Available online at www.sciencedirect.com
SCIENCE DIRECT®
Infection, Genetics and Evolution 5 (2005) 67–77
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Differentiation of highly prevalent IS6110 single-copy strains of *Mycobacterium tuberculosis* from a rural community in South India with an ongoing DOTS programme

Sulochana D. Das[a], Sujatha Narayanan[a], Lalitha Hari[a], S.L. Hoti[b], R.K. Thangathurai[b], Niruparani Charles[a], K. Jaggarajamma[a], P.R. Narayanan[a,*]

[a]*Tuberculosis Research Centre (ICMR), Mayor V.R. Ramanathan Road, Chetput, Chennai 600 031, India*
[b]*Vector Control Research Centre (ICMR), Pondicherry, India*

Received 14 November 2003; received in revised form 14 June 2004; accepted 16 June 2004
Available online 30 July 2004

Abstract

We have prospectively analysed the DNA fingerprinting of *Mycobacterium tuberculosis* strains in a rural community from high prevalence area in South India with an ongoing DOTS programme. Strains from 451 culture-positive cases, diagnosed during July 1999–December 2000, were fingerprinted initially by both IS6110 and DR probes followed by polymorphic GC-rich repeat sequences (PGRS) typing only on low-copy strains. The results were correlated with selected epidemiological and clinical data. Forty one percent of strains showed single copy of IS6110, which further got differentiated into 62 DR and 27 PGRS patterns. One predominant DR pattern (5B/2) was found in 20% of the low-copy strains and was also involved in clusters. In all, 183 patients out of 451 (40%) were clustered in total 44 clusters when analysed by IS6110 and DR probes. With additional PGRS typing, the number of patients clustered was further reduced to 106 (23%). More number of patients (131) were clustered in IS6110 single-copy group. The maximum number of clusters was found with two or three patients. Only a small percentage (16%) of the patients reported direct epidemiological links while remaining patients might have had indirect links or casual contacts. Thus, a combination of two to three genetic markers is able to differentiate the most endemic strains of *M. tuberculosis* in areas with a high incidence of tuberculosis. The epidemiological data do not suggest any major outbreaks or a hot-spot hypothesis of transmission in this region. Phylogenetic analysis using IS6110, DR and PGRS RFLP (restriction fragment length polymorphism, RFLP) fingerprints showed that isolates exhibited clonal evolutionary pattern. The predominance of certain genotypes and agreement between the phylogenetic trees indicated that these strains were closely related and might have evolved or propagated from the common ancestor.

Keywords: RFLP; *Mycobacterium tuberculosis*; IS6110 single copy; Clonal evolution; DOTS

1. Introduction

Restriction fragment length polymorphism (RFLP) of *Mycobacterium tuberculosis* based on the insertion sequence IS6110 has been proved to be a potent and versatile tool for molecular typing of this organism. Researchers, the world over have used this technique extensively for the discrimination of *M. tuberculosis* strains (Cave et al., 1991; van Soolingen et al., 1993), the identification of transmission chains (Barnes et al., 1997; Raviglione et al., 1995), the investigation of TB outbreaks (Kline et al., 1995; Valway et al., 1998), and distinguishing between reactivation and reinfection (Das et al., 1995; Small et al., 1993a; van Rie et al., 1999). In addition, the technique has been used successfully to prove laboratory cross-contamination (Bauer et al., 1997; Braden et al., 1997a; Small et al., 1993b). The insertion sequence IS6110 is relatively stable within a strain over a period and yet shows great variability in its number and position in the genome among different strains (Cave et al., 1994; van Soolingen et al., 1991). Hence, this technique is being widely used to display variable polymorphisms (Raviglione et al., 1995). The major limitation of the IS6110

* Corresponding author. Tel.: +91-44-2836-2525; fax: +91-44-2836-2528.
E-mail address: nrparanj@md2.vsnl.net.in (P.R. Narayanan).

1567-1348/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2004.06.007

fingerprinting method is its low discriminatory power for isolates with fewer than six copies of IS6110 (Bauer et al., 1999; Yang et al., 2000). Hence, many genotyping methods have been developed to differentiate these low-copy-number isolates. These secondary typing methods are based on different short repetitive DNA sequences associated with some degree of genetic diversity and includes spoligotyping on direct repeat (DR) (Kamerbeek et al., 1997; Cronin et al., 2001), polymorphic GC-rich repeat sequences (PGRS) (Chaves et al., 1996), variable number tandem repeats (VNTR) (Frothingham and Meeker-O'Connell, 1998) and mycobacterial interspersed repetitive units (MIRU) (Supply et al., 2001).

