Prediction of Local Transmission of Mycobacterium tuberculosis Isolates of a Predominantly Beijing Lineage by Use of a Variable-Number Tandem-Repeat Typing Method Incorporating a Consensus Set of Hypervariable Loci

Yoshiro Murase,a Kiyohiko Izumi,b Akihiro Ohkado,b Akio Aono,a Kinuyo Chikamatsu,a Hiroyuki Yamada,a Yuriko Igarashi,a Akiko Takaki,a Satoshi Mitarai,a

aDepartment of Mycobacterium Reference and Research, The Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo, Japan
bDepartment of Epidemiology and Clinical Research, The Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association, Tokyo, Japan

ABSTRACT Strain genotyping based on the variable-number tandem repeat (VNTR) is widely applied for identifying the transmission of Mycobacterium tuberculosis. A consensus set of four hypervariable loci (1982, 3232, 3820, and 4120) has been proposed to improve the discrimination of Beijing lineage strains. Herein, we evaluated the utility of these four hypervariable loci for tracing local tuberculosis transmission in 981 cases over a 14-month period in Japan (2010 to 2011). We used six different VNTR systems, with or without the four hypervariable loci. Patient ages and weighted standard distances (a measure of the dispersion of genotype-clustered cases) were used as proxies for estimating local tuberculosis transmission. The highest levels of isolate discrimination were achieved with VNTR systems that incorporated the four hypervariable loci (i.e., the Japan Anti-Tuberculosis Association [JATA]18-VNTR, mycobacterial interspersed repetitive unit [MIRU]28-VNTR, and 24Beijing-VNTR). The clustering rates by JATA12-VNTR, MIRU15-VNTR, JATA15-VNTR, JATA18-VNTR, MIRU28-VNTR, and 24Beijing-VNTR systems were 52.2%, 51.0%, 39.0%, 24.1%, 23.1%, and 22.0%, respectively. As the discriminative power increased, the median weighted standard distances of the clusters tended to decrease (from 311 to 80 km, \( P < 0.001 \), Jonckheere-Terpstra trend test). Concurrently, the median ages of patients in the clusters tended to decrease (from 68 to 60 years, \( P < 0.001 \), Jonckheere-Terpstra trend test). These findings suggest that strain typing using the four hypervariable loci improves the prediction of active local tuberculosis transmission. The four-locus set can therefore contribute to the targeted control of tuberculosis in settings with high prevalence of Beijing lineage strains.

KEYWORDS Beijing lineage, local transmission, molecular epidemiology, Mycobacterium tuberculosis, variable-number tandem-repeat typing

The development of molecular genotyping methods for Mycobacterium tuberculosis has significantly improved our understanding of tuberculosis (TB) epidemiology over the last 2 decades (1). Genotyping methods, IS6110 restriction fragment length polymorphism, variable-number of tandem-repeat (VNTR) typing, and spoligotyping are used to elucidate TB transmission dynamics in human populations. Due to the methodological difficulty and inconvenience of the former method, a simple and faster method, VNTR typing, is extensively used as a current standard for the genotyping of M. tuberculosis isolates (1, 2). However, recent studies demonstrated suboptimal discrimination of the international standardized VNTR methods, i.e., 15- and 24-locus-optimized mycobacterial interspersed repetitive unit (MIRU)-VNTR, for genetically sim-
ilar M. tuberculosis, such as Beijing lineage strains (3–7). Clusters falsely defined by this method include TB cases that lack epidemiological links and hence hamper the evaluation of TB transmission in a community (6, 8–10). A more reliable genotyping method is thus required for contact investigation to control TB transmission.

To circumvent the lack of discriminatory power of the VNTR methods, whole-genome sequencing (WGS) has been introduced for the molecular epidemiology of M. tuberculosis (11). Several studies demonstrated the high-resolution power of WGS to trace an ongoing transmission, even among TB patients with genetically similar Beijing lineage strains (9, 12). However, because of its high running cost and need for expensive equipment, the routine application of WGS remains impractical. Furthermore, standardization and quality assurance schemes are needed before the WGS method is implemented worldwide (13).

