compared, VNTR typing, mixed-linker PCR and spoligotyping were highly reproducible between different laboratories.The double repetitive PCR (DRE-PCR), IS6110 inverse PCR, IS6110 ampliprinting and arbitrarily primed PCR were not reproducible. Despite the development of different typing methods, RFLP using IS6110 is being widely used and considered the Gold standard to which other methods are compared[39]. Thus implementation of multiple molecular techniques in a single study provided better discrimination between strains and insight for phylogenetic groupings[40]. Today, most of the molecular epidemiologic studies rely on IS6110 RFLP typing and a secondary typing method such as PGRS or spoligotyping for isolates with less than 6 bands in the IS6110 RFLP band pattern.

There is rising interest in identifying relationships between strains that have a specific phenotype such as increased infectivity, virulence, or hypermutability. Direct comparison of genomic DNA sequences of strains of *M. tuberculosis* would be the best way of quantitatively determining whether the two strains are similar or different, but DNA sequencing is still too expensive and complex to be applied in practical situations to large numbers of isolates. Currently, it is possible to analyze short segments of DNA for sequence similarities and differences. Genomic fragments can be amplified using PCR, and an automated DNA-sequencing procedure involving fluorescent dye-labeled terminators can be used to directly sequence the PCR-amplified DNA fragment[41]. This approach allows a DNA fragment of 300 to 500 bp to be sequenced in 24 h. In future, improvements in automation of target amplification and direct sequence analysis may lead to practical implementation of this method in laboratories.

Another approach is to evaluate the relatedness of strains based on the whole genome sequence using DNA microarrays and DNA chip technology. These techniques allow simultaneous detection of genetic variation at various genomic sites by analysis of the amount and specific location of mycobacterial DNA. Conceptually, they use oligonucleotide arrays containing thousands of oligonucleotides on a limited surface[42].

Deletion microarray approach will potentially provide information both on phylogenetic relationships and information about specific biologically relevant phenotypes. Briefly, the genome of a strain is compared against that of a known, sequenced reference strain, using a microarray. Any deletions that have occurred will be detected in the comparison. Since deletions rarely occur independently at exactly the same chromosomal locus, they can be considered unique and irreversible genetic events. The number and distribution of these deletions provide a genomic pattern that can be used to construct phylogenetic relationships. The genomic patterns can also be used to determine whether the loss of specific genes is related to the phenotype of a strain, such as its transmissibility or antigenicity.

Molecular epidemiology as a tool to identify outbreaks and to analyse the transmission dynamics of TB

Outbreak situation usually involves person-to-person spread or simultaneous infection from a common source. By definition, all isolates involved in outbreak of an infection would be expected to be clonal. Non clonality, which is often easier to determine, eliminates an isolate from consideration in a specific chain of transmission. Ideally, strain typing will provide a clear, objective basis for identifying the outbreak strain and distinguishing it from epidemiologically unrelated isolates. Many

studies on TB have extended these assumptions to define clusters of patients in the community based on identical DNA fingerprinting patterns from the isolates of *M. tuberculosis*. Conventional TB contact investigations use circuitous approaches to collect information and to screen spouses, partners, other household members, co-workers and increasingly distinct contacts for TB infection and disease[43]. Several studies have added molecular typing of the isolates of contacts who were also TB cases, in order to trace the source of infection. Molecular epidemiological data overlaid with conventional epidemiology data would help in knowing the transmission dynamics. In a high incidence area in Barcelona, Spain (163 TB cases/100,000 population), there was 61.5 per cent concordance between the DNA fingerprint results (IS6110 RFLP and PGRS) and conventional contact tracing[44]. In this study the authors concluded that conventional contact tracking was useful for identifying new TB cases, but it did not provide much information about the chains of TB transmission and how to block or prevent that.

In a five-year population-based study in the Netherlands, contact investigations of persons in five of the largest clusters identified epidemiological links between them based on time, place and risk factors. However TB transmission also occurred only through short term, casual contact that was not easily detected in routine contact investigations[45].

In low-incidence areas such as San Francisco (California, USA)[43], Zurich[45] and Amsterdam[46], a relatively small percentage (5-10%) of cases having identical RFLP patterns were actually identified as a contact by the source case. This suggests that unsuspected transmission of TB occurs and is not easily traced by conventional contact tracing investigations[47]. In a contact tracing study done at Thiruvallur near Chennai, India, only 10 per cent concordance was seen between conventional epidemiology and molecular epidemiology using IS6110 and DR probes[14]. Among the patients in the clusters having identical fingerprints by IS6110 and DR, only 10 per cent could name the contact which could be a source case[14].

In summary, DNA fingerprinting is a useful tool to confirm or rule out the possibility of recent TB transmission between two or more persons. It has also shown that TB transmission can occur through short, casual and unsuspected contacts. Molecular epidemiologic studies suggest that the traditional or classical contact tracing approaches such as DNA fingerprinting could be particularly useful to guide contact tracing strategies in low incidence areas, where its predictive value would be high.

