ORIGINAL ARTICLE BACTERIOLOGY

Genetic diversity of *Mycobacterium tuberculosis* isolates from foreign-born and Japan-born residents in Tokyo

M. Kato-Miyazawa[1], T. Miyoshi-Akiyama[1,2], Y. Kanno[3], J. Takasaki[3], T. Kirikae[1] and N. Kobayashi[3,4]

1) *Department of Infectious Diseases,* 2) *Pathogenic Microbe Laboratory, Research Institute, National Center for Global Health and Medicine,* 3) *Department of Respiratory Medicine, National Center for Global Health and Medicine and* 4) *National Hospital Organization Tokyo National Hospital, Tokyo, Japan*

Abstract

Sequences of the full genomes of 259 clinical isolates of *Mycobacterium tuberculosis*, obtained from foreign-born and Japan-born patients in Tokyo, Japan, were determined, and a phylogenetic tree constructed by concatenated single-nucleotide polymorphism (SNP) sequences. The 259 isolates were clustered into four clades: Lineage 2 (East Asian or "Beijing" genotype; *n* = 182, 70.3%), Lineage 4 (Euro-American, *n* = 46, 17.8%), Lineage 1 (Indo-Oceanic, *n* = 23, 8.9%), and Lineage 3 (East African-Indian, *n* = 8, 3.1%). Of the 259, 36 (13.9%) were resistant to at least one drug. There was no multi-drug-resistant isolate. Drug resistance was greater for the strains in Lineage 2 than the non-Lineage 2. The proportion of Lineage 2 isolates was significantly smaller in foreign-born (*n* = 43/91, 47.3%) than in Japan-born (*n* = 139/168, 82.7%) patients, whereas the proportion of Lineage 1 isolates was significantly larger in foreign-born (*n* = 19/91, 20.9%) than in Japan-born (*n* = 4/168, 2.4%) patients. We also found eight SNPs specific to the typical Beijing sub-genotype in Lineage 2, including 4 non-synonymous SNPs. Of the 259 isolates, 244 had strain-specific SNP(s) and small (1–30-bp) insertions and deletions (indels). The numbers of strain-specific SNPs and indels per isolate were significantly larger from foreign-born (median 89, range 0–520) than from Japan-born (median 23, range 0–415) (p 3.66E-15) patients. These results suggested that *M. tuberculosis* isolates from foreign-born patients had more genetic diversity than those from Japan-born patients.

Keywords: Epidemiology, foreign-born, lineage, *Mycobacterium tuberculosis*, whole genome sequencing
Original Submission: 2 July 2014; **Revised Submission:** 28 September 2014; **Accepted:** 29 September 2014
Editor: Prof. D. Raoult
Article published online: 13 October 2014

Corresponding author: T. Kirikae, Department of Infectious Diseases, Research Institute, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku, Tokyo 162-8655, Japan
E-mail: tkirikae@ri.ncgm.go.jp

Introduction

The World Health Organization has estimated that there are 8.6 million newly diagnosed patients and 1.3 million tuberculosis (TB)-associated deaths worldwide each year [1]. The highest TB burden region is in Asia, with 58% of the estimated numbers of patients in 2013 [1]. All Asian countries except for Japan are categorized as high or relatively high burden countries, defined as ≥100 patients per 100 000 population [1]. The national surveillance of Japan (http://www.mhlw.go.jp/english/wp/wp-hw5/dl/23010226e.pdf) showed that the TB prevalence rate in Tokyo was 25 patients per 100 000 population and more than twice that of rural areas. Foreign-born individuals account for over 10% of the total population of Shinjuku, one of the central parts of Tokyo, with most coming from high/relatively high burden countries, including, in descending order, China (36.9% of the foreign-born residents), Korea (30.8%), Vietnam (6.3%), Nepal (5.7%), and Burma (Myanmar) (3.4%). To date, however, there has been no epidemiological study of foreign-born TB patients in Tokyo.

The purpose of this study is to clarify whether *Mycobacterium tuberculosis* isolates from foreign-born TB patients in Japan have the potential risk of spreading to other residents in Japan. Whole

http://dx.doi.org/10.1016/j.cmi.2014.09.013

genome sequencing is a powerful epidemiological tool in TB epidemiology [2,3]. This study, therefore, was designed to determine the entire genome sequences of 259 clinical isolates of *M. tuberculosis* obtained from foreign-born and Japan-born residents in Tokyo. Comparisons showed bacterial genome differences in TB strains obtained from foreign- and Japan-born hosts.

