# Hypervariable loci that enhance the discriminatory ability of newly proposed 15-loci and 24-loci variable-number tandem repeat typing method on *Mycobacterium tuberculosis* strains predominated by the Beijing family

Tomotada Iwamoto[1], Shiomi Yoshida[2], Katsuhiro Suzuki[2], Motohisa Tomita[2], Riyo Fujiyama[3], Noriko Tanaka[3], Yasuto Kawakami[3] & Masahiro Ito[1]

[1]Department of Microbiology, Kobe Institute of Health, Kobe, Japan; [2]National Hospital Organization Kinki-chuo Chest Medical Center, Osaka, Japan; and [3]Kobe City Public Health Center, Kobe, Japan

**Correspondence:** Tomotada Iwamoto, Department of Microbiology, Kobe Institute of Health, Kobe, Japan, Tel.: +81 78 302 6251; fax: +81 –78 –302 0894; e-mail: kx2t-iwmt@asahi-net.or.jp

## Abstract

The newly proposed 15- and 24-loci mycobacterial interspersed repetitive unit (MIRU)–variable-number tandem repeat (VNTR) typing method was evaluated for its ability to differentiate 181 *Mycobacterium tuberculosis* Beijing family strains. Compared with the original 12-loci MIRU–VNTR typing method, the 15-loci system dramatically improved the discriminatory power for Beijing strains; however, large clusters that could be further differentiated by IS*6110* restriction fragment length polymorphism (RFLP) were still obtained. The clonal stability and allelic diversity of a total of 31 VNTR loci were evaluated. VNTRs 3232, 3820, and 4120 were identified as the effective hypervariable VNTR set for the second-line typing of clustered strains following the 15-loci based scheme. Consequently, the discriminatory power of the new scheme (18 loci) equaled that of IS*6110* RFLP.

## Introduction

Tuberculosis remains a major public health threat worldwide, with an estimated one-third of the world population being infected with the *Mycobacterium tuberculosis* complex. A better understanding of tuberculosis epidemiology is essential for controlling existing tuberculosis epidemics and preventing the development of new ones. The introduction of the mycobacterial interspersed repetitive unit (MIRU)–variable-number tandem repeat (VNTR) typing method based on 12 loci (Mazars *et al.*, 2001; Supply *et al.*, 2001) has enabled the study of certain aspects of interest in tuberculosis molecular epidemiological studies, including real-time tracing of ongoing transmission, interlaboratory data comparison, and global database construction (Frothingham & Meeker-O'Connell, 1998; Supply *et al.*, 2001). However, the application of this method has revealed its limited power for discriminating the Beijing family strains of *M. tuberculosis* (Kam *et al.*, 2005; Surikova *et al.*, 2005). As the current worldwide tuberculosis epidemic is, for a significant part, determined by the worldwide spread of the Beijing family of *M. tuberculosis* (Glynn *et al.*, 2002; Kam *et al.*, 2005), the requirement of an effective VNTR typing method for the characterization and differentiation of this family is greatly felt.

Apart from the 12 MIRU loci, the use of various other VNTR loci has been suggested for further improving the discrimination of unrelated stains (Smittipat & Palittapongarnpim, 2000; Roring *et al.*, 2002; Skuce *et al.*, 2002; Smittipat *et al.*, 2005; Surikova *et al.*, 2005; Kam *et al.*, 2006; Nikolayevskyy *et al.*, 2006; Supply *et al.*, 2006). Recently, Supply *et al.* (2006) investigated the resolution power, clonal stability, and technical applicability of 29 VNTR loci using 824 isolates, including worldwide representatives of the main *M. tuberculosis* lineages. They selected 15 loci as the new standard for routine epidemiological discrimination of *M. tuberculosis* isolates and 24 loci as a high-resolution tool for phylogenetic studies. Although this VNTR loci set may currently be the most promising one and is expected to facilitate the sharing of data across the

international community in the field of tuberculosis molecular epidemiology, its feasibility in areas where Beijing genotype strains are prevalent needs to be assessed.

The aim of the present study was to investigate the differentiation ability of the newly proposed 15- and 24-loci VNTR typing method for the Beijing family strains. The allelic diversity and clonal stability of a total of 31 loci were also investigated to identify those that could be applied for the second-line typing of clustered strains nondifferentiated by the newly proposed 15- and 24-loci VNTR typing method.

## Materials and methods

### Bacterial strains

From 2002 to 2004, a total of 238 *M. tuberculosis* clinical strains were obtained from either the Nishi Kobe Medical Center or National Hospital Organization Hyogo-chuo Hospital. These strains were obtained from different patients, including most of the newly diagnosed smear-positive patients at these two hospitals, during the strain-collection period. Of these strains, 235 were used for evaluating the allelic polymorphisms of the VNTR loci. Three other strains were excluded from the analysis because two distinct alleles were observed in VNTR 3232. The clonal stability of each VNTR locus was investigated using 33 isolates from eight different patients with persistent tuberculosis (Table 3) and 18 isolates from six individual outbreak settings (Table 4); these isolates were collected at the Kobe Institute of Health or National Hospital Organization Kinki-chuo Chest Medical Center. Genomic DNA was extracted from these isolates using the Isoplant kit (Nippon Gene, Toyama, Japan), according to the manufacturer's instructions.