Molecular epidemiological studies have provided novel insights into the transmission dynamics of tuberculosis[48]. Such an approach has shown that a drug-susceptible strain of *M. tuberculosis* (the C or J strain) which was first identified as causing a large outbreak in 1990 in a homeless shelter[49] has become widely prevalent in New York city[50]. The availability of standardized genotyping technique for *M. tuberculosis* and the existence of extensive collections of fingerprints made it possible to do a molecular epidemiological assessment of tuberculosis transmission between different geographic regions[51]. Daley *et al*[52] described 12 cases of TB that occurred in a housing facility in San Francisco, USA, among HIV infected people. The demonstration of transmission of *M. tuberculosis* in nosocomial settings[53-55], congregate living facilities[52] and among persons at high risk such as the homeless[56,57] and those who are HIV infected[54-55] has been especially important. Fingerprinting in the context of geographic studies has shown the acquisition of *M. tuberculosis* of Tunisian or Ethiopian genotypes by Dutch persons who resided in Tunisia or Ethiopia[58] as well as spread of the organisms between Greenland and Denmark[59].

Exogenous infection vs endogenous reactivation

Post-primary TB which occurs many years after a primary infection, may develop as the result of reactivation of the endogenous primary infection or as a result of a recent exogenous infection. In this era of effective treatment regimens, the notion that multiple episodes of TB in one patient are almost always caused by endogenous reactivation may be questioned. It is now possible to characterize the genotype of *M. tuberculosis* by DNA fingerprinting,

which can show whether a new episode of the disease is caused by infection with the same strain that caused a previous episode or by a different strain. Thus, molecular epidemiology using DNA fingerprinting can determine the proportion of cases due to recent infection and the proportion due to reactivation.

RFLP studies conducted in Hong Kong[60] showed that the patterns of 88 per cent of the isolates from patients with relapses matched those for their pre-treatment counterparts indicating a high frequency of occurrence of infections caused by endogenous reactivation of *M. tuberculosis*. A study conducted at the Tuberculosis Research Centre (TRC), Chennai on pre- and post-treatment isolates by DR-RFLP analysis indicated (69% of the isolates by DR probe and 50% by IS61100) a high degree of endogenous reactivation among patients who have relapses after successful completion of chemotherapy[61,62]. Small *et al*[63] used IS6110 typing to trace exogenous reinfection with multidrug-resistant *M. tuberculosis* in patients with advanced HIV infection. Recently, molecular epidemiological study undertaken in a rural area near Chennai, India as part of the model DOTS (directly observed therapy short course) programme using fingerprinting with two probes (IS6110 and DR) and cluster analysis revealed more of endogenous reactivation than exogenous reinfection in the community[14]. Similar observations were made by the molecular biological study conducted in New York City from 1989 to 1992[64] and in San Francisco, California during 1991 and 1992[43].

Laboratory contamination

It is very important to determine whether a group of culture positive isolates represents a true outbreak of TB or a pseudo outbreak based on false positive laboratory cultures of *M. tuberculosis*. DNA fingerprinting analysis is a very good tool to identify false positive laboratory cultures. Earlier investigations focused on the isolates of *M. tuberculosis* that were processed together in the laboratory and had identical IS6110 RFLP patterns, but were from at least one otherwise asymptomatic patient[65,66]. In a study conducted in New York City[67], an isolate was collected from every patient with a positive culture for *M. tuberculosis* during a one-month period, including both incident and prevalent cases, and RFLP analyses were performed. The DNA fingerprinting of all *M. tuberculosis* isolates from a 700-bed urban hospital in Chicago, USA, revealed only one possible instance of nosocomial transmission and five false-positive *M. tuberculosis* cultures out of 183 patients[68]. In another study[69], isolates collected prospectively over 5 yr from a municipal health department laboratory, underwent DNA fingerprinting using IS6110 and pTBN12 sequences, clinical and laboratory records of all isolates with matching DNA fingerprints and processed within 42 days of each other, were reviewed, and 4.0 per cent of the culture-positive patients were identified as probable or definite false-positives. In a convenience sample of isolates from three other mycobacterial laboratories, 12 per cent were found to be definite or probable false-positive. The reasons for laboratory cross-contamination are careless specimen processing and contaminated reagents[69]. A small, but non-negligible proportion of cases with laboratory cross-contamination was detected in every institution that looked for it[70-73]. As a result, DNA fingerprinting is now used in some settings to routinely evaluate all specimens for possible laboratory cross-contamination.

In general, laboratory cross-contamination should be considered if isolates were cultured within one week of each other and had identical DNA fingerprints. Laboratory contamination should be suspected when *M. tuberculosis* is grown from smear-negative specimens, from low-yield cultures, and from patients who are otherwise asymptomatic. A single positive culture in clinically well patients with negative acid fast bacilli (AFB) smears and no other evidence of TB may not always need therapy. Laboratory cross-contamination should also be suspected when there is a sudden increase in culture positive isolates, without an epidemiological or clinical explanation. For example, adopting more rapid and sensitive methods may increase the contamination rate. The isolates should be analyzed by reliable molecular typing techniques, and compared with specimens that were originally processed during the same time period. Many

investigators used IS6110 RFLP typing, VNTR typing[73] or spoligotyping[74] to detect and evaluate laboratory cross-contamination.

Simultaneous infection by more than one strain of *M. tuberculosis* by RFLP

It has been understood from the recent reports that a single patient could be infected with more than one strain of *M. tuberculosis* at any given time as more reports are confirming infections by multiple strains. Phage typing method was used in the 1970s to detect the presence of more than one strain (phage types) in a single patient[75-78]. Due to technical complexity of the assay method, the results were not reliable. With the advent of newer methods of genotyping in early nineties, like IS6110-based DNA fingerprinting together with secondary typing methods, it is possible now to precisely identify specific strains of *M. tuberculosis* isolated from clinical samples.

