Phylogeographical particularity of the *Mycobacterium tuberculosis* Beijing family in South Korea based on international comparison with surrounding countries

Hee Yoon Kang, Takayuki Wada, Tomotada Iwamoto, Shinji Maeda, Yoshiro Murase, Seiya Kato, Hee Jin Kim and Young Kil Park

Correspondence
Young Kil Park
ypark7@empal.com

1Korean Institute of Tuberculosis, Korean National Tuberculosis Association, Seoul, Republic of Korea
2Department of Microbiology, Osaka City Institute of Public Health and Environmental Sciences, Osaka, Japan
3Department of Microbiology, Kobe Institute of Health, Kobe, Japan
4Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo, Japan

To understand the domestic population structure of *Mycobacterium tuberculosis* clinical isolates in the Republic of Korea, we genotypically analysed 80 isolates obtained from various geographical origins in the country. Of these, 64 (80.0 %) isolates were identified as Beijing family strains. It is particularly interesting that their phylogenetic classification, based on the ancient/modern separation and the presence/absence of the genomic region RD181, revealed a majority of the ancient (RD181 +) subfamily in the population. The 15 loci of variable number of tandem repeat(s) of mycobacterial interspersed repetitive units (15-MIRU-VNTR) were also analysed. Combination with the previous VNTR data reported from surrounding countries revealed that the topology of the minimum spanning tree was linked tightly not to the geographical origins of the patients but to the phylogenetic characteristics of the isolates. These results show that the phylogeographical distribution of the *M. tuberculosis* Beijing family around far-eastern Asia could be estimated using international accumulation and comparison of VNTR genotyping data.

INTRODUCTION

The Beijing family, a lineage of *Mycobacterium tuberculosis*, is well known for its highly endemic prevalence around East Asian countries and as a causative agent of tuberculosis (TB) (van Soolingen *et al.*, 1995). In the Republic of Korea (South Korea), TB is still a major public health concern, with 34 157 (70.3 per 100 000) registered new TB patients in 2008 and 2376 deaths attributable to TB in 2007 (Korea Centers for Disease Control & Prevention, 2008). More than 70 % of *M. tuberculosis* strains isolated from Korean pulmonary TB patients belong to the Beijing family. ‘K strain’ (Park *et al.*, 2000), one of the Beijing family strains, has been reported as the cause of a severe outbreak of TB in South Korea (Kim *et al.*, 2001). Although their phylogenetic position in the Beijing family lineage has been unclear, Shamputa *et al.* (2010) reported recently that K strains show genetic diversity by some genotyping methods, even in isolates obtained from a single hospital.

The detailed phylogenetic variation of the Beijing family has been unveiled by various genetic markers such as single nucleotide polymorphisms, regions of difference and variable number of tandem repeat (VNTR) loci (Filliol *et al.*, 2006; Tsolaki *et al.*, 2005; Wada & Iwamoto, 2009). A notable characteristic of the lineage is the insertion of IS6110 in a genomic region named NTF (Mokrousov *et al.*, 2005; Plikaytis *et al.*, 1994). This phylogenetic marker can classify the Beijing family into two distinct sublineages: ancient and modern. The modern Beijing sublineage has been more predominant than the ancient sublineage throughout the world, including countries surrounding South Korea (Bifani *et al.*, 2002; Dou *et al.*, 2008; Kremer *et al.*, 2009; Mokrousov *et al.*, 2005, 2006; van Soolingen & Kremer, 2009). However, Wada *et al.* (2009) reported that the ancient Beijing sublineage has been observed to be exceptionally predominant in Japan.

**Abbreviations:** MIRU-VNTR, variable number of tandem repeats of mycobacterial interspersed repetitive units; MST, minimum spanning tree; TB, tuberculosis.

Supplementary tables giving information on and genotypic profiles of the isolates analysed in this study are available with the online version of this paper.
The goal of this study was to elucidate the phylogenetic distinctiveness and the genetic diversity of Beijing family strains around the Republic of Korea, including K family strains, for international comparison with strains from surrounding countries. For this purpose, we analysed a collection of *M. tuberculosis* isolates obtained from public health centres in the country using well-known phylogenetic markers and an international standard set of VNTR loci (Supply *et al.*, 2006).