### Molecular typing methods

IS*6110* restriction fragment length polymorphism (RFLP) typing and spoligotyping were performed in accordance with standardized protocols (van Embden *et al.*, 1993; Kamerbeek *et al.*, 1997). The original spoligotyping patterns were converted into octal formats (Dale *et al.*, 2001). VNTR analysis was carried out by the amplification of 31 genomic loci in nine different multiplex PCRs and seven different simplex PCRs using the primers listed in Table 1. The PCR reactions for all VNTR loci were performed in a reaction volume of 15 µL containing *c.* 1 ng of DNA template, 1 × GC buffer I (Takara Bio Inc., Shiga, Japan), 0.375 U of Ex Taq (Takara Bio Inc.), 200 µM of each dNTP, and 5 pmol of each primer set. The thermocycling conditions were as follows: 95 °C for 3 min, followed by 35 cycles at 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 1 min, with a final extension step at 72 °C for 7 min. One microliter of PCR product was mixed with 12 µL of formamide and 0.5 µL of

the GeneScan™ 2500 ROX size standard (Applied Biosystems, Foster City, CA). After the denaturation, DNA fragments were separated by capillary electrophoresis using the ABI PRISM 310 Genetic Analyzer (Applied Biosystems). VNTR numbers were scored based on the amplicon size using the GENEMAPPER software (Applied Biosystems). The allelic diversity of each VNTR locus was evaluated by Nei's diversity index, i.e. the polymorphic information content (PIC) (Keim *et al.*, 2000). The level of genotypic diversity based on each VNTR loci set was calculated using the Hunter–Gaston discriminatory index (HGI) (Hunter & Gaston, 1988).

## Results and discussion

Of the 235 strains, 173 had the Beijing family-specific spoligotyping signature and eight had Beijing-like spoligo patterns. We treated these 181 strains as Beijing family strains. As reported by many other researchers (Kam *et al.*, 2005; Surikova *et al.*, 2005; Nikolayevskyy *et al.*, 2006), the 12 original loci (old 12) and their combination with three exact tandem repeat loci (ETR) showed a low discriminatory power for the Beijing strains (Table 2). On the other hand, the newly proposed 15-loci set dramatically improved the discrimination in this study population predominated by Beijing family strains (Table 2). The two largest clusters comprised 34 (old 12-loci profile 223325173533) and 24 strains (old 12-loci profile 221325173533) that were commonly identified as the popular 12-loci pattern in many regions (Mokrousov *et al.*, 2004, 2005; Kremer *et al.*, 2005; Nikolayevskyy *et al.*, 2006) were further discriminated by the new 15-loci set. Although this supports the usefulness of the newly proposed 15-loci set as a first-line method for large-scale, prospective genotyping of *M. tuberculosis*, the present result implied that IS*6110* fingerprinting is still required in many cases for typing clustered isolates (Table 2). About 40% of the clustered strains were subtyped as unique patterns by IS*6110* RFLP. Further, this issue could not be resolved with the 24-loci set even after including nine additional loci (Table 2). The large clusters of unrelated strains would confound the results of outbreak investigations. Therefore, a genotyping technique that provides still higher differentiation is required to allow more effective contact investigation in tuberculosis control programs. Considering this aspect, the allelic diversity of 31 loci including 24-loci set (Supply *et al.*, 2006) was investigated in order to identify those highly polymorphic (Table 1). Further, their added value to the recently proposed standardized 15-loci set was evaluated for further discrimination of the closely related strains of the Beijing family.

Among the 31 loci evaluated, VNTRs 3232, 3820, and 4120 were ranked in the top 3 in the hierarchy of allelic diversities in the Beijing strains that showed a PIC value of