Few reports have shown the simultaneous infection with two or more strains of *M. tuberculosis* by RFLP [79]. Yeh *et al*[80] demonstrated the existence of simultaneous infection with two strains of *M. tuberculosis* using IS6110 DNA fingerprinting, based on the relative intensities of the band patterns. Infections from multiple strains of *M. tuberculosis* are sometimes mistaken to be due to laboratory cross-contamination. It is important to identify "true" mixed infections to gain insights into the patterns of transmission of the disease in the community. Molecular epidemiological approaches have provided novel insights. Adoption of more rigorous reporting standards in studies of the molecular epidemiology of tuberculosis would improve the comparability of studies and help investigators to assess the implications of their results[81].

Risk factors and settings for recent transmission

Molecular typing techniques in combination with conventional epidemiological methods, can be used to identify the risk factors associated with recent transmission. Cases defined as patients whose isolates have clustered RFLP patterns, and controls are defined as patients whose isolates have unique band patterns. The risk factors that are associated with recent infection are specific to a particular community and others are common to TB patients in geographical areas. In San Francisco, among persons < 60 yr of age, Hispanic ethnicity, birth in the United States and a diagnosis of AIDS were independently associated with being in a cluster[43]. Specific interventions were directed at persons with one or more of the independent risk factors, and consequently the proportion of TB cases that were clustered decreased over time[82]. In a recent study in New York city birth outside the United States, age $\geq$ 60 yr, and diagnosis after 1993 were independently associated with reaction of latent tuberculosis infection (LTBI), while homelessness was associated with clustering or recent transmission. TB among the foreign-born persons was more likely to result from the reactivation of LTBI among those who were not infected with HIV[82]. The researchers recommended that TB prevention and control strategies need to be targeted to the large number of foreign born persons in New York city who have latent TB infection. However, HIV was not associated with clustering among TB patients in a university teaching hospital Rio de Janerio, Brazil[83] and HIV was not a risk factor for clustering among South African gold miners[84].

The limited numbers of molecular epidemiological studies conducted in India were laboratory-based and comprised small numbers of patients[61,62]. The recent study from Tuberculosis Research Centre[14] was the first in India to combine molecular and conventional epidemiologic techniques to investigate the mechanism and risk factors of transmission. They reported several characteristics of the molecular epidemiology of TB in the rural settings at Chennai India using IS6110 and DR probes which differ from previously reported findings in other settings. Forty one per cent of *M. tuberculosis* isolates harboured a single IS6110 copy. Such a high proportion of single-copy isolates has not been reported elsewhere except south India[84]. The proportion of clustering in this study ranged from 9 to 38 per cent depending on whether single-copy strains were excluded or included in the analysis. Clustering was higher in older patients contrary to the observation by many other investigators[86-89].

Geographical distribution and dissemination of tuberculosis

There may be a link between geographic location and IS6110 number. Some isolates of *M. tuberculosis* contains no or very few copies of IS6110. One early study based on 1S typing claimed that *M. tuberculosis* strains from regions in Central Africa, where tuberculosis is highly endemic, are generally related to each other than isolates from the Netherlands, where the transmission rate is slow and where the majority of TB cases are presumed to be the result of reactivation of LTBI[90].

Several of the strains identified in outbreaks have been associated with large clusters that are widely dispersed both geographically and temporarily, suggesting they are either more transmissible or they are more likely to cause disease once transmitted than are other strains. The most commonly cited and reviewed example of the geographical dissemination of a particular clone of *M. tuberculosis* is that of Beijing/W strains[91,92] which is a multidrug-resistant strain of *M. tuberculosis*, responsible for causing many cases of TB and deaths attributable to TB among patients and health care workers in nosocomial outbreaks and other institutional settings in New York city during the 1990s[93-95]. This strain was later found in other parts of USA. By the late 1990s the W strain was recognized as the member of Beijing genotype family strains. A study performed in Beijing, China reported that 85 per cent of the isolates were strains with more than 66 per cent similarity among their IS6110 RFLP patterns[96]. This "Beijing family" of strains was also detected in high proportions among strains in other parts of Asia[97], the former Russian Federation[98-100] and Estonia Latin America[101]. Beijing stains including the W strain and its variants, have an insertion of IS 6110 in the dnaA-dnaN locus[102]. Based on several early technical studies and a review of 16 studies of the Beijing or W strains that gave results on spoligotyping, the W family and Beijing family strains have as identical, characteristic spoligotype based on DNA polymorphism in the direct repeat region that contains spacers 35-43[91,94,103-106]. The true proportion of TB cases attributable to the Beijing family of strains and an association between the Beijing family of strains and drug resistance is hard to assess.

Transmission of drug resistant strain

There is no evidence of a lower risk of infection among contacts exposed to TB patients with drug resistant pulmonary TB[107]. A population based study in Mexico reported that MDR-TB were less likely to be in clusters relative to persons with drug-susceptible TB[107]. Similar results were reported by studies among South African gold miners[84] and in the Netherlands[108]. Except in localized areas with poor cure rates, and a high prevalence of HIV, it is unlikely that drug resistance strains spread fast. This has been shown by mathematical modeling of the relative transmission of drug resistant versus drug sensitive strains[109]. The studies showing reducing bacterial transmissibility are predominantly for strains resistant to isoniazid. Isoniazid is a key component of the short-course regimen for treatment of TB. Studies with animal models showed that isoniazid resistant strains caused significantly less disease in guinea pigs than did drug susceptible strains[110-112]. Specific mutations or deletions within the *KatG* gene of isoniazid resistant stains of *M. tuberculosis* have been associated with decrease in its pathogenecity[113,114]. The most commonly occurring *KatG* mutation were [serine 315 replaced by threonine (S 315T] is associated with clinically significant levels of isoniazid resistance. Mycobacterial genome sequence and molecular epidemiology reveal the phenotypic and genotypic associations[114].