Another approach constitutes the inclusion of hypervariable loci into VNTR typing. Recently, a four-hypervariable-locus set (VNTR 1982, VNTR 3232, VNTR 3820, and VNTR 4120) has been proposed as a consensus set for subtyping Beijing clonal complexes and clusters, as a follow-up to the standard VNTR typing (5). Although high discrimination of the hypervariable loci is expected, its utility for the detection of local TB transmission in a setting with high prevalence of the Beijing lineage strains remains unknown.

In this study, we evaluated the utility of the hypervariable loci in the detection of local TB transmission in Japan where the Beijing strains are highly prevalent. To compensate for the lack of epidemiological data that would validate TB transmission between patients, we evaluated the spatial aggregation of isolates that had identical VNTR profiles (clustered isolates), using geographic and genotypic data for 981 cases recruited in the country during 14 months (years 2010 to 2011).

MATERIALS AND METHODS

Study samples. From October 2010 to November 2011, 987 M. tuberculosis isolates were acquired from 987 patients with TB (pulmonary and extrapulmonary), regardless of the smear results, from 37 of the 47 prefectures in Japan (see Fig. S1 in the supplemental material). All cases with M. tuberculosis-positive cultures were eligible. The clinical isolates were collected as a part of routine diagnostic examination. Six isolates were excluded because of an insufficient quantity of extracted DNA.

Patient characteristics and bacterial lineage. The study population consisted of 981 patients. Among these, 353 and 198 cases were excluded from the analysis of age and sex, respectively, because of missing records. For the remaining analyses, all 981 cases were used. The characteristics of the study population were as follows: 484 (61.8%) were male and 299 (38.2%) were female; the median age was 72 years (range, 2 to 104 years). The study population was enrolled from 37 of 47 prefectures (Fig. S1). Among the clinical isolates, the proportion of the Beijing and non-Beijing lineage isolates were 70.6% (693/981) and 29.4% (288/981), respectively.

Ethics statement. All samples used in this study were anonymized, and personally identifiable data were removed beforehand. Because of the anonymous nature of the data, the requirement for informed consent was waived. The study protocol was approved by the ethics committee of the Research Institute of Tuberculosis, Japan Anti-Tuberculosis Association (approval no. RIT/IRB 28-9).

Genotyping methods. The usefulness of the hypervariable loci for strain discrimination was evaluated by comparing the discriminatory power of six VNTR systems that either incorporated the hypervariable loci or did not: optimized-MIRU15-VNTR, optimized-MIRU28-VNTR, JATA12-VNTR, JATA15-VNTR, JATA18-VNTR, and 24Beijing-VNTR (Table S1) (14–16). The latter four systems are widely used in Japan for efficient discrimination of Beijing lineage strains. From an international perspective, we used optimized-MIRU15-VNTR to represent the minimum set for routine epidemiological discrimination of M. tuberculosis (17). Optimized-MIRU28-VNTR was used as a new gold standard set for subtyping Beijing clonal complexes and clusters (5).

The copy numbers of tandem repeats of the six VNTR systems were calculated on the basis of the molecular sizes of PCR products determined using a 3500 genetic analyzer with a GeneMapper program (Applied Biosystems, Foster, MA) or 1.8% agarose gel electrophoresis (16, 18). M. tuberculosis H37Rv was used as a new gold standard set for subtyping Beijing clonal complexes and clusters (5).