**Table 1.** Locus designations, PCR primer sequences and polymorphic information content (PIC) of each VNTR locus

| Multiplex | Locus | Alias | Copy no. in H37Rv* | All (n = 235) | PCI for: Beijing (n = 181) | Non-Beijing (n = 54) | Primer pairs (5′–3′) (Labeling)† |
|---|---|---|---|---|---|---|---|
| Mix 1 | 580 | MIRU 4 ETR D | 3′ | 0.1973 | 0.0859 | 0.4691 | GCGCGAGAGCCCGAACTGC (FAM) GCGCAGCAGAAACGTCAGC |
| | 2996 | MIRU 26 | 3 | 0.5748 | 0.3830 | 0.7394 | TAGGTCTACCGTCGAAATCTGTGAC CATAGGCGACCAGGCGAATAG (HEX) |
| | 802 | MIRU 40 | 1 | 0.4988 | 0.3268 | 0.7517 | GGGTTGCTGGATGACAACGTGT (NED) GGGTGATCTCGGCGAAATCAGATA |
| Mix 2 | 960 | MIRU 10 | 3 | 0.5649 | 0.4189 | 0.7936 | GTTCTTGACCAACTGCAGTCGTCC GCCACCTTGGTGATCAGCTACCT (FAM) |
| | 1644 | MIRU 16 | 2 | 0.411 | 0.3104 | 0.6097 | TCGGTGATCGGGTCCAGTCCAAGTA CCCGTCGTGCAGCCCTGGTAC (HEX) |
| | 3192 | MIRU 31 | 3 | 0.5603 | 0.3215 | 0.537 | ACTGATTGGCTTCATACGGCTTTA GTGCCGACGTGGTCTTGAT (NED) |
| Mix 3 | 154 | MIRU 2 | 2 | 0.0252 | 0.0000 | 0.1049 | TGGACTTGCAGCAATGGACCAACT TACTCGGACGCCGGCTCAAAAT (FAM) |
| | 2531 | MIRU 23 | 6 | 0.3108 | 0.1761 | 0.5974 | CTGTCGATGGCCGCAACAAAACG (HEX) AGCTCAACGGGTTCGCCCTTTTGTC |
| | 4348 | MIRU 39 | 2 | 0.4542 | 0.2212 | 0.4973 | CGCATCGACAAACTGGAGCCAAAC CGGAAACGTCTACGCCCCACACAT (NED) |
| Mix 4 | 2059 | MIRU 20 | 2 | 0.0737 | 0.0219 | 0.2257 | TCGGAGAGATGCCCTTCGAGTTAG (FAM) GGAGACCGCGACCAGGTACTTGTA |
| | 2687 | MIRU 24 | 1 | 0.0252 | 0.0000 | 0.1049 | CGACCAAGATGTGCAGGAATACAT GGGCGAGTTGAGCTCACAGAA (HEX) |
| | 3007 | MIRU 27 QUB-5 | 3 | 0.0976 | 0.1148 | 0.0364 | TCGAAAGCCTCTGCGTGCCAGTAA GCGATGTGAGCGTGCCACTCAA (NED) |
| Mix 5 | 2165 | ETR A | 3 | 0.4407 | 0.1470 | 0.5535 | AAATCGGTCCCATCACCTTCTTAT (FAM) CGAAGCCTGGGGTGCCCGCGATTT |
| | 2461 | ETR B | 3 | 0.122 | 0.0326 | 0.3697 | GCGAACACCAGGACAGCATCATG (HEX) GGCATGCCGGTGATCGAGTGG |
| | 577 | ETR C | 4 | 0.0742 | 0.0219 | 0.2298 | GTGAGTCGCTGCAGAACCTGCAG (NED) GGCGTCTTGACCTCCACGAGTG |
| Mix 6 | 424 | Mtub04 | 2 (51 × 2 + 30) | 0.5982 | 0.4587 | 0.4712 | CTTGGCCGGCATCAAGCGCATTATT GGCAGCAGAGCCCGGGATTCTTC (FAM) |
| | 2401 | Mtub30 | 2 (58 × 1 + 53) | 0.4926 | 0.4034 | 0.5796 | CTTGAAGCCCCGGTCTCATCTGT (NED) CGAAGCCTGGGGTGCCCGCGATTT |
| | 3690 | Mtub39 | 5* | 0.3574 | 0.1859 | 0.7346 | CGGTGGAGGCGATGAACGTCTTC (HEX) TAGAGCGGCACGGGGGAAAGCTTAG |
| Mix 7 | 2347 | Mtub29 | 4 (57 × 3 + 48) | 0.0982 | 0.0434 | 0.2627 | GCCAGCCGCCGTGCATAAACCT (FAM) AGCCACCCGGTGTGCCTTGTATGAC |
| | 3171 | Mtub34 | 3 (54 × 2 + 51) | 0.0582 | 0.0647 | 0.0364 | GGTGCGCACCTGCTCCAGATAA (HEX) GGCTCTCATTGCTGGAGGGTTGTAC |
| Mix 8 | 1451 | QUB-1451 | 4 (57 × 3 + 46) | 0.3077 | 0.0327 | 0.631 | CGGTGGTGTGCACGGTAA (FAM) AGCCCAGAAGGCAACAATG |
| | 1895 | QUB-1895 | 4 (57 × 4 + 11) | 0.4285 | 0.3637 | 0.5919 | GCGTGACTGCCTCGAACA (HEX) CAATCACACCCAGCCCAAT |
| Mix 9 | 1955 | Mtub21 | 1 (57 × 1 + 34) | 0.5835 | 0.3927 | 0.5988 | AGACGTCAGATCCCAGTT (NED) ACCCGACAACAAGCCCA |
| | 4156 | QUB-4156 | 3 (59 × 2 + 51) | 0.6899 | 0.6106 | 0.6646 | CGTCCGAGCGACATCAC (HEX) AGGATCGAGCGGTCCAG |
| Single | 2163a | QUB-11A | 2 (62 × 2 + 8) | 0.7922 | 0.6854 | 0.8745 | CGTGATGTTGATCGGGATGT (HEX) ACCCTGGAGTCTGGCATC |
| Single | 2163b | QUB-11B | 5 (69 × 5 + 10) | 0.8309 | 0.7716 | 0.7476 | CCGATGTAGCCCGTGAAGA (NED) AGGGTCTGATTGGCTACTCA |
| Single | 3336 | QUB-3336 | 8* | 0.6619 | 0.4870 | 0.8848 | CCACCGCGATCCAGGAAT (FAM) CGGGATTCACCACGATCTC |
| Single | 3232 | QUB-3232 | 4 (56 × 3 + 48) | 0.9051 | 0.8799 | 0.7901 | CCCCAGCCTTACGACTGA (FAM) GTCGGGCTTGGTGAAGG |