Materials and methods

Patients and bacterial strains

Of TB patients admitted to a hospital (National Center for Global Health and Medicine; NCGM) in Tokyo from January 2001 through June 2012, 91 were foreign born and culture confirmed. Bacterial strains obtained from the patients were kept in storage. Patients' gender, age, nationality, and period of stay in Japan were documented. They comprised 56 (61.5%) males and 35 (38.5%) females, mean age 32.1 years; of these, 20 (22.0%), 46 (50.5%), 21 (23.1%), and 4 (4.4%) were aged 15–24, 25–34, 35–54, and >54 years, respectively. Two age-/gender-matched Japan-born patients were selected at random for each foreign-born patient. Another 168 isolates were obtained from Japan-born patients, comprising 113 (67.3%) males and 55 (32.7%) females, mean age 32.9 years; of these, 44 (26.2%), 66 (39.3%), 49 (29.2%), and 9 (5.4%) were aged 15–24, 25–34, 35–54, and >54 years, respectively. Foreign-born patients were from 19 different countries/regions, and their median time of stay in Japan was 36 months (interquartile range 12–102 months) (Table 1). A total of 259 isolates of *M. tuberculosis* were obtained, one from each patient. These isolates were tested for drug susceptibility using Vite Spectrum SR (Kyokuto Pharmaceutical Industrial, Tokyo, Japan), which is a conventional proportion method on egg-based medium.

TABLE 1. Country/region of origin of foreign-born patients and length of stay in Japan

Areas and countries/regions, and time of stay in Japan	Foreign patients $n = 91$	
	n	%
Country/region		
Asia		
China	22	24.2
Korea	18	19.8
Philippines	10	11.0
Burma (Myanmar)	8	8.8
Nepal	7	7.7
Vietnam	4	4.4
Indonesia	3	3.3
India	3	3.3
Mongolia	3	3.3
Taiwan	2	2.2
Iran	1	1.1
Europe[a]	4	4.4
Africa[b]	3	3.3
America[c]	3	5.5
Time of stay in Japan (month)		
<1	2	2.2
1–6	8	8.8
7–36	34	37.4
<37	27	29.7
Unknown	20	22.0

[a]European countries were France ($n = 2$),UK ($n = 1$), and Russia ($n = 1$).
[b]African countries were Ethiopia ($n = 1$), Somalia ($n = 1$), and Nigeria ($n = 1$).
[c]American countries were Peru ($n = 2$) and the USA ($n = 1$).

Whole genome sequencing of *M. tuberculosis* isolates, mapping of sequencing reads, and single-nucleotide polymorphism calling

Genomic DNA of all isolates from the 259 TB patients was extracted by standard procedures [4]. Paired-end multiplexed Illumina sequencing using MiSeq (Illumina Inc., San Diego, CA, USA) and mapping of sequence reads was performed as described [5]. Briefly, generated paired-end reads from each genome were assembled with CLC Genomics Workbench ver. 6.5 (CLC Bio, Aarhus, Denmark). The output Illumina reads generated in this study were submitted under accession number DRA001219. Sequencing reads from each isolate were mapped to the *M. tuberculosis* reference genome H37Rv (AL123456). The average length of individual mapped sequence reads was 229.9 (range 129.3–246.2). Median genome sequence coverage was 57.9 of the reference genome (range 21.5–1472.2, Supplementary Table 1). Single-nucleotide polymorphisms (SNPs), small (1–30-bp) insertions and deletions (indels), structural variants, and large (>30-bp) indels were called in the CLC genomics workbench. Whole genome sequences of 18 strains were obtained from the National Center for Biotechnology Information database: Beijing_NITR203, CCDC5180, CCDC5079, CDC1551, CTRI-2, EAI5, EAI5-NITR206, Erdman, F11, KZN605, KZN1435, KZN2407, H37Ra, H37Rv, RGRB327, RGTB423, UT205, and 7199-99. SNPs of the 18 strains against H37Rv were called using by MUMmer, a software for large-scale DNA and protein sequences [6]. The identified SNPs (coverage >11 and frequency >95%) were concatenated to obtain concatenated SNP sequences to be aligned with MAFFT [7].

Phylogenetic analyses

A dataset was comprised of the 259 clinical isolates and the 18 published strains. Maximum-likelihood and neighbor-joining methods were used for phylogenetic analyses. A maximum-likelihood tree was generated from aligning SNPs with PhyML3.0 [8] using general time reversible (GTR) model. The best maximum-likelihood tree was selected by using TREE-PUZZLE (http://www.tree-puzzle.de) and CONSEL (http://www.is.titech.ac.jp/~shimo/prog/consel). A neighbor-joining tree was constructed using the GTR model with a bootstrap support analysis (100 iterations) in MEGA 5.07 [9].