Spatial analysis. Because of the anonymous nature of this study, only patients’ age and sex information were obtained. The spatial analysis was performed on the basis of the addresses of the capital cities of the prefectures of the patients’ hospitals. The degree of spatial aggregation of TB cases was assessed for genotype-cluster groups that consisted of four or more cases. The spatial aggregation was determined by weighted standard distance (in km); this parameter is used to express the dispersion of cases around the mean center (20). The weighted standard distance for each genotype-cluster group was calculated using the following formula:

\[ D = \frac{1}{k} \sum_{i=1}^{k} d_i \]
weighted standard distance $= \sqrt{\sum_{i=1}^{n} w_i (x_i - X_w)^2 + \sum_{i=1}^{n} w_i (y_i - Y_w)^2} / \sum_{i=1}^{n} w_i$

where $x_i$ and $y_i$ are the coordinates for the location of capital city $i$, $w_i$ is the weight using the number of cases at $i$, $X_w$ and $Y_w$ represent the weighted mean center for the capital cities with cases, and $n$ is the total number of cities with cases. A small value indicates a high aggregation of TB cases; a large value indicates a dispersed distribution of TB cases. The spatial aggregation of each cluster was represented by drawing a circle with a radius equal to the weighted standard distance, expected to cover ca. 68% of cases within clusters. The spatial analysis was conducted with ArcGIS v. 10.2 (ESRI Inc., Redlands, CA).

**Definitions.** A genotype cluster was defined as a set of isolates with identical VNTR profiles. Noncluster referred to isolates with a unique VNTR profile that was not shared by any other isolate. The clustering rate was calculated as a proportion of clustered isolates among all isolates. Clusters that contained isolates from more than one prefecture were defined as regional clusters; those with isolates from the same prefecture were defined as local clusters. Clusters with a weighted standard distance of up to 80 km were classified as geographically small clusters; those with a weighted standard distance of over 80 km were classified as geographically large clusters.

**Statistical analysis.** Either chi-square tests or Fisher’s exact tests were performed to compare categorical variables, and Mann-Whitney U tests were performed for continuous variables. When a trend was observed, a Jonckheere-Terpstra trend test was used. A $P$ value of $<0.05$ was considered to indicate a significant difference. All statistical analyses were performed with EZR (version 1.35), a graphical interface for R (v. 3.3.2) (21).

**RESULTS**

**Discriminatory power of the six VNTR systems.** The following systems were evaluated: Japan Anti-Tuberculosis Association (JATA)12-VNTR, MIRU15-VNTR, JATA15-VNTR, JATA18-VNTR, MIRU28-VNTR, and 24Beijing-VNTR (see Table S1 in the supplemental material) (14–16). As anticipated, the highest levels of isolate discrimination, expressed by the clustering rate (i.e., the higher the discriminatory power, the lower the clustering rate), were achieved with 24Beijing-VNTR, MIRU28-VNTR, and JATA18-VNTR, all of which incorporate the four hypervariable loci (clustering rates, 22.0%, 23.1%, and 24.1%, respectively) (Fig. 1). An intermediate level of isolate discrimination was obtained with JATA15-VNTR, which includes only one of the four hypervariable loci (clustering rate, 39.0%). Optimized-MIRU15-VNTR and JATA12-VNTR were characterized by the lowest levels of isolate discrimination (clustering rates, 51.0% and 52.2%, respectively). The clear improvement of isolate discrimination using the four hypervariable loci was consistent among both Beijing and non-Beijing isolates. Beijing isolates formed clusters more frequently than non-Beijing isolates, regardless of the VNTR system used (e.g., clustering rate by 24Beijing-VNTR in Beijing versus non-Beijing isolates, 24.1% versus 17.0%, respectively; $P = 0.014$, Fisher’s exact test).
Spatial aggregation of clustered isolates according to analyses with six VNTR systems. Since we anticipated the clonal tuberculosis transmission would occur in a geographically restricted area rather than over a large region (22, 23), we next evaluated the spatial aggregation of cases in genetic clusters defined by the six VNTR systems. Weighted standard distances of clusters were calculated; this parameter indicates the level of spatial aggregation of TB cases within each cluster (20). A decreasing trend for cluster number with increasing discriminatory power of VNTR was observed (121, 109, 97, 71, 66, and 66 clusters for JATA12-VNTR, optimized-MIRU15-VNTR, JATA15-VNTR, JATA18-VNTR, optimized-MIRU28-VNTR, and 24Beijing-VNTR, respectively). At the same time, the median weighted standard distances of clusters displayed a significantly decreasing trend with increasing discriminatory power of VNTR systems (for the systems mentioned above: 311, 258, 208, 94, 99, and 80 km, respectively; \( P < 0.001 \), Jonckheere-Terpstra trend test) (Fig. 2 and 3). These observations indicated that isolates with identical low-discriminatory VNTR profiles are highly prevalent throughout the country. In contrast, isolates with high discriminatory VNTR profiles tended to distribute locally. This suggested that the VNTR systems incorporating the four hypervariable loci potentially narrow down a suspected area where a clonal TB transmission occurred.