**Table 1.** Continued.

| Multiplex | Locus | Alias | Copy no. in H37Rv* | All (n = 235) | PCI for: Beijing (n = 181) | Non-Beijing (n = 54) | Primer pairs (5′–3′) (Labeling)[†] |
|---|---|---|---|---|---|---|---|
| Single | 3820 | VNTR 3820 | 3 (57 × 3 + 26) | 0.8609 | 0.8000 | 0.7963 | ACCTTCATCCTTGGCGAC (HEX) TGCGCGGTGAATGAGACG |
| Single | 4052 | QUB-26 | 5 (111 × 5 + 0) | 0.7808 | 0.7409 | 0.7977 | GAGGTATCAACGGGCTTGT (FAM) GAGCCAAATCAGGTCCGG |
| Single | 4120 | VNTR 4120 | 2 (57 × 2 + 23) | 0.9173 | 0.9022 | 0.6749 | GTTCACCGGAGCCAACC (NED) GAGGTGGTTTCGTGGTCG |

*Except old 12-loci, ETR and two loci indicated with asterisk, repeat unit size (bp) × repeat no. + partial repeat size (bp) which were used for copy no. determination in H37Rv were indicated in parenthesis. The repeat number at VNTR 3336 and VNTR 3690 were found to be different from the ones deduced by the genomic sequence (five and two copies, respectively).

[†]References for the primer sequences are as follows: old 12-loci, Supply *et al*. (2001); ETR A–C, Frothingham & Meeker-O'Conell (1998); VNTR 424, 2401, 3690, 2347, 3171, Supply *et al*. (2006); VNTR 1955, 3820, 4052, 4120, Smittipat *et al*. (2005); VNTR 1451, 4156, 3336, 3232, Roring *et al*. (2002); VNTR 2163a, 2163b, Skuce *et al*. (2002). The primers of VNTR 1451, 4156, 2163a, 2163b, 3336, 3232 were redesigned in this study.

more than 0.8. Of these, VNTR 3232 has been well characterized by many other researchers as a highly polymorphic locus (Kremer *et al*., 2005; Kam *et al*., 2006; Nikolayevskyy *et al*., 2006). The other two loci VNTRs 3820 and 4120 were recently identified by Smittipat *et al*. (2005) as potentially highly polymorphic loci, but they were not evaluated in the study conducted for selecting the standardized VNTR loci (Supply *et al*., 2006). So far, this study is the first to confirm that VNTRs 3820 and 4120 are highly polymorphic loci using many Beijing strains from the same geographic region. When these three loci were used as an adjunct to the 15-loci set, the discriminatory power for the strains used in this study reached almost the maximum resolution that was achieved using the 31-loci set (Table 2). Moreover, the resolution was equivalent to that achieved with IS*6110* RFLP, which is the current gold standard typing method for *M. tuberculosis*. The results support the use of an additional set of hypervariable loci to supplement and enhance the ability of the 15-loci VNTR typing method for discriminating strains in a population predominated by the Beijing family.

In the study for standardization of MIRU–VNTR typing (Supply *et al*., 2006), serious concerns were raised regarding the use of hypervariable loci for routine epidemiological discrimination; these included the absence of PCR products, double alleles in a locus, uninterpretable PCR patterns, and presumably low stability in long-term epidemiological investigations. Therefore, the highly polymorphic VNTRs 3232, 3336, and 2163a were excluded from those selected for the standardized MIRU–VNTR typing in order to minimize the probability of incorrect exclusion of an epidemiological link. However, it was realized that in the present population predominated by strains of the Beijing family, the clustered strains determined using the newly proposed standardized set need to be subtyped to ensure direct clonal transmission; such subtypes would be required in outbreak investigations that usually require labor-intensive careful contact investigation.