**METHODS**

**Bacterial isolates.** Eighty *M. tuberculosis* clinical isolates were analysed in this study. They were randomly selected from smear-positive and culture-positive pulmonary TB patients who were under 46 years of age with no epidemiological links during 2006. All patients were primary cases. Information such as geographical origin, age and sex of the patients, date of isolation and drug susceptibility for each of the 80 isolates is presented in Supplementary Table S1 in JMM Online. The geographical distribution of the population is presented in Fig. 1.

**Identification of Beijing family strains and their phylogenetic subdivision.** The Beijing family strains were defined by the deletion of RD207 in this study (Tsolaki *et al.*, 2004). The deletion of RD207 corresponds with the absence of signals 1–34 in spacer oligonucleotide (spoligo) genotyping, which is the most standard definition of the lineage (van Soolingen *et al.*, 1995). The presence or absence of RD207 was analysed using PCR according to a previous report (Warren *et al.*, 2004).

The classification of the ancient and modern subfamilies of the 64 Beijing family isolates was determined according to a previous report (Wada *et al.*, 2009). The presence or absence of RD181 was also verified as described previously (Tsolaki *et al.*, 2004). The genomic deletion of RD181 has been considered to have occurred during evolution of the ancient Beijing sublineage (Tsolaki *et al.*, 2004; Maeda *et al.*, 2010). The sets of primer sequences were described in these previous reports. They were designed to detect insertions or deletions of regions by the difference in sizes of amplified DNA fragments.

**IS6110 DNA fingerprinting.** For all 80 isolates, DNA isolation and IS6110 RFLP typing were performed as described previously (Park *et al.*, 2000). The IS6110 RFLP patterns were compared with UPGMA using the Dice coefficient to find RFLP clusters (GelCompar v. 5.1; Applied Maths). An RFLP cluster was defined by the identification of two or more isolates with identical band patterns. K family strains were identified according to the previously reported definition (Kim *et al.*, 2001; Park *et al.*, 2000).

**VNTRs.** The standard 15 VNTR loci for routine epidemiological discrimination (hereinafter, 15-MIRU-VNTR) were also analysed for all 80 isolates (Supply *et al.*, 2006). Their copy number was calculated from their size and assigned according to the number of repeats for each locus, and in agreement with published allelic tables (Iwamoto *et al.*, 2007). The accuracy of the size of amplified PCR fragments was confirmed using a capillary electrophoresis system (SV1210; Hitachi High Technologies). A VNTR cluster was defined by the identification of all 15 loci in two or more isolates.

**Minimum spanning tree (MST).** A MST was generated based on the 15-MIRU-VNTR types using software (Bionumerics v. 4.6; Applied Maths) for clustering analysis. We used the reconstruction rules as follows. A categorical coefficient was selected. The priority rule was set such that the type that had the highest number of single-locus variants would be linked first. Creation of hypothetical types was not allowed. The VNTR types of Beijing family strains from neighbouring countries were retrieved from previous reports (Jiao *et al.*, 2008; Wada *et al.*, 2009). All 15-MIRU-VNTR types published in these reports were incorporated into the construction of the MST tree.

**RESULTS**

**Phylogeographical specificity of the Beijing family in South Korea**

To elucidate the population structure of *M. tuberculosis* in South Korea, we identified 64 (80.0 %) Beijing family strains from 80 isolates obtained from various geographical origins (Fig. 1) and subdivided them into three phylogenetic sublineages [ancient (RD181 +), ancient (RD181−) and modern]. A high proportion of the lineage is cordant with previous reports from South Korea and surrounding East Asian countries (Mokrousov *et al.*, 2006; Park *et al.*, 2000, 2005; van Soolingen *et al.*, 1995; Yun *et al.*, 2009). Phylogenetic subdivision of the 64 Beijing isolates revealed a higher proportion of the ancient Beijing sublineage (46 strains; 71.9 %) than the modern sublineage...
(18 strains; 28.1%). The ancient (RD181+) sublineage, having diverged from the evolutionary process towards the modern Beijing sublineage before the deletion of RD181, was observed to be predominant (29 strains; 45.3% of the Beijing strains) in the population. This study identified 13 K family strains (16.3% of all; Fig. 2): all were found to belong to the ancient (RD181+) sublineage.