To test this possibility further, we next compared the median ages of clustered cases across the VNTR systems. Young age is one of the key risk factors of recent TB transmission; younger individuals are at a lower risk of disease reactivation and their social behaviors are more likely to be transmissible (24, 25). We found that the median ages of clustered cases tended to become lower with increasing discriminatory power...
of the VNTR systems (68, 68, 66, 61, 63, and 60 years for JATA12-VNTR, optimized-MIRU15-VNTR, JATA15-VNTR, JATA18-VNTR, optimized-MIRU28-VNTR, and 24 Beijing-VNTR, respectively; \( P < 0.001 \), Jonckheere-Terpstra trend test) (Fig. 4). Hence, the age trend also supported the possibility that the VNTR systems incorporating the four hypervariable loci potentially improve the prediction of an active local TB transmission. Consequently, the most discriminative VNTR system, 24Beijing-VNTR, was used for further analyses.

**Genotype clusters.** Of the 981 isolates, 216 isolates (22.0%) formed 66 genotype clusters according to 24Beijing-VNTR typing. More than half of the clusters (54.5% [36/66]) contained two isolates (Fig. 5). Most clusters (80.3% [53/66]) contained at least one isolate from a different prefecture (regional clusters). The number of local clusters was limited (19.7% [13/66]); most of them (92.3% [12/13]) contained two isolates. We detected two cases of a possible clonal transmission of drug-resistant isolates: seven isoniazid-monoresistant isolates and 14 streptomycin-monoresistant isolates, which had identical 24Beijing-VNTR profiles within a cluster.

**Characteristics of clustered cases.** We analyzed the association between genetic clustering by 24Beijing-VNTR with patients’ characteristics and bacteriological factors (Table 1). Patients included in clustered cases were younger than ones in nonclustering cases (60 versus 75 years, respectively; \( P < 0.001 \)). Beijing isolates formed clusters more frequently than non-Beijing isolates (odds ratio [OR], 1.55; 95% confidence interval [CI], 1.08 to 2.25). Streptomycin resistance was significantly associated with clustering (OR, 1.91; 95% CI, 1.10 to 3.25). Other factors, such as sex and resistance to other drugs, were not significantly associated with clustering.

To further evaluate factors involved in cluster formation, we compared the characteristics of geographically small (≤80 km) and large (>80 km) clusters (Table 2). Patients representing cases involved in geographically small clusters were younger than those
in large clusters (57 versus 65 years, respectively; \( P = 0.019 \)). Isoniazid resistance was associated with geographically small clusters (OR, \( \infty \); 95% CI, 2.61 to \( \infty \)), while streptomycin resistance was associated with large clusters (OR, 9.69; 95% CI, 2.06 to 92.35). Other factors were not associated with geographically small clusters.

