Tuberculosis 90 (2010) 9–15

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

MOLECULAR ASPECTS

Variable-number tandem repeats typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS*6110* in Thailand

Arunee Thong-On [a,f], Nat Smittipat [b], Tada Juthayothin [b], Hideki Yanai [c,g], Norio Yamada [c,g], Jutaporn Yorsangsukkamol [d], Angkana Chaiprasert [d], Dhanida Rienthong [e], Pamaree Billamas [b], Prasit Palittapongarnpim [a,b,*]

[a] *Department of Microbiology, Faculty of Science, Mahidol University, Rama VI Road, Bangkok, Thailand*
[b] *National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Pathumthani, Thailand*
[c] *TB/HIV Research Project, The Research Institute of Tuberculosis (RIT), Tokyo, Japan*
[d] *Department of Microbiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand*
[e] *Tuberculosis Division, Department of Disease Control, Ministry of Public Health, Bangkok, Thailand*

ARTICLE INFO

Article history:
Received 30 July 2009
Received in revised form
22 October 2009
Accepted 22 October 2009

Keywords:
Mycobacterium tuberculosis
Low-copy-number isolates
Genotyping
Spoligotyping
VNTR typing
IS*6110*-RFLP

SUMMARY

Spoligotyping and variable-number tandem repeats (VNTR) typing have been increasingly used for differentiating *Mycobacterium tuberculosis* strains with low copy numbers of IS*6110*. However, there are few studies comparing their potential to type the strains originating from South and Southeast Asia where many of the isolates have only a few copies, or even single copy, of IS*6110*. Here, we evaluated the genotyping of 187 *M. tuberculosis* isolates harboring 1–6 copies of IS*6110*, available from a population-based study in Chiangrai, northern Thailand during 1998–2000, using spoligotyping and VNTR typing. The low-copy-number isolates constituted about 34% of all *M. tuberculosis* isolated in the province. Discriminating capacities and cluster identification by the two methods were compared with each other and to those obtained by the standard IS*6110*-restriction fragment length polymorphism (RFLP) method. We found that VNTR typing based on the studied 10-loci set generated more distinct patterns (151 patterns) than spoligotyping (54 patterns) and IS*6110*-RFLP (65 patterns). Most of the RFLP- or spoligotyping-defined clusters were subdivided by VNTR typing. Combining IS*6110*-RFLP with VNTR typing produced 164 distinct patterns and 21.9% of clustered isolates whereas the combination of IS*6110*-RFLP and spoligotyping gave 103 different patterns and 59.4% of clustered isolates. Our results confirm the utility of VNTR typing as the secondary method of choice for investigating the epidemiology of *M. tuberculosis* with low copy numbers of IS*6110*.

1. Introduction

Tuberculosis is a global health problem with increased concerns due to the spread of HIV infection and increased drug-resistance.[1] Although studies on the epidemiology of tuberculosis have been done for decades, many still have to be learned, particularly regarding the behaviors of subpopulation of the bacteria. The ability to genotype the bacteria in the last two decades has opened the opportunities to study epidemiological questions that could not be addressed before. Some important questions include the significance of exogenous reinfections as the cause of chronic pulmonary tuberculosis in adults and the atypical epidemiological and pathological properties of the Beijing strains,[2,3] as well as the study of the correlation between virulence, immune response and genotypes.[4–12]

Strains of *Mycobacterium tuberculosis* are typically differentiated by Southern hybridization with IS*6110*, due to its high discriminating power.[13] However, many strains of the bacteria in South and Southeast Asia, including Thailand and India, have only a few copies or even single copy of IS*6110*,[14,15] making the discriminating power of the method too low to be useful for inferring epidemiological linkages. The extent of recent transmission of *M. tuberculosis* in Thailand was therefore conventionally estimated by excluding the data of isolates with low copy numbers of IS*6110*, which unfortunately account for almost half of the population.[15,16]

* Corresponding author. Department of Microbiology, Faculty of Science, Mahidol University, Rama VI Road, Bangkok 10400, Thailand. Tel.: +66 2 5646700x3444; fax: +66 2 6445411.
E-mail addresses: prasit@nstda.or.th, prasit@biotec.or.th (P. Palittapongarnpim).
[f] Present address: College of Medicine and Public Health, Ubon Rajathanee University, Ubon Rajathanee, Thailand.
[g] Present address: Nagasaki University, Nagasaki, Japan.