In silico genotyping

M. tuberculosis lineages were defined based on large sequence polymorphism and SNP markers [10,11]: RD239 for Lineage 1

(Indo-Oceanic), RD105 for Lineage 2 (East Asian or Beijing), RD750 for Lineage 3 (East African-Indian), 7-bp deletion in *pks15/1* and/or t1388g in *katG* for Lineage 4 (Euro-American), and RD711 for Lineage 5 (West African I) and RD702 for Lineage 6 (West African II). An insertion in the *dnaA-dnaN* region [12] was also used to identify Lineage 2. Isolates belonging to the typical ("modern") Beijing sub-genotype in Lineage 2 were those with an insertion categorized as an IS*6110* insertion in the noise transfer function (NTF) chromosomal region (H37Rv position 3493689-3493990) [13,14], and with a Gly58Arg substitution in *mutT2* [7,8,15–18]. IS*6110* was detected as a structural variant (an insertion) in the *dnaA-dnaN* and/or NTF regions since IS*6110* sequence could not be determined by short sequence reads of MiSeq in this study. The remaining isolates of Lineage 2 were defined as being in the atypical ("ancestral") Beijing sub-genotype. Substitutions in other mutator genes, including *ogt* and *mutT4* [7,8,15–18], were analyzed.

Spoligotyping was performed *in silico*, with a spoligo-international-type (SIT) number determined by the international spoligotyping database (SITVIT and SpolDB4.0).

Statistics analysis

Data were summarized as mean, median, range, and interquartile range, as appropriate, and compared using Fisher's exact test and the nonparametric Mann-Whitney *U*-test. All tests were two-sided, and $p < 0.05$ was considered statistically significant.

Ethical considerations

This study protocol was reviewed and approved by the ethics committee of NCGM (approval number NCGM-G-001467-00). Patient information was made anonymous and tightly controlled according to the Japanese government ethics guidelines.

Results

In silico genotyping of isolates

A maximum-likelihood tree was constructed from the 259 *M. tuberculosis* isolates from foreign- and Japan-born patients, and the 18 published strains (Fig. 1). The tree was congruent with the tree selected by CONSEL (Supplementary Fig. 1 and Supplementary Table 2) or constructed with the neighbour-joining method (Supplementary Fig. 2). Phylogenetic analysis revealed four clades, in good agreement with four of the lineages categorized by genotyping based on large sequence polymorphism/SNP markers and an insertion in the *dnaA-dnaN* locus, as indicated by previous studies [2,3]. Of the 259 isolates, 182 (70.3%) were of Lineage 2 (East Asian or Beijing), 46 (17.8%) of Lineage 4 (Euro-American), 23 (8.9%) of Lineage 1 (Indo-Oceanic), and 8 (3.1%) of Lineage 3 (East African-Indian).

Of the 182 isolates belonging to Lineage 2 (East Asian or Beijing), 81 (44.5%) belonged to the typical Beijing sub-genotype in Lineage 2, i.e. they had structural variants of the NTF region and mutations of Gly58Arg in *mutT2*. These 81 isolates were sub-clustered in the whole genome phylogenic tree (Fig. 1). The remaining 101 (55.5%) isolates of Lineage 2 belonged to the atypical Beijing sub-genotype in Lineage 2, i.e. they had neither an insertion in the NTF region nor a Gly58Arg substitution in *mutT2*. These 101 isolates were also sub-clustered in the whole genome phylogenic tree (Fig. 1). Substitutions in other mutator genes, *ogt* and *mutT4*, are shown in Supplementary Table 1.

Two of the isolates did not belong to any cluster in the phylogenic tree (Fig. 1). Both had characteristic markers for Lineage 2 (East Asian or Beijing), including an insertion in the *dnaA-dnaN* locus and a deletion of Rv0069-0075 including the RD105 locus, but both had different spoligo patterns than the Beijing genotype (Supplementary Table 1). Neither of these isolates had an insertion in the NTF region or mutations in *mutT2*, *ogt*, and *mutT4*.

Spoligotypes could be determined for 248 of the 259 isolates, but not for the other 11. Eighteen spoligotypes obtained from 152 isolates had known SIT numbers, whereas 84 spoligotypes from 96 isolates had unknown SIT numbers (Supplementary Table 3).