**DISCUSSION**

A high-resolution genotyping tool is needed to efficiently control TB infections, especially in a setting where the highly homogeneous Beijing strains are prevalent, e.g.,
in Japan. In this study, using 981 M. tuberculosis isolates collected throughout Japan over a 14-month period, we found a higher spatial aggregation of TB cases within clusters based on 24Beijing-VNTR, optimized-MIRU28-VNTR, or JATA18-VNTR than within clusters generated with other VNTR methods that did not include the four hypervariable loci (Fig. 2 and 3). Furthermore, cases clustered based on 24 Beijing-VNTR, optimized-MIRU28-VNTR, or JATA18-VNTR were associated with lower patient ages (Fig. 4). These results suggest that a high-resolution VNTR system incorporating the four hypervariable loci potentially narrows down a suspected area where an active TB transmission occurred.

This is the first study to demonstrate the utility of the four hypervariable loci for the prediction of an active local TB transmission over a broad area with high prevalence of Beijing lineage strains. The four hypervariable loci have recently been proposed to constitute a consensus set to improve the discrimination of Beijing lineage strains (5). Although their highly polymorphic nature has been shown in China, Albania, and Vietnam, geographical relations between the clustered cases, essential to assess local transmission, remain unknown (3, 13, 14). Our study has added novel evidence that the four-locus set improves infection tracking. Therefore, we recommend the usage of the four hypervariable loci to facilitate the identification of geographically linked isolate groups in settings with high Beijing lineage prevalence.

Spatial analysis in conjunction with genotyping may be used to identify active TB transmission that occurs over a small area and to distinguish that from one that would be expected by chance (22, 23). However, such studies have never been conducted in Japan. Hence, the present study was conducted using 981 isolates collected throughout Japan over a year (2010 to 2011). The choice of this setting is justified by three arguments. First, Japan is surrounded by sea and has no physical borders with other countries. Hence, human migration from other countries, which would affect the interpretation of transmission studies, is limited. Second, the proportion of foreign nationals among all TB cases notified is low (7.6% in 2016) (26). Third, the proportion

### TABLE 1 Factors associated with case clustering by 24Beijing-VNTR

<table>
<thead>
<tr>
<th>Factor</th>
<th>24Beijing-VNTR clustered</th>
<th>OR (95% CI) or P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age (yrs [range])</td>
<td>Yes (n = 216)</td>
<td>No (n = 765)</td>
</tr>
<tr>
<td>Sex (n [%])</td>
<td>Female</td>
<td>51 (30)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>119 (70)</td>
</tr>
<tr>
<td>INH resistance (n [%])</td>
<td>R</td>
<td>13 (6.0)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>203 (94)</td>
</tr>
<tr>
<td>RIF resistance (n [%])</td>
<td>R</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>216 (100)</td>
</tr>
<tr>
<td>EB resistance (n [%])</td>
<td>R</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>215 (100)</td>
</tr>
<tr>
<td>STR resistance (n [%])</td>
<td>R</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>191 (88)</td>
</tr>
<tr>
<td>Lineage (n [%])</td>
<td>non-Bj</td>
<td>49 (23)</td>
</tr>
<tr>
<td></td>
<td>Bj</td>
<td>167 (77)</td>
</tr>
</tbody>
</table>

®The total number of cases was 981, except for patient age (n = 628) and sex (n = 783) analyses. INH, isoniazid; RIF, rifampin; EB, ethambutol; STR, streptomycin; R, resistant; S, sensitive; Bj, Beijing lineage isolates; non-Bj, non-Beijing lineage isolates.

®P value calculated by Mann-Whitney U test.
of cases with the Beijing lineage in Japan is high (71% in this study). Thus, the chosen setting was ideal in that it enabled us to focus on the transmission dynamics of the Beijing strains over a broad area.

In this study, we identified clusters of various geographical sizes by using 24Beijing-VNTR (the range of weighted standard distance, 20 to 400 km) (Fig. 3). Clustered strains distributed over a large area are probably more likely to be epidemiologically unrelated than those distributed locally. The former strains may have arisen from common endemic strains or reactivation diseases (7, 9, 10, 12, 22). In those cases, the strains have microevolved over time via multiple transmission events or latent infection periods, but their VNTR profiles may not have changed.