In many large-scale studies world over, people have reported from one-fifth to one-third of *M. tuberculosis* isolates as low-copy-number strains (Bauer et al., 1998; Braden et al., 1997b). We have reported a high percentage of single-copy strains and no copy strains from an endemic urban area of Chennai, South India (Das et al., 1995). Recently, Radhakrishnan et al. (2001) have also reported high percentage of single- and no-copy strains from Kerala, another southern state of India. However, such high percentage of single-copy/no-copy strains was not found in northern part of India as reported recently (Bhanu et al., 2002). When a high percentage of low-copy strains are prevalent in the area, it becomes mandatory to use secondary typing methods to differentiate the low-copy strains, especially the single-copy strains. The most popular secondary typing methods are PGRS fingerprinting using recombinant plasmid pTBN12 and DR spoligotyping.

In this study, we analysed RFLP patterns of *M. tuberculosis* isolates from tuberculosis patients identified in a rural community from high prevalence area in South India where DOTS-based Tuberculosis Control Programme is implemented. The broader objectives were to study the polymorphism of *M. tuberculosis* strains from the rural set-up and the frequency of low-copy strains in this community and to establish tuberculosis transmission by conventional epidemiology. The low-copy-number strains were further subjected to secondary typing methods, and the results were compared with the IS6110 fingerprinting to establish the discriminating power of the DR and PGRS typing methods in this set-up. We also assessed the clonality of *M. tuberculosis* strains by constructing dendrograms based on RFLP fingerprints generated using IS6110, DR and PGRS markers, which has implications on chemotherapeutic strategies, drug resistance, vaccine development and containment of spread of infection.

2. Materials and methods

2.1. Study population

Total 1034 cases were registered during June 1999–December 2000 period of DOTS programme. All the patients included in the study were residents of five Panchayat unions of Tiruvallur district, Tamil Nadu at the time of their diagnosis. Among them, 577 (56%) were culture positive cases of which 523 (90%) were processed for IS6110 and DR fingerprinting. Additional RFLP with PGRS was done only on low-copy strains (292 strains). The final results were available for total 451 cultures (88%). The loss of remaining cultures was due to various reasons like contamination of the original culture, mixed infection or non-availability of the results from either of the probes.

2.2. DNA fingerprinting

Clinical isolates of *M. tuberculosis* were cultured on Lowenstein–Jensen medium for 6–8 weeks. Chromosomal DNA was extracted from the isolates as per the standardized method (van Embden et al., 1993). IS6110 fingerprinting was performed on *Pvu*II-digested DNA samples, as described previously (Das et al., 1995). The IS6110 probe used was a PCR product (245 bp) complementary to the sequence on the right side of the *Pvu*II site within IS6110 and was labeled with the ECL direct labeling system (Amersham Pharmacia Biotech, England).

The secondary typing method like DR typing was performed on *Alu*I-digested DNA samples, according to standard procedures. The DR probe, 36-base pair oligonucleotide (DR-r) (5′-GTT CCG TCC CC TCT CGG GGG TTT TTG GGT C GAC GAC-3′), was labeled using ECL 3′ oligolabeling kit (Amersham Pharmacia Biotech, England) as described in previous protocol (Sahadevan et al., 1995). PGRS RFLP using pTBN12 probe was performed as described by Ross et al., (1992).

2.3. Analysis of genotyping results

IS6110 fingerprints were analysed visually, as described earlier (Das et al., 1995). RFLP patterns were grouped from 1 to 17 based on IS6110 band numbers (total 17 groups). In each group, the patterns were further subdivided based on the position of the bands and were designated from A to Z, depicting the polymorphic types in each group. Fingerprints generated by the DR and PGRS probes were also compared visually and numbered as above. Each new DR and PGRS pattern was further designated with an Arabic numeral.

2.4. Cluster analysis

To compare the diversity of DNA polymorphism associated with each of these genetic markers, genotype clustering of 451 isolates was done initially based on IS6110 genotyping. A cluster was defined as a group of two or more patients with isolates whose fingerprints were identical with respect to both the number and the size of all bands. To evaluate the usefulness of secondary typing methods for subtyping of low-copy-number strains of IS6110, isolates