The completed, published genome sequence of *M. tuberculosis* provides an enormous amount of information that will widen research in molecular epidemiology and mycobacteria genomics[115]. There are a number of molecular typing techniques available which will enable individual stains or clonal groups to be identified by specific phenotypic traits to study the genetic basis of these important traits using gene expression profiling with microarrays. The strains are being examined for specific differences in correlation with bacterial phenotypes such as tissue tropism, virulence, transmissibility, pathogenesis, antogenecity, resistance to antimicrobial agents and immunogenecity. The casual relationships can be established if we understand the specific polymorphism, deletions or other changes in the genotypes of the strains.

Future research should focus on phenotypic characteristics, gene expression and genotype-phenotype correlations in *M. tuberculosis* strains. Molecular epidemiological methods will continue to play an important role to identify appropriate public health interventions and to measure their impact. However, most of these studies are being conducted only in industrialized countries and resource-rich areas that have a relatively low incidence of TB. Therefore, the inferences drawn and their applications are limited. There is a strong need for additional studies in different geographical areas and populations with a high burden of disease. There is a need for a better understanding of the epidemiology of tuberculosis; instead of using molecular epidemiology only as a tool for molecular typing, we need to find ways to enlist this tool to answer questions of major public health importance.

References

1. Manca C, Tsenova L, Barry CE III, Bergtold A, Freeman S, Haslett PA, *et al*. *Mycobacterium tuberculosis* CDC 1551 induces a more vigorous host response *in vivo* and *in vitro*, but is not more virulent than other clinical isolates. *J Immunol* 1999; *162* : 6740-6.

2. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, *et al*. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; *393* : 537-44.

3. Hermans PWM, van Soolingen D, Dale JW, Schuitema ARJ, McAdam RA, Catty D, *et al*. Insertion element IS986 from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. *J Clin Microbiol* 1990; *28* : 2051-8.

4. Heersma HF, Kremer K, van Embden JD. Computer analysis of IS6110 RFLP patterns of *Mycobacterium tuberculosis*. *Methods Mol Biol* 1998; *101* : 395-422.

5. Thierry D, Cave MD, Eisenach KD, Crawford JT, Bates JH, Gicquel B, *et al*. IS6110, an IS-like element of *Mycobacterium tuberculosis* complex. *Nucleic Acids Res* 1990; *18* : 188.

6. Agasino CB, Ponce de Leon A, Jasmer RM, Small PM. Epidemiology of *Mycobacterium tuberculosis* strains in San Francisco that do not contain IS6110. *Int J Tuberc Lung Dis* 1998; *2*: 518-20.

7. Cave MD, Eisenach KD, McDermott PF, Bates JM, Crawford JT. IS6110: conservation of sequence in the *Mycobacterium tuberculosis* complex and its utilization in DNA fingerprinting. *Mol Cell Probes* 1991; *5* : 73-80.

8. McAdam RA, Hermans PW, van Soolingen D, Zainuddin ZF, Catty D, van Embden JD, *et al*. Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS3 family. *Mol Microbiol* 1990; *4* : 1607-13.

9. 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* : 2695-705.

10. Collins DM, Stephens DM. Identification of an insertion sequence, IS1081, in *Mycobacterium bovis*. *FEMS Microbiol Lett* 1991; *67*: 11-5.

11. Goguet dela Salmoniere YO, Li HM, Torrea G, Bunschoten A, Embden JDA, Gicquel B. Evaluation of spoligotyping in a study of the transmission of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1997; *35* : 2210-4.

12. Warren RM, Streicher EM, Charalambous S, Churchyard G, van der Spuy GD, Grant AD, *et al*. Use of spoligotyping for accurate classification of recurrent tuberculosis. *J Clin Microbiol* 2002; *40* : 3851-3.

13. Narayanan S, Sahadevan R, Narayanan PR, Krishnamurthy PV, Paramasivan CN, Prabhakar R. Restriction fragment length polymorphism of *Mycobacterium tuberculosis* strains from various regions of India using direct repeat probe. *Indian J Med Res* 1997; *106* : 447-54.

14. Narayanan S, Das S, Garg R, Hari L, Rao VB, Frieden TR, *et al*. Molecular epidemiology of tuberculosis in a rural area of high prevalence in south India : implications for disease control and prevention. *J Clin Microbiol* 2002; *40* : 4785-8.

15. Mistry NF, Iyer AM, D'Souza DT, Taylor GM, Young DB, Antia NH. Spoligotyping of *Mycobacterium tuberculosis* isolates from multiple-drug-resistant tuberculosis patients from Bombay, India. *J Clin Microbiol* 2002; *40* : 2677-80.

16. Bauer J, Andersen AB, Kremer K, Miorner H. Usefulness of spoligotyping to discriminate IS6110 low-copy-number *Mycobacterium tuberculosis* complex strains

cultured in Denmark. *J Clin Microbiol* 1999; *37*: 2602-6.