**Genotypic diversity of isolates in South Korea**

We analysed all 80 isolates by 15-MIRU-VNTR to investigate the detailed genotypic diversity of the population. This genotyping method has been used as a standard discrimination tool for *M. tuberculosis* because of its high resolution among isolates from cosmopolitan origins (Supply et al., 2006). The VNTR profiles and other genotypic characteristics analysed in this study are combined and listed in Supplementary Table S2. The genotyping was able to classify our 80 isolates into 71 distinct genotypes. There were four clusters, which comprised 16 (20.0%) isolates. The clustering rate was slightly higher than that of RFLP genotyping (10 isolates; 12.5%), which was concordant with the previous reports on the Beijing family (Iwamoto et al., 2007; Yokoyama et al., 2007). Six of 13 K family strains belonging to two RFLP clusters were separated further into unique genotypes using 15-MIRU-VNTR (Fig. 3).

Recently, it was reported that cluster modelling of 15-MIRU-VNTR genotypes of the *M. tuberculosis* Beijing strains was highly concordant with their phylogenetic subdivision into some sublineages (Wada & Iwamoto, 2009). MST clustering analysis was performed for the 64 Beijing strains of our population (Fig. 4). The tree topology was observed to be highly associated with the sublineage classification.

**Comparison of VNTR profiles with those of neighbouring countries**

The MST clustering of 15-MIRU-VNTR genotypes was sufficient to visualize the phylogenetic differences among the Beijing family successfully in our study (Fig. 4), which prompted us to make an international comparison of the 15-MIRU-VNTR types with those of strains from the neighbouring countries (China and Japan) reported previously. The 202 15-MIRU-VNTR types of Beijing strains from Japan (Wada et al., 2009) and the 64 types from Beijing City, China (Jiao et al., 2008), were combined with our current data to construct a mixed clustering tree using MST (Fig. 5). Results showed that three branches including types of strains from South Korea were readily apparent, two of which included mostly South Korean types (indicated by arrowheads). Both these branches comprised strains belonging to the ancient (RD181+) sublineage, isolated in Japan and South Korea. A remaining branch included types from all three countries (indicated by an arrow). The results of the ancient/modern classification in the previous report by Wada et al. (2009) and this study showed that this branch was occupied by modern Beijing strains.

**DISCUSSION**

In general, the ancient Beijing sublineage has been considered to be an atypical Beijing genotype which has been only rarely observed (Bifani et al., 2002; Milan et al., 2004; Strauss et al., 2008). The sole exception has been the population study of Beijing family strains from Japan (Wada et al., 2009). In this study, we found that the ancient sublineage was predominant in South Korea. Contrary to the situation in Japan, it is unique that the ancient (RD181+) sublineage was the most prevalent in our 80 isolates. This sublineage has been reported to be a minor component of the Beijing family population, even in Japan (Maeda et al., 2010). Therefore, our data also revealed the particularity of the population structure of the *M. tuberculosis* Beijing family in South Korea. These observations suggest that the distribution of the Beijing family sublineages is broadly variable in different regions. It still remains unclear whether the higher prevalence of the sublineages in certain areas has been caused by the difference of fitness or by occasional clonal expansion in the past.

K family strains, derivatives of the Beijing family, have been isolated frequently in both a population-based study and outbreaks of TB in South Korea (Kim et al., 2001; Shamputa et al., 2010). In this study, they were also detected with a high clustering rate (46.2%) using RFLP genotyping (Fig. 2) despite no epidemiological link. They belonged to the ancient (RD181+) sublineage. Therefore,
the domestic prevalence of K family strains may be related to the predominance of the sublineage in the country. The 15-MIRU-VNTR genotyping was able to classify the RFLP clusters as single-locus to three-locus variants comprising the K family strains (Fig. 3). This result suggests that frequent isolation of K family strains in South Korea may be caused not by a recent expansion of a single strain but by endemic fixation in the past. The genetic diversity within the K family strains should be analysed in more detail to uncover the history of their prevalence.