1472-9792/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tube.2009.10.006

In order to infer epidemiological linkage between the low-copy-number isolates, another strain typing method is needed. Many of the methods have been reported but only a few are practically useful.[17–19] Currently, the most popular ones are spoligotyping and VNTR typing.[20–30] Spoligotyping allows for the detection of sequence variations within the direct repeat (DR) region of the chromosome, whereas VNTR typing is used to assay variations in copy numbers of tandem repeats (as equivalent to micro- or mini-satellite loci in eukaryotes). Both PCR-based typing methods have several advantages over the classical IS*6110*-RFLP method including the generation of portable digital data that can be compared reliably between different laboratories. However, there are few studies comparing the potential of the two methods for typing the low-copy-number isolates.[30–32] Most of the previous studies were done in U.S., where the population genetic structures of the isolates may be different from genotypes of the isolates originating from South and Southeast Asia.[19,21,25,32] The previous report from Southeast Asia included only a small number of the low-copy-number isolates.[33]

Here, we evaluated the genotypes of *M. tuberculosis* harboring 1–6 copies of IS*6110* in Thailand using spoligotyping and VNTR typing methods. Discriminating powers and cluster identification by both methods were compared with each other and to those obtained by the classical IS*6110*-RFLP analysis. We confirmed that VNTR typing had higher capability for differentiating *M. tuberculosis* isolates with low copy numbers of IS*6110* than spoligotyping and IS*6110*-RFLP. We also demonstrated high levels of polymorphisms of some VNTR loci not yet included in the standard VNTR typing set previously proposed by Supply and colleagues.[44]

2. Materials and methods

M. tuberculosis strains. Two hundred and fourteen isolates of *M. tuberculosis* having 1–6 IS*6110*-RFLP bands were included in this study. They represented all available low-copy-number isolates of a population-based study, sponsored by Ministry of Health, Welfare and Labor of Japan through the Research Institute of Tuberculosis (RIT). The samples were all from pulmonary tuberculosis patients and obtained in 6 districts of Chiangrai, the northernmost province in Thailand, during 1998–2000. The bacteria were identified by the Ziehl–Neelson staining method, the niacin production test, the nitrate reduction test and the catalase test at 68 °C. During the period, the low-copy-number isolates, containing 1–5 copies of IS*6110*, contributed to about 34% of the total isolates, while no isolate without IS*6110* or *M. bovis* was identified. *M. tuberculosis* H37Rv and Mt14323 were used as references in VNTR-PCR and IS*6110*-RFLP analyses and were originally from the National Reference Center for Tuberculosis, Canada, and Jan D.A. van Embden, respectively.

DNA extraction and IS6110-RFLP typing. *M. tuberculosis* isolates were grown on Lowenstein–Jensen medium for 3–6 weeks at 37 °C. The chromosomal DNA was extracted by an enzymatic lysis method, digested with *Pvu*II and Southern hybridized by IS*6110* as described earlier.[15] Briefly, 2 μg of DNA of each isolate, or the Mt14323 strain, which was used as a marker, was digested with *Pvu*II and Southern blotted to a nylon filter. The filter was then hybridized with digoxigenin-labeled plasmid pDC73, which contained a segment of IS*6110*. The IS*6110* hybridization patterns were analyzed by Gelcompar II version 1.5 (Applied Maths, Kortrijk, Belgium).