For this purpose, the hypervariable loci set, i.e. VNTRs 3232, 3820, and 4120, showed tremendous potential for the discrimination of unrelated strains when it was used as a second-line typing method after the 15-loci VNTR typing method (Table 2). Therefore, an attempt was made to clarify the above-mentioned issues in relation to hypervariable loci for ensuring their reliability in the second typing. As Supply *et al*. (2006) mentioned, the alleles of hypervariable loci are distributed in a wide range of repeat numbers that are frequently amplified as large-size fragments (> 1000 bp). Such large-size fragments hinder the accurate scoring of alleles. To resolve the issue, the stutter peaks could be considered as efficient indicators for accurately determining the repeat number (Fig. 1). By combining base size calling and stutter peak counting, the allele scores for the three hypervariable loci could be obtained unambiguously with high reproducibility even when the fragment size exceeded 1000 bp. For stutter peak counting, the peak position and allele score of the H37Rv strain should be used as the standard. Confirmatory agarose gel electrophoresis of the PCR products obtained using nonlabeled primers is recommended to confirm the accuracy of scoring for these hypervariable loci.

The clonal stability of the individual 31 VNTR loci was evaluated using 33 serial isolates obtained from eight different patients with persistent disease over a period of 53 months (Table 3) and 18 isolates obtained from six individual outbreak settings (Table 4). Differences in the 31-loci VNTR patterns between the initial and follow-up isolates were observed in patients 3 (Beijing strain), 4 (non-Beijing strain), and 6 (non-Beijing strain). All changes were single-locus variants in VNTR 3232 in patients 3 and 4 and VNTR 3690 in patient 6. In patient 3, the IS*6110* RFLP patterns also changed. The study of strains spread by person-to-person transmission (Table 4) revealed no allele changes in the 31 VNTR loci, but an RFLP pattern change was observed in 1 strain. In the clonal stability study conducted by Supply

**Table 2.** Discriminatory power of various combination of VNTR loci

| Strains typed (*n*) | Cluster size | No. of clusters and unique types by MIRU–VNTR locus set | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Old 12 | Old 12+ETR | New 15 | New 24 | New 15+HV3* | All 31 loci | RFLP† |
| All (235) | 2 | 17 | 19 | 13 | 13 | 13 | 12 | 13 |
| | 3 | 8 | 6 | 4 | 4 | 10 | 8 | 8 |
| | 4 | 4 | 3 | 5 | 5 | 1 | 1 | 2 |
| | 5 | 4 | 3 | 1 | 4 | 0 | 0 | 0 |
| | 6 | 1 | 2 | 1 | 0 | 0 | 0 | 0 |
| | 7 | 1 | 0 | 2 | 2 | 0 | 0 | 0 |
| | 8 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 9 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 22 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 24 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 32 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 34 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Unique | 62 | 78 | 135 | 143 | 175 | 183 | 170 |
| | HGI | 0.9634 | 0.9684 | 0.9932 | 0.9950 | 0.9982 | 0.9985 | 0.9981 |
| Beijing (181) | 2 | 13 | 14 | 11 | 11 | 10 | 10 | 11 |
| | 3 | 6 | 4 | 3 | 3 | 8 | 7 | 6 |
| | 4 | 3 | 2 | 4 | 4 | 0 | 0 | 2 |
| | 5 | 3 | 3 | 1 | 3 | 0 | 0 | 0 |
| | 6 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 7 | 0 | 0 | 1 | 2 | 0 | 0 | 0 |
| | 8 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 9 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 22 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 24 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 32 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 34 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Unique | 38 | 50 | 99 | 105 | 137 | 140 | 133 |
| | HGI | 0.9411 | 0.9486 | 0.9905 | 0.9929 | 0.9980 | 0.9982 | 0.9975 |

*Three hypervariable loci set consisting VNTR 3232, 3820, and 4120.

†Seven isolates which were clustered by either one or two copies of IS *6110* were excluded from the calculation in all 235 isolates.
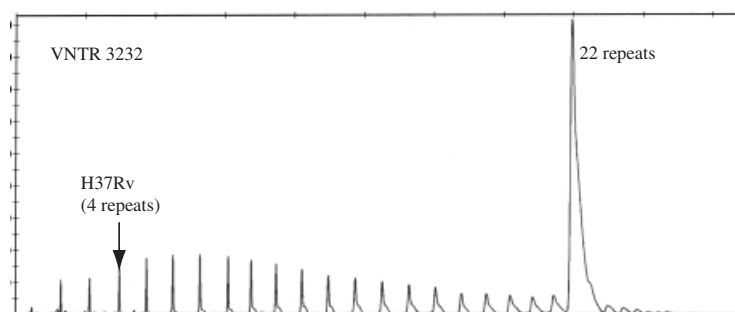