17. Cousins DV, Williams SN, Ross BC, Ellis TM. Use of a repetitive element isolated from *Mycobacterium tuberculosis* in hybridization studies with *Mycobacterium bovis*: a new tool for epidemiological studies of bovine tuberculosis. *Vet Microbiol* 1993; *37* : 1-17.

18. Yang Z, Chaves F, Barnes PF, Burman WJ, Koehler J, Eisenach KD, *et al*. Evaluation of method for secondary DNA typing of *Mycobacterium tuberculosis* with pTBN12 in epidemiologic study of tuberculosis. *J Clin Microbiol* 1996; *34* : 3044-8.

19. Singh SP, Salamon H, Lahti CJ, Farid-Moyer M, Small PM. Use of pulsed-field gel electrophoresis for molecular epidemiologic and population genetic studies of *Mycobacterium tuberculosis. J Clin Microbiol* 1999; *37* : 1927-31.

20. Qian L, van Embden JDA, van der Zanden AGM, Weltevreden EF, Duanmu H, Douglas JT. Retrospective analysis of the Beijing family of *Mycobacterium tuberculosis* in preserved lung tissues. *J Clin Microbiol* 1999; *37* : 471-4.

21. Taylor GM, Goyal M, Legge AJ, Shaw RJ, Young D. Genotypic analysis of *Mycobacterium tuberculosis* from medieval human remains. *Microbiology* 1999; *145* : 899-904.

22. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, *et al*. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997; *35* : 907-14.

23. Bonora S, Gutierrez MC, Di Perri G, Brunello F, Allegranzi B, Ligozzi M, *et al*. Comparative evaluation of ligation-mediated PCR and spoligotyping as screening methods for genotyping of *Mycobacterium tuberculosis* strains. *J Clin Microbiol* 1999; *37*: 3118-23.

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

25. 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.

26. Butler WR, Haas WH, Crawford JT. Automated DNA fingerprinting analysis of *Mycobacterium tuberculosis* using fluorescent detection of PCR products. *J Clin Microbiol* 1996; *34*: 1801-3.

27. Prodhom G, Guilhot C., Gutierrez MC, Vernent A, Giequel B, Vincent V. Rapid discrimination of *Mycobacterium tuberculosis* complex strains by ligation-mediated PCR fingerprint analysis. *J Clin Microbiol* 1997; *35* : 3331-4.

28. Kearns AM, Barrett A, Marshall C, Freeman R, Magee JG, Bourke SJ, *et al*. Epidemiology and molecular typing of an outbreak of tuberculosis in a hostel for homeless men. *J Clin Pathol* 2000; *53*: 122-4.

29. Patel S, Wall S, Saunders NA. Heminested inverse PCR for IS6110 fingerprinting of *Mycobacterium tuberculosis* strains. *J Clin Microbiol* 1996; *34*:1686-90.

30. Goyal M, Young D, Zhang Y, Jenkins PA, Shaw RJ. PCR amplification of variable sequence upstream of katG gene to subdivide strains of *Mycobacterium tuberculosis* complex. *J Clin Microbiol* 1994; *32*: 3070-1.

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. Skuce RA, McCorry TP, McCarroll JF, Roring SM, Scott AN, Brittain D, *et al*. Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. *Microbiology* 2002; *148*: 519-28.

33. 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.

34. 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.

35. Mazars E, Lesjean S, Banuls AL, Gilbert M, Vincent V, Gicquel B, *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*: 1901-6.

36. 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.

37. Fillol I, Ferdinand S, Negroni L, Sola C, Rastogi N, Molecular typing based on variable number of tandem DNA repeats used alone and in association with spoligotyping. *J Clin Microbiol* 2000; *38* : 2520-4.

38. Kremer K, van Soolingen D, Frothingham R, Haas WH, Hermans PWM, Martin C, *et al*. Comparison of methods based on different molecular epidemiologic markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* 1999; *37*: 2607-18.

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

40. van Soolingen D, Hermans PWM, de Haas PEW, Soll DR, van Embden JDA. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol* 1991; *29*: 2578-86.

41. Tracey TE, Mulcny LS. A simple method for direct automated sequencing of PCR fragments. *Biotechniques* 1991; *1* : 68-75.

42. Kato-Maeda M, Rhee JT, Gingeras TR, Salamon H, Drenkow J, Smittipat N, *et al*. Comparing genomes within species of *Mycobacterium tuberculosis*. *Genome Res* 2001; *11*: 547-54.

43. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC, *et al*. The epidemiology of tuberculosis in San Francisco: a population-based study using conventional and molecular methods. *N Engl J Med* 1994; *330* : 1703-9.

44. Jones WD Jr. Bacteriophage typing of *Mycobacterium tuberculosis* cultures from incidents of suspected laboratory cross-contamination. *Tubercle* 1988; *69*: 43-6.

45. Pfyffer GE, Strassle A, Rose N, Wirth R, Brandli O, Shang H. Transmission of tuberculosis in the metropolitan area of Zurich: a 3 year survey based on DNA fingerprinting. *Eur Respir J* 1998; *11*: 804-8.

46. van Deutekom H, Gerritsen JJJ, van Soolingen D, van Ameijden EJC, van Embden JD, Coutinho RA. A molecular epidemiological approach to studying the transmission of tuberculosis in Amsterdam. *Clin Infect Dis* 1997; *25*: 1071-7.