In contrast, M. tuberculosis isolates from geographically restricted clusters are likely genetically closely related, that is, they arise as a result of recent transmission events (9, 10). In the present study, the patients grouped in geographically small clusters were younger than those in dispersed clusters, supporting this notion (57 versus 65 years, respectively; \( P = 0.018 \)) (Table 2). When considering a targeted TB control program, geographically linked clusters should be prioritized for contact investigation surveys. Indeed, as a consequence of a 3-year evaluation project, the routine investigation of regional clusters (i.e., geographically large clusters) will no longer be conducted in the United Kingdom (27).

A significant association between streptomycin resistance and geographically large clustering was found (Table 1). This association was due to a single large-size cluster consisting of 14 mono-streptomycin-resistant isolates with a weighted standard distance of 248 km and an average patient age of 45.8 years (data not shown). Interestingly, the VNTR profile of this cluster perfectly matched that of M strains, previously reported as nation-wide endemic strains with high genetic clonality (12). Therefore, we concluded that our study included M strains that are spreading over a wide area of Japan, forming the cluster identified in the study.

Another example of a drug resistance-related cluster was an association with

### TABLE 2 Factors associated with geographically small (≤80 km) clustering

<table>
<thead>
<tr>
<th>Factor</th>
<th>Data at weighted standard distance of:</th>
<th>OR (95% CI) or ( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>≤80 km ((n = 52))</td>
<td>&gt;80 km ((n = 53))</td>
</tr>
<tr>
<td>Median age (yrs [range])</td>
<td>57 [19–91]</td>
<td>65 [16–92]</td>
</tr>
<tr>
<td>Sex ((n\ [%]))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>5 (13)</td>
<td>12 (29)</td>
</tr>
<tr>
<td>Male</td>
<td>33 (87)</td>
<td>30 (71)</td>
</tr>
<tr>
<td>INH resistance ((n\ [%]))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>10 (19)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>S</td>
<td>42 (81)</td>
<td>53 (100)</td>
</tr>
<tr>
<td>RIF resistance ((n\ [%]))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>S</td>
<td>52 (100)</td>
<td>53 (100)</td>
</tr>
<tr>
<td>EB resistance ((n\ [%]))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0 (0)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>S</td>
<td>52 (100)</td>
<td>52 (98)</td>
</tr>
<tr>
<td>STR resistance ((n\ [%]))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>2 (3.8)</td>
<td>15 (28)</td>
</tr>
<tr>
<td>S</td>
<td>50 (96)</td>
<td>38 (72)</td>
</tr>
<tr>
<td>Lineage ((n\ [%]))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-Bj</td>
<td>6 (12)</td>
<td>4 (7.5)</td>
</tr>
<tr>
<td>Bj</td>
<td>46 (88)</td>
<td>49 (92)</td>
</tr>
</tbody>
</table>

^aThe total number of cases was 105, except for patient age (\(n = 67\)) and sex (\(n = 80\)) analyses. INH, isoniazid; RIF, rifampin; EB, ethambutol; STR, streptomycin; R, resistant; S, sensitive; Bj, Beijing lineage isolates; non-Bj, non-Beijing lineage isolates.

^b\(P\) value calculated by Mann-Whitney U test.
isoniazid resistance. This cluster consisted of seven mono-isoniazid-resistant isolates with a geographic link (weighted standard distance, 38 km). These patients were relatively young compared to general TB patients in Japan (average age, 42.3 versus 66.0 years, respectively) (data not shown). Although local transmissions could explain this cluster, official notifications discussing such an outbreak had not been reported to the national government from these local governments until 2016 (26). Therefore, we concluded that this isoniazid-resistant cluster reflects potentially unidentified local transmissions that have been overlooked by conventional public health practice.