**Fig. 1.** Example of stutter peaks on VNTR 3232 that were judged as 22 repeats.

*et al.* (2006), allele changes were identified in six loci among the 24 robust loci. These results suggested that the hypervariable loci (VNTRs 3232, 3820, and 4120) were not extraordinarily unstable when compared with other loci and those of the IS*6110* RFLP pattern. Overall, the clonal stability of the studied 31 loci could be considered as equivalent to IS*6110* RFLP and thus reasonably stable for tracing the transmission of the targeted clone; however, caution should be exercised in the case of single locus variants as in the case of minor changes in IS*6110* RFLP patterns. The relatively low clonal stability of VNTRs 3232 and 3336 in the previous study (Supply *et al.*, 2006) might be due to the technical difficulty of allele scoring for such hypervariable loci or may be simply due to the differences in the strain population. The method of using stutter peaks appeared to minimize the potential technical error of mis-scoring.

While the clonal stability of each of the 31 VNTR loci was being evaluated, interesting insights were obtained into the interpretation of double peaks for a locus and the absence of PCR products. The simultaneous occurrence of two

**Table 3.** Stability of VNTRs in serial patient isolates

| Patient no. | Isolate no. | Time span (months) | Spoligotyping octal code | No. of IS6110 bands | MIRU*–ETR†–VNTR‡ |
|---|---|---|---|---|---|
| 1 | 169 – OK | 0 | 000000000003771 | 13 | 223325173533–424–57A24I753C844343 |
| | 170 – OK | 13 | 000000000003771 | 13 | 223325173533–424–57A24I753C844343 |
| | 171 – OK | 40 | 000000000003771 | 13 | 223325173533–424–57A24I753C844343 |
| 2 | 185 – SY | 0 | 000000000003771 | 11 | 221325173533–424–97822K753CC54343 |
| | 186 – SY | 47 | 000000000003771 | 11 | 221325173533–424–97822K753CC54343 |
| | 187 – SY | 52 | 000000000003771 | 11 | 221325173533–424–97822K753CC54343 |
| 3 | 188 – WN | 0 | 000000000003771 | 9 | 221325143524–424–97724I753EB44343 |
| | 189 – WN | 9 | 000000000003771 | 9 | 221325143524–424–97724I753EB44343 |
| | 190 – WN | 18 | 000000000003771 | 9 | 221325143524–424–97724J753EB44343 |
| | 191 – WN | 24 | 000000000003771 | 9+1§ | 221325143524–424–97724J753EB44343 |
| | 192 – WN | 30 | 000000000003771 | 9 | 221325143524–424–97724J753EB44343 |
| | 193 – WN | 36 | 000000000003771 | 9 | 221325143524–424–97724J753EB44343 |
| | 194 – WN | 43 | 000000000003771 | 9 | 221325143524–424–97724J753EB44343 |
| 4 | 195 – YY | 0 | 677777607760771 | 13 | 323226163321–223–44724CB325431243 |
| | 196 – YY | 7 | 677777607760771 | 13 | 323226163321–223–44724(C/D)B325431243 |
| | 197 – YY | 28 | 677777607760771 | 13 | 323226163321–223–44724(C/D)B325431243 |
| 5 | 207 – SM | 0 | 000000000003771 | 16 | 223425153533–424 – 85222C7536B34343 |
| | 208 – SM | 22 | 000000000003771 | 16 | 223425153533–424–85222C7536B34343 |
| | 209 – SM | 28 | 000000000003771 | 16 | 223425153533–424–85222C7536B34343 |
| | 210 – SM | 34 | 000000000003771 | 16 | 223425153533–424–85222C7536B34343 |
| | 211 – SM | 37 | 000000000003771 | 16 | 223425153533–424–85222C7536B34343 |
| 6 | 212 – IT | 0 | 777737777760771 | 11 | 223326153311–323–143445D315222643 |
| | 213 – IT | 4 | 777737777760771 | 11 | 223326153311–323–143445D315222643 |
| | 214 – IT | 16 | 777737777760771 | 11 | 223326153311–323–143445D315222643 |
| | 215 – IT | 23 | 777737777760771 | 11 | 223326153311–323–143445D315222443 |
| | 216 – IT | 40 | 777737777760771 | 11 | 223326153311–323–143445D315222443 |
| | 217 – IT | 44 | 777737777760771 | 11 | 223326153311–323–143445D315222643 |
| | 218 – IT | 51 | 777737777760771 | 11 | 223326153311–323–143445D315222643 |
| | 219 – IT | 53 | 777737777760771 | 11 | 223326153311–323–143445D315222643 |
| 7 | K1 – SH | 0 | 777637777777771 | 15 | 228225173733–524–P86429123B342143 |
| | K2 – SH | 12 | 777637777777771 | 15 | 228225173733–524–P86429123B342143 |
| 8 | K3 – KK | 0 | 000000000003771 | 18 | 222325193543–424–86624C734LB44343 |
| | K4 – KK | 24 | 000000000003771 | 18 | 222325193543–424–86624C734LB44343 |