47. Sebek M. DNA fingerprinting and contact investigation. *Int J Tuberc Lung Dis* 2000; *4*: (2 Suppl 1): S45-8.

48. Small PM, van Embden JDA. Molecular epidemiology of tuberculosis. In: Bloom BR, editor, *Tuberculosis: Pathogenesis, protection and control*. Washington, DC: American Society for Microbiology Press; 1994 p. 6569-82.

49. Dobbs KG, Lok KH, Bruce F, Mulcahy D, Benjamin WH, Dunlap NE. Value of *Mycobacterium tuberculosis* fingerprinting as a tool in a rural state surveillance program. *Chest* 2001; *120* :1877-82.

50. Friedman CR, Stoeckle MY, Kreiswirth BN, Johnson WD Jr, Manoach SM, Berger J, *et al*. Transmission of multidrug-resistant tuberculosis in a large urban setting. *Am J Respir Crit Care Med* 1995; *152*: 355-9.

51. van Embden JDA, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, *et al*. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting : recommendations for a standardized methodology. *J Clin Microbiol* 1993; *31* : 406-9.

52. Daley CL, Small PM, Schecter GF, Schoolnik GK, McAdam RA, Jacobs WR Jr. 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*: 231-5.

53. Frieden TR, Sherman LF, Maw KL, Crawford JT, Nivin B, *et al*. A multi-institutional outbreak of highly drug-resistant tuberculosis. Epidemiology and clinical outcomes. *JAMA* 1996; *276*: 1229-35.

54. Edlin BR, Tokars JI, Grieco MH, Crawford JT, Wiliams J, Sordillo EM, *et al*. An outbreak of multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome. *N Engl J Med* 1992; *326*: 1514-21.

55. Jereb JA, Burwen DR, Dooley SW, Haas WH, Crawford JT, Geiter LJ, *et al*. Nosocomial outbreak of tuberculosis in a renal transplant unit: application of new technique for restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolates. *J Infect Dis* 1993; *168*: 1219-24.

56. Genewein A, Telenti A, Bernasconi C, Mordasin C, Weiss S, Mourer AM, *et al*. Molecular approach to identifying route of transmission of tuberculosis in the community. *Lancet* 1993; *342*: 841-4.

57. Barnes PF, el-Hajj H, Preston-Martin S, Cave MD, Jones BE, Otaya M, *et al*. Transmission of tuberculosis among the urban homeless. *JAMA* 1996; *275*: 305-7.

58. Hermans PW, Messadi F, Guebrexabher H, van Soolingen D, de Haas PE, Heersma 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*: 1504-13.

59. Yang ZH, de Haas PEW, van Soolingen D, van Embden JDA, Andersen AB. Restriction fragment length polymorphism of *Mycobacterium tuberculosis* strains isolated from Greenland during 1992; evidence of tuberculosis transmission between Greenland and Denmark. *J Clin Microbiol* 1994; *32*: 3018-25.

60. Das S, Chan S L, Allen B W, Mitchison D A, Lowrie D B. Application of DNA fingerprinting with IS986 to sequential mycobacterial isolates obtained from pulmonary tuberculosis patients in Hong Kong before, during and after short-course chemotherapy. *Tuberc Lung Dis* 1993; *74*: 47-51.

61. Sahadevan R, Narayanan S, Paramasivan CN, Prabhakar R, Narayanan PR. Restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in madras, India, by use of Direct-Repeat probe. *J Clin Microbiol* 1995; *33*: 3037-9.

62. Das S, Paramasivan C N, Lowrie D B, Prabhakar R, Narayanan P R. IS6110 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.

63. Small PM, Shafer RW, Hopewell PC, Singh SP, Murphy MJ, Desmond E, *et al*. Exogenous reinfection with multidrug resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. *N Engl J Med* 1993; *328*: 1137-43.

64. Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, *et al*. Transmission of tuberculosis in New York City: an analysis by DNA fingerprinting and conventional epidemiologic methods. *N Engl J Med* 1994; *330*: 1710-6.

65. Dunlap NE, Harris RH, Benjamin WH Jr, Harden JW, Hafner D. Laboratory contamination of *Mycobacterium tuberculosis* cultures. *Am J Respir Crit Care Med* 1995; *152*: 1702-4.

66. 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. *Tuber Lung Dis* 1996; *77*: 407-13.

67. French AL, Welbel SF, Dietrich SE, Mosher LB, Breall PS, Paul WS, *et al*. Use of DNA fingerprinting to assess tuberculosis infection control. *Ann Intern Med* 1998; *129*: 856-61.

68. Burman WJ, Stone BL, Reves RR, Wilson ML, Yang Z, El-Hajj H, *et al*. The incidence of false-positive cultures for *Mycobacterium tuberculosis*. *Am J Respir Crit Care Med* 1997; *155*: 321-6.

69. Nivin B, Fujiwara PI, Hannifin J, Kreiswirth BN. Cross-contamination with *Mycobacterium tuberculosis*: an epidemiological and laboratory investigation. *Infect Control Hosp Epidemiol* 1998; *19*: 500-3.

70. Perfecto B, Dorronsoro I, Lopez-Goni I. Confirmation by molecular typing of cross-contamination in a mycobacteria laboratory. *Enferm Infecc Microbiol Clin* 2000; *18*: 12-5.