Combining our data with the reported VNTR types of strains from the surrounding countries for MST clustering analysis revealed that the tree topology (the branch formation) was consistent with the phylogenetic classification of strains, irrespective of their origins (Fig. 5). The types of strains originating in China (Beijing) reported by Jiao et al. (2008) were concentrated in the branch of the modern Beijing sublineage of South Korea and Japan. Although their data did not include information on sublineages of strains, the result is consistent with those

![Fig. 3. 15-MIRU-VNTR profiles of 10 isolates belonging to four clusters by IS6110 RFLP genotyping. The order of the VNTR loci is: MIRU 04, MIRU 10, MIRU 16, MIRU 26, MIRU 31, MIRU 40, ETR A, ETR C, QUB-2163b, QUB-26, QUB-4156, Mtub04, Mtub21, Mtub30 and Mtub39. The alleles that differed from those of other isolates in the respective clusters are shaded.]

![Fig. 4. Minimum spanning tree based on the 15-MIRU-VNTR genotypes of the 64 M. tuberculosis Beijing isolates. The 56 circles depicted correspond to the different types discriminated by the 15-MIRU-VNTR genotyping. Their sizes correspond to the number of isolates with a particular genotype. They were coloured according to the phylogenetic sublineages: ancient (RD181+), ancient (RD181−) and modern. Heavy lines connecting two types denote single-locus variants; thin lines connect double-locus variants; and dotted lines (black) connect triple-locus variants. The grey dotted lines represent the most likely connection between two types differing by more than three VNTR loci.]

### Table 1

<table>
<thead>
<tr>
<th>ID</th>
<th>IS6110 RFLP</th>
<th>15-MIRU-VNTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>291</td>
<td>![Image 109x557 to 478x715]</td>
<td>233353442722343</td>
</tr>
<tr>
<td>881</td>
<td>![Image 42x51 to 298x313]</td>
<td>233353442722343</td>
</tr>
<tr>
<td>481</td>
<td>![Image 42x51 to 298x313]</td>
<td>2333753345834443</td>
</tr>
<tr>
<td>571</td>
<td>![Image 42x51 to 298x313]</td>
<td>2333753345834443</td>
</tr>
<tr>
<td>681</td>
<td>![Image 42x51 to 298x313]</td>
<td>2333742445943443</td>
</tr>
<tr>
<td>1831</td>
<td>![Image 42x51 to 298x313]</td>
<td>2333752445843543</td>
</tr>
<tr>
<td>1961</td>
<td>![Image 42x51 to 298x313]</td>
<td>2333842445943443</td>
</tr>
<tr>
<td>1461</td>
<td>![Image 42x51 to 298x313]</td>
<td>133742445943443</td>
</tr>
<tr>
<td>201</td>
<td>![Image 42x51 to 298x313]</td>
<td>2433742445953443</td>
</tr>
<tr>
<td>781</td>
<td>![Image 42x51 to 298x313]</td>
<td>2333742445943443</td>
</tr>
</tbody>
</table>
of previous reports indicating that most of the isolated *M. tuberculosis* strains in the Beijing area (76–93%) belonged to the modern Beijing sublineage (Mokrousov *et al.*, 2006; van Soolingen & Kremer, 2009). These results mean that similarity of 15-MIRU-VNTR types has been preserved in the ancient/modern sublineages of the Beijing family over the three East Asian countries. The phylogenetic validity of similar VNTR types must be verified in *M. tuberculosis* lineages because these genotypes are generally highly homoplastic. However, it is notable that similarity of VNTR types of the Beijing family has been observed based not on their geographical origins but in a phylogenetic manner. One merit of VNTR genotyping is its convenience for comparison of data from different countries (Allix-Béguec *et al.*, 2008; Mokrousov, 2008). Worldwide data accumulation of VNTR types of *M. tuberculosis* may shed light on the microevolution and genetic diversity of the species. Such global characterization of the phylogeographical distribution of the Beijing family may be useful in providing fundamental information about the ongoing worldwide expansion of the lineage.

**ACKNOWLEDGEMENTS**

This work was supported by a grant from the Ministry of Health, Labour and Welfare, Japan (Research on Emerging and Re-emerging
Infectious Diseases, Health Sciences Research Grants) and the Korean Centers for Disease Control & Prevention.

REFERENCES