*Order of MIRU loci: 2, 4, 10, 16, 20, 23, 24, 26, 27, 31, 39, 40.
†Order of ETR loci: A, B, C.
‡Order of VNTR loci: VNTRs 2163a, 2163b, 4052, 1451, 1895, 3232, 3336, 4156, 1955, 3820, 4120, 424, 2401, 3690, 2347, 3171.
§One additional band was found.
The MIRU–ETR–VNTR profiles consists of digits (when nine or less repeats were detected at a particular locus) and letters; the letter A represents 10 repeats, the letter B represents 11 repeats, and so on.
VNTR profile which differed from the first isolates were enclosed by square.

different alleles in VNTR 3232 (12 repeats shown as C and 13 repeats shown as D) was clearly detected in the second (196-YY) and third (197-YY) isolates obtained from patient 4 (Table 3). The first isolate (195-YY) from this patient displayed only repeat C in this locus; thus, the double band would indicate the appearance/coexistence of a clonal variant strain that differs from the parent strain. The clonal variant strain coexisted in the patient at least for 21 months. The same change, i.e. the coexistence of a clonal variant strain, was reported in *Mycobacterium bovis* (Allix *et al.*, 2006). In the presence of such evidence, the presence of dual peaks for a locus may be interpreted as the coexistence of a

clonal variant if dual peaks were observed only for a single locus as is the case in this study.

As shown in previous studies (Kam *et al.*, 2006; Supply *et al.*, 2006), PCR failure was found for VNTRs 2162a (nine out of 235 strains, 3.8%), 2163b (7/235, 3.0%), 3336 (2/235, 0.9%), and 3820 (3/235, 1.3%). These PCR failures were repeated even though the tests were performed under different PCR conditions and with different primer sets. Interestingly, consistent PCR failures were observed in VNTRs 2163a and 3336 from three epidemiologically linked strains (hospital contact in Table 4). Therefore, it was assumed that no PCR product would reflect the

**Table 4.** Stability of VNTRs in strains that undergo a person-to-person transmission

| Type of transmission | No. of patients | Spoligotyping octal code | No. of IS*6110* bands (no. of patients) | MIRU–ETR–VNTR* |
|---|---|---|---|---|
| Family contact | 2 | 000000000003771 | 13 (2) | 243325163533–424–83822J743GG42343 |
| Family and social contact | 4 | 000000000003771 | 10 (4) | 221325163533–424–978248733EC44343 |
| Work contact | 3 | 000000000003771 | 18 (3) | 222325193543–424–86624C734LB44343 |
| Work contact | 3 | 000000000003771 | 14 (3) | 223325173524–224–83724C743E542543 |
| Social contact | 3 | 000000000003771 | 12 (2), 13 (1)[†] | 223425173533–424–87222A753C934343 |
| Hospital contact | 3 | 741737777760371 | 2 | 215125113322–314–$28245$428424343 |

*No PCR amplification was designated as $ repeat.
[†]One additional band was found.

characteristics of certain strains. Although the cause of the PCR failures could not be determined, the presence of an unexpectedly large number of repeat units or deletion of this region is suspicious. Rather than discarding the data that did not yield any PCR product, such data would be useful for indicating the characteristics that could be used as reference information for genotyping. Under the conditions where no PCR amplification was confirmed after repeated PCR attempts and DNA integrity was assured based on PCR at other loci, no PCR amplification was designated as $ repeat in this study.

In conclusion, hypervariable loci can be potentially applied for VNTR typing at least as a second-line typing when more detailed subtyping is required in outbreak investigations and for confirmation of laboratory cross-contamination. The proposed standardized MIRU–VNTR typing using the new 15- and 24-loci sets will definitely contribute to better sharing of data across the international community in the field of tuberculosis molecular epidemiology. It was confirmed that this method improves the discrimination of Beijing family strains, which has been the main drawback of the original 12-loci MIRU–VNTR typing method. Thus, the new system could be useful for population-based epidemiological studies in populations predominated by Beijing family strains. Besides the standardized loci set, the use of hypervariable loci such as VNTRs 3232, 3820, and 4120 may be recommended as a second-line typing method if more detailed genotyping is necessary to differentiate Beijing genotype strains.

## Acknowledgements

## References

Allix C, Walravens K, Saegerman C, Godfroid J, Supply P & Fauville-Dufaux M (2006) Evaluation of the epidemiological relevance of variable-number tandem-repeat genotyping of *Mycobacterium bovis* and comparison of the method with IS*6110* restriction fragment length polymorphism analysis and spoligotyping. *J Clin Microbiol* **44**: 1951–1962.