We found that VNTR systems with higher discriminatory power decreased the proportions of clustered cases (from 52.2% to 22.0%) (Fig. 1), the numbers of clusters (from 121 to 66 clusters), and the geographical dispersion of clusters (median weighted standard distance, from 311 to 80 km) (Fig. 3). Therefore, we would expect that the higher discriminatory power of the VNTR system, the more efficient the contact investigation survey can be based on the identification of clusters that include actual transmission events.

Although the inclusion of hypervariable loci improves the discriminatory power in the case of Beijing lineage isolates, a lower reproducibility of such analyses was reported (17). Compared with the other standard loci, the frequencies of large alleles in the four hypervariable loci are higher, rendering them more difficult to analyze. To resolve this issue, accurate sizing methods involving capillary DNA analyzers were reported (5, 18). Consequently, the four-locus set was proposed as an international consensus for subtyping Beijing clonal isolates (5). In fact, we have confirmed the high reproducibility of the four hypervariable loci among >30 participant laboratories through external quality assurance programs in the years 2015 and 2016 in Japan, although some laboratories required improvements in typing performance for the four hypervariable loci (Y. Murase and S. Mitarai, unpublished data). A continued quality control effort is particularly required for the analysis of hypervariable loci.

In Japan, the genotyping of TB isolates is independently conducted by public health laboratories in 47 prefectures, and the data are not routinely shared. In the present study, 80% of the clusters were spread across the prefecture borders (Fig. 5); consequently, a segment of the transmission chain would have been missed by a single facility-based isolate-genotyping program. Hence, a database that crosses internal national administrative borders is urgently required, and the introduction of such a database is under review by the government of Japan. Furthermore, an international database is needed to understand the transmission dynamics of TB across different countries.

There are several limitations to our study. First, we were unable to obtain epidemiological linkage data that would validate patient-to-patient transmission among genotypic clusters. As the purpose of the study was not to directly intervene in TB control practice by public health centers, we were not allowed to use personal identifying information. The lack of an epidemiological link is a problem universal to many TB studies, because conventional contact investigation surveys often fail to identify epidemiological links between TB cases that occur as a result of casual contacts or between TB cases in which the sources of the infections are missing (9, 28). To partially circumvent this issue, we conducted a spatial analysis of the clustered cases on a large population scale. Second, the sampled portion (the proportion of the enrolled cases among new TB cases with bacterial-positive TB during the study period) was low (8.0%) and the study period was short (14 months); both of these may potentially result in missed TB cases that were involved in the transmission chain. Third, we performed a spatial analysis based on the addresses of the capital cities of the prefectures of patients’ hospitals. The spatial aggregation of TB cases may therefore have been overestimated, especially in the instances when the patients visited the same tertiary TB hospitals. Lastly, the application of WGS analysis to estimate TB transmission in more detail was beyond the scope of this study, because sequence data were not available. Further research is therefore necessary to address these limitations.

Despite the limitations, this study nonetheless demonstrated the usefulness of using...
the hypervariable loci for assessing active TB transmission in a setting where the Beijing strains are the dominant lineage. This conventional molecular epidemiological investigation has contributed to the understanding of TB transmission dynamics; however, epidemiological links were not established for some cases, even ones identified as belonging to the same genotype clusters. Nevertheless, we showed that VNTR genotyping that incorporates the four hypervariable loci potentially improves the prediction of a local and actively occurring TB transmission. This method can be applied in combination with contact investigation for improved and targeted TB control in a setting where the Beijing lineage strains are highly prevalent.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JCM.00126-16.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.

SUPPLEMENTAL FILE 2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Tomoko Kato for DNA preparation and VNTR analysis. This work was supported by the Research Program on Emerging and Re-emerging Infectious Diseases from the Japan Agency for Medical Research and Development (AMED) grant number 16fk0108301h0003). All authors have declared that no competing interests exist.

REFERENCES