VNTR-PCR analysis. PCR was done for primers targeting at ten loci of VNTR; VNTR0569, 0580, 0960, 1955, 2461, 3192, 3690, 4052, 4120, and 4155 as described previously.[34] These loci were chosen on the basis of their potential to differentiate the isolates with low copy numbers of IS*6110*.[34] The final PCR mixture was composed of 10 mM Tris–HCl, pH 9.0, 50 mM KCl, 0.1% Triton X-100, 1.5 mM $MgCl_2$, 200 μM of each dNTP, 0.5 μM of each primer, 2.5 U of *Taq* polymerase, and 40 ng of DNA template in a total volume of 50 μl. Thermocycling condition included a denaturation step at 95 °C for 1 min, an annealing step at 55–65 °C for 1 min, and an extension step at 72 °C for 2 min. PCR was done for 30 cycles in a PCR gradient machine (Eppendorf, Germany). Amplicons were separated by electrophoresis in 1–2% agarose gels at 100 V and visualized after ethidium bromide staining. The number of repeats for each VNTR locus was calculated from the size of the PCR amplicon in comparison to that of H37Rv.[34]

Spoligotyping analysis. Spoligotyping was performed as previously described.[35] Spoligotyping data were recorded as 15 digits using the octal-code transformation.[36] The 43-digit binary result, representing the 43 spacers, was divided into 14 sets of three digits (spacers 1–42) plus one additional digit (spacer 43). Each three-digit set was transformed into octal code as follows: 000 = 0, 001 = 1, 010 = 2, 011 = 3, 100 = 4, 101 = 5, 110 = 6, and 111 = 7. The final digit remained either 1 or 0.

Analysis of VNTR allelic diversity and genetic relationships. The allelic diversity of each VNTR locus was evaluated by Nei's diversity index (polymorphic information content [PIC]), which is equal to $1 - \sum(\text{allele frequency})^2$.[37] Genetic relationships among the isolates were estimated by the unweighted pair group method with arithmetic averages in PAUP version 4.0 b1 software using the total number of different copy numbers of VNTR loci between pairs of the isolates.

3. Results

Two hundred and fourteen isolates of *M. tuberculosis* were included in this study. The numbers of the isolates with 1–6 copies of IS*6110* were 91, 10, 20, 40, 39, and 14 respectively. Amplifications of all the 10 VNTR loci were successfully done for all the isolates except one, showing no amplicon of VNTR0580 despite several attempts. Based on the amplification data, the allelic diversities of VNTR loci were analyzed. Of the 214 isolates, 187 were also spoligotyped. Due to the lack of the DNA samples and sustainable cultures, the spoligotyping of the remaining 27 isolates was not done. The spoligotyping data of the 187 isolates were compared to the data obtained from VNTR typing and IS*6110*-RFLP.

3.1. Allelic diversity of VNTR loci

The PIC of each VNTR locus was calculated and compared to that of IS*6110*-RFLP (calculated from the diversity of RFLP fingerprinting) for groups of isolates with different copy numbers of IS*6110* (Table 1). As expected, the PIC value of IS*6110*-RFLP typing was particularly low for the single-banded isolates, but became much higher when the copy numbers of IS*6110* increased and was more than 0.9 for isolates with 5–6 copies of IS*6110*.

Seven VNTR loci had high overall PIC (>0.6), whereas the remaining three loci had overall PIC between 0.4 and 0.5. There was only one VNTR, VNTR4155, which by itself was more polymorphic (PIC = 0.903) than the RFLP typing (PIC = 0.835). To increase the discriminating power of VNTR typing, we combined all the 10 VNTR loci together. Combining the results of 10 loci gained the PIC value of more than 0.9 for any group of isolates with different copy numbers of IS*6110* and the overall PIC of 0.992. In all groups, the 10-loci VNTR set was more polymorphic than IS*6110*-RFLP. However, the differences between the PICs of both methods were relatively small when the copy numbers of IS*6110* were 5 or 6.