71. Breese PE, Burman WJ, Hildred M, Stone B, Wilson ML, Yang Z, *et al*. The effect of changes in laboratory practices on the rate of false-positive cultures for *Mycobacterium tuberculosis*. *Arch Pathol Lab Med* 2001; *125*: 1213-6.

72. Anon. Misdiagnoses of tuberculosis resulting from laboratory cross-contamination of *Mycobacterium tuberculosis* cultures - New Jersey, 1998. *MMWR Morb Mortal Wkly Rep* 2000; *49*: 413-6.

73. Gascoyne-Binzi DM, Barlow RE, Frothingham R, Robinson G, Xollyns TA, Gelleflie R, *et al*. Rapid identification of laboratory contamination with *Mycobacterium tuberculosis* using variable number tandem repeat analysis. *J Clin Microbiol* 2001; *39* : 69-74.

74. Nivin B, Driscoll J, Glaser T, Bifani P, Munsiff S. Use of spoligotype analysis to detect laboratory cross-contamination. *Infect Control Hosp Epidemiol* 2000; *21*: 525-7.

75. Bates JH, Stead WW, Rado TA. Phage type of tubercle bacilli isolated from patients with two or more sites of organ involvement. *Am Rev Respir Dis* 1976; *114*: 353-8.

76. Mankiewicz E, Liivak M. Phage types of *Mycobacterium tuberculosis* in cultures isolated from Eskimo patients. *Am Rev Respir Dis* 1975; *111*: 307-12.

77. Crawford J, Bates JH. Phage typing of mycobacteria In: Kubica GP, Wayne LG, editors, *The Mycobacteria: a sourcebook: Part A*. New York: Marcel Dekker, 1984 p.123-32.

78. Glynn JR, Jenkins PA, Fine PEM, Ponnighans JM, Sterne JAC, Mkandwire PK, *et al*. Pattenrs of initial and acquired antituberculosis drug resistance in Karonga district, Malawi. *Lancet* 1995; *345*: 907-10.

79. Pavlic M, Allerberger F, Dierich MP, Prodinger WM. Simultaneous infection with two drug-susceptible *Mycobacterium tuberculosis* strains in an immunocompetent host. *J Clin Microbiol* 1999; *37*: 4156-7.

80. 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*: 537-9.

81. Glynn JR, Vynnycky E, Fine PEM. Influence of sampling on estimates of clustering and recent transmission of *Mycobacterium tuberculosis* derived from DNA fingerprinting techniques. *Am J Epidemiol* 1999; *149*: 366-71.

82. Geng E, Kreiswirth B, Driver C, Li J, Burzynski J, DellaLatta P, *et al*. Changes in the transmission of tuberculosis in New York City from 1990 to 1999. *N Engl J Med* 2002; *346*: 1453-8.

83. Godfrey-Faussett P, Sonnenberg P, Shearer SC, Bruce MC, Mee C, Morris L, *et al*. Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet* 2000; *356*: 1066-71.

84. Fandinho FC, Kritski AL, Hofer C, Conde H Jr., Ferreira RMC, Saad MHF, *et al*. RFLP patterns and risk factors for recent tuberculosis transmission among hospitalized tuberculosis patients in Rio de Janeiro, Brazil. *Trans R Soc Trop Med Hyg* 2000; *94*: 271-5.

85. Radhakrishnan I, Manju YK, Kumar RA, Mundayoor S. Implications of low frequency of IS6110 in fingerprinting field isolates of *Mycobacterium tuberculosis* from Kerala, India. *J Clin Microbiol* 2001; *39*: 1683.

86. Friedman CR, Quinn GC, Kreiswirth BN, Perlman DC, Salomon N, Schlinger N, *et al*. Widespread dissemination of a drug-susceptible strain of *Mycobacterium tuberculosis. J Infect Dis* 1997; *176*: 478-84.

87. Frieden TR, Fujiwara PI, Washko RM, Hamburg MA. Tuberculosis in New York City- turning the tide. *N Engl J Med* 1995; *333*: 229-33.

88. Snider DE Jr, Kelly GD, Cauthen G, Thompson NJ, Kilburn JO. Infection and disease among contacts of tuberculosis cases with drug-resistant and drug-susceptible bacilli. *Am Rev Respir Dis* 1985; *132*: 125-32.

89. Warren R, Hauman J, Bayers N, Richardson M, Schaaf HS, Donald P, *et al*. Unexpectedly high strain diversity of *Mycobacterium tuberculosis* in a high-incidence community. *S Afr Med J* 1996; *86*: 45-9.

90. Casper C, Singh SP, Rane S, Daley CL, Schecter GS, Riley LW, *et al*. The transcontinental transmission of tuberculosis: a molecular epidemiological assessment. *Am J Public Health* 1996; *86*: 551-3.

91. 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.

92. Daley CL, Small PM, Schecter GF, Schoolnik GK, McAdam RA, Jacobs WR, Jr., Hopewell PC. An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphism. *N Engl J Med* 1992; *326* : 231-5.

93. Bifani PJ, Plikaytis BB, Kapur V, Stockbauer K, Pan X, Lutfey ML, *et al*. Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. *JAMA* 1996; *275*: 452-7.

94. Bifani PJ, Mathema B, Liu Z, Moghazeh S, Shobsin B, Tempalski B, *et al*. Identification of a 'W' variant outbreak of *Mycobacterium tuberculosis* via population-based molecular epidemiology. *JAMA* 1999; *282*: 2321-7.