Dale JW, Brittain D, Cataldi AA *et al.* (2001) Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: recommendations for standardised nomenclature. *Int J Tuberc Lung Dis* **5**: 216–219.

Frothingham R & Meeker-O'Connell WA (1998) Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology* **144**: 1189–1196.

Glynn JR, Bifani PJ, Kremer K & van Soolingen D (2002) Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg Infect Dis* **8**: 843–849.

Hunter PR & Gaston MA (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* **26**: 2465–2466.

Kam KM, Yip CW, Tse LW, Wong KL, Lam TK, Kremer K, Au BK & van Soolingen D (2005) Utility of mycobacterial interspersed repetitive unit typing for differentiating multidrug-resistant *Mycobacterium tuberculosis* isolates of the Beijing family. *J Clin Microbiol* **43**: 306–313.

Kam KM, Yip CW, Tse LW, Leung KL, Wong KL, Ko WM & Wong WS (2006) Optimization of variable number tandem repeat typing set for differentiating *Mycobacterium tuberculosis* strains in the Beijing family. *FEMS Microbiol Lett* **256**: 258–265.

Kamerbeek J, Schouls L, Kolk A *et al.* (1997) Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* **35**: 907–914.

Keim P, Price LB, Klevytska AM, Smith KL, Schupp JM, Okinaka R, Jackson PJ & Hugh-Jones ME (2000) Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J Bacteriol* **182**: 2928–2936.

Kremer K, Au BK, Yip PC, Skuce R, Supply P, Kam KM & van Soolingen D (2005) Use of variable-number tandem-repeat typing to differentiate *Mycobacterium tuberculosis* Beijing family isolates from Hong Kong and comparison with IS*6110*

restriction fragment length polymorphism typing and spoligotyping. *J Clin Microbiol* **43**: 314–320.

Mazars E, Lesjean S, Banuls AL, Gilbert M, Vincent V, Gicquel B, Tibayrenc M, Locht C & Supply P (2001) High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. *Proc Natl Acad Sci USA* **98**: 1901–1906.

Mokrousov I, Narvskaya O, Limeschenko E, Vyazovaya A, Otten T & Vyshnevskiy B (2004) Analysis of the allelic diversity of the mycobacterial interspersed repetitive units in *Mycobacterium tuberculosis* strains of the Beijing family: practical implications and evolutionary considerations. *J Clin Microbiol* **42**: 2438–2444.

Mokrousov I, Ly HM, Otten T, Lan NN, Vyshnevskyi B, Hoffner S & Narvskaya O (2005) Origin and primary dispersal of the *Mycobacterium tuberculosis* Beijing genotype: clues from human phylogeography. *Genome Res* **15**: 1357–1364.

Nikolayevskyy V, Gopaul K, Balabanova Y, Brown T, Fedorin I & Drobniewski F (2006) Differentiation of tuberculosis strains in a population with mainly Beijing-family strains. *Emerg Infect Dis* **12**: 1406–1413.

Roring S, Scott A, Brittain D, Walker I, Hewinson G, Neill S & Skuce R (2002) 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* **40**: 2126–2133.

Skuce RA, McCorry TP, McCarroll JF, Roring SM, Scott AN, Brittain D, Hughes SL, Hewinson RG & Neill SD (2002) Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. *Microbiology* **148**: 519–528.

Smittipat N & Palittapongarnpim P (2000) Identification of possible loci of variable number of tandem repeats in *Mycobacterium tuberculosis*. *Tuber Lung Dis* **80**: 69–74.

Smittipat N, Billamas P, Palittapongarnpim M, Thong-On A, Temu MM, Thanakijcharoen P, Karnkawinpong O & Palittapongarnpim P (2005) Polymorphism of variable-number tandem repeats at multiple loci in *Mycobacterium tuberculosis*. *J Clin Microbiol* **43**: 5034–5043.

Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D & Locht C (2001) Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *J Clin Microbiol* **39**: 3563–3571.

Supply P, Allix C, Lesjean S *et al.* (2006) Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* **44**: 4498–4510.

Surikova OV, Voitech DS, Kuzmicheva G, Tatkov SI, Mokrousov IV, Narvskaya OV, Rot MA, van Soolingen D & Filipenko ML (2005) Efficient differentiation of *Mycobacterium tuberculosis* strains of the W-Beijing family from Russia using highly polymorphic VNTR loci. *Eur J Epidemiol* **20**: 963–974.

van Embden JD, Cave MD, Crawford JT *et al.* (1993) Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* **31**: 406–409.

## Supplementary material

The following supplementary material is available for this article:

Figure S1. IS*6110* RFLP profiles and dendrogram of all clustered strains by either IS*6110* RFLP or 15-loci VNTR

Table S2. All other 16 VNTR profiles on 235 strains

Table S3. Spolipotyping results on 235 strains

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1574.6968.2007.00658.x (This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.