95. Agerton TB, Valway SE, Blinkhorn RJ, Shilkret KL, Reves R, Schluter W, *et al*. Spread of strain W, a highly drug-resistant strain of *Mycobacterium tuberculosis*, across the United States. *Clin Infect Dis* 1999; *29*: 85-92.

96. van Soolingen D, Qian L, de Haas PEW, Douglas TJ, Traore H, Portaels F, *et al*. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia. *J Clin Microbiol* 1995; *33*: 3234-8.

97. Anh DD, Borgdorff MW, Van LN, Lan NT, van Gookom T, Kremer K, *et al*. *Mycobacterium tuberculosis* Beijing genotypes emerging in Vietnam. *Emerg Infect Dis* 2000; *6*: 302-5.

98. Kurepina N, Shashkina E, Sloutsky A, Naroditskaya V, Safonova S, Chernousova L, *et al*. RFLP and DST data on *Mycobacterium tuberculosis* cultures isolated from patients in prison and civilian populations in Russia.

Abstract 1537, European Respiratory Society Congress 2000, Florence, Italy, Aug 30-Sept 3, 2000.

99. Kruuner A, Hoffner SE, Sillastu H, Danilovits M, Levina K, Svenson SB, *et al*. Spread of drug-resistant pulmonary tuberculosis in Estonia. *J Clin Microbiol* 2001; *39*: 3339-45.

100. Diaz R, Kremer K, de Haas PE, Gomez RI, Marrero A, *et al*. Molecular epidemiology of tuberculosis in Cuba outside of Havana, July 1994-June 1995: utility of spoligotyping versus IS6110 restriction fragment length polymorphism. *Int J Tuberc Lung Dis* 1998; *2*: 743-50.

101. Niemann S, Rusch-Gerdes S, Richter E. IS6110 fingerprinting of drug-resistant *Mycobacterium tuberculosis* strains isolated in Germany during 1995. *J Clin Microbiol* 1997; *35*: 3015-20.

102. Kurepina NE, Sreevatsan S, Plikaytis BB, Bifani PJ, Connell ND, Donnelly RJ, *et al*. Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS6110 elements in *Mycobacterium tuberculosis*: non-random integration on the dnaA-dnaN region. *Tuber Lung Dis* 1998; *79*: 31-2.

103. van Crevel R, Nelwan RH, de Lenne W, Veeraragu Y, van der Zanden AG, Amin Z, *et al*. *Mycobacterium tuberculosis* Beijing genotype strains associated with febrile response to treatment. *Emerg Infect Dis* 2001; *7*: 880-3.

104. Soini H, Pan X, Amin A, Graviss EA, Siddiqui A, Masser JM. Characterization of *Mycobacterium tuberculosis* isolates from patients on Houston, Texas, by spoligotyping. *J Clin Microbiol* 2000; *38*: 669-76.

105. Sola C, Devallois A, Horgen L, Maisetti J, Filliol I, Lagrand E, *et al*. Tuberculosis in the Caribbean: using spacer oligonucleotide typing to understand strain origin and transmission. *Emerg Infect Dis* 1999; *5*: 404-14.

106. Nardell E, McInnis B, Thomas B, Weidhaas S. Exogenous reinfection with tuberculosis in a shelter for the homeless. *N Engl J Med* 1986; *315*: 1570-4.

107. Garcia-Garcia ML, Ponce de León A, Jimenez-Corona ME, Jimenez-Corona A, Palaicios-Marnez M, Balandrano-Cambos S, *et al*. Clinical consequences and transmissibility of drug-resistant tuberculosis in southern Mexico. *Arch Intern Med* 2000; *160*: 630-6.

108. van Soolingen D, Borgdorff MW, de Haas PE, Sebek MMGG, Veen J, Dessens M, *et al*. Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 1999; *180*: 726-36.

109. Dye C, Williams BG, Espinal MA, Raviglione MC. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. *Science* 2002; *295*: 2042-6.

110. 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*: 641-64.

111. Riley RL, Mills CC, O'Grady F, Sultan LU, Wittstadt F, Shivpuri DN. Infectiousness of air from a tuberculosis ward.Ultra irradiation of infected air: comparative infectiousness of different patients. *Am Rev Respir Dis* 1962; *85*: 511-25.

112. Li Z, Kelley C, Collins F, Rouse D, Morris S. Expression of katG in *Mycobacterium tuberculosis* is associated with its growth and persistence in mice and guinea pigs. *J Infect Dis* 1998; *177*: 1030-5.

113. Pym AS, Domenech P, Honore N, Song J, Deretic V, Cole ST. Regulation of catalase-peroxidase (*katG*) expression, isoniazid sensitivity and virulence by furA of *Mycobacterium tuberculosis*. *Mol Microbiol* 2001; *40*: 879-89.

114. Pym AS, Saint-Joanis B, Cole ST. Effect of katG mutations on the virulence of *Mycobacterium tuberculosis* and the implication for transmission in humans. *Infect Immun* 2002; *70*: 4955-60.

115. Ambroon CB, Kadlubar FF. Toward an integrated approach to molecular epidemiology. *Am J Epidemiol* 1997; *146*: 912-8.

Reprint requests: Dr Sujatha Narayanan, Assistant Director, Department of Immunology, Tuberculosis Research Centre (ICMR) Mayor V.R. Ramanathan Road, Chetput, Chennai 600031, India
e-mail: sujatha36@hotmail.com