Lineage comparisons between foreign-born and Japan-born patient isolates

The distribution of isolates among *M. tuberculosis* lineages differed significantly between foreign- and Japan-born patients (Table 2). The largest number of isolates from foreign-born patients belonged to Lineage 2 (East Asian or Beijing), Lineage 4 (Euro-American), and Lineage 1 (Indo-Oceanic), whereas the majority of isolates from Japan-born patients belonged to Lineage 2. In patients with Lineage 2, most of the isolates from foreign-born patients were of the typical Beijing sub-genotype (67.4%), whereas most of the isolates from Japan-born patients were of the atypical sub-genotype (62.8%). Of the isolates from foreign-born patients, the majority from patients born in East Asian countries or regions including China, Korea, Mongolia, and Taiwan were of typical Beijing sub-genotype in Lineage 2, whereas most isolates from patients born in non-East Asian countries were of Lineage 1 or 4 (Supplementary Table 4). As shown in Fig. 1(b), there were clusters (sub-clades) comprising isolates specific to either foreign- or Japan-born patients (indicated as F and J in Fig. 1(b)). Whereas there were clusters comprising isolates shared by both groups, which were especially represented by typical Beijing sub-genotype in Lineage 2, and Lineage 4. No significant change in

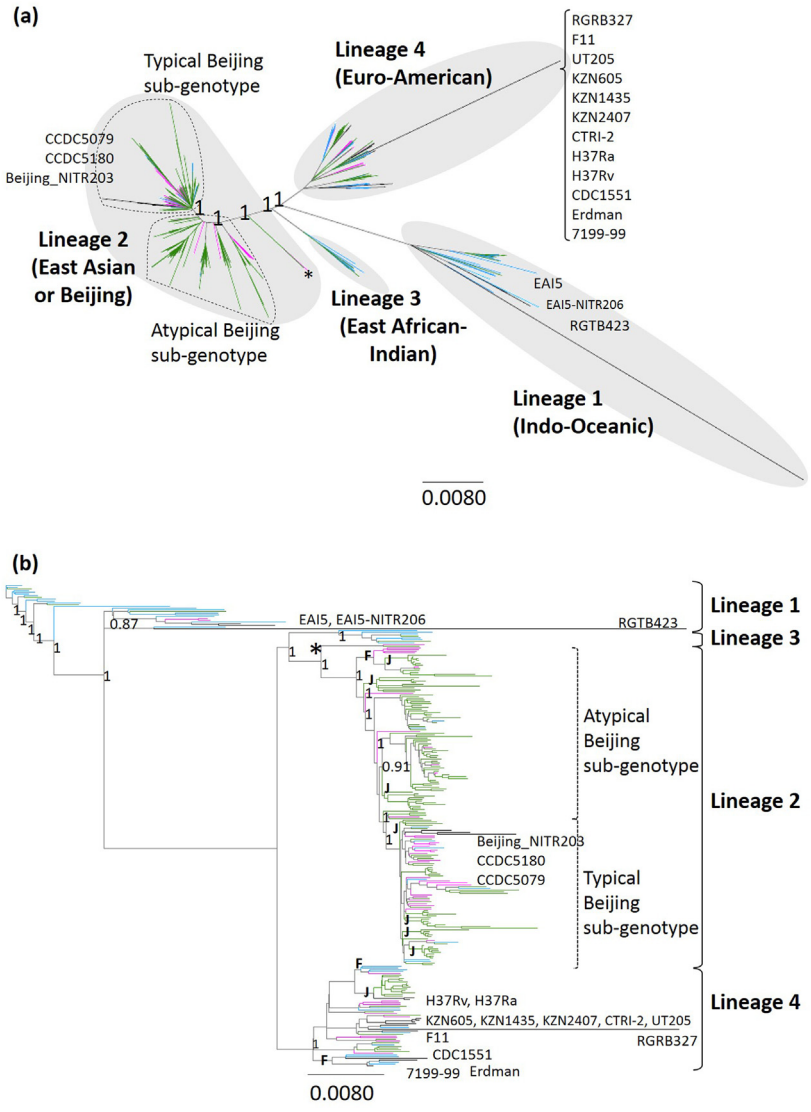

FIG. 1. Maximum-likelihood phylogenic trees of 259 clinical isolates and 18 published strains based on whole genome analysis. Radial and rectangular tree layouts are shown in (a) and (b), respectively. Branches are colored according to isolates from foreign-born patients (pink: born in China, Korea, Mongolia and Taiwan; blue: other countries and regions), Japan-born (green) and published strains (black). Clinical isolates in oval (a) and square brackets (b) belong to indicated lineages/sub-lineages. Bootstrap values are shown in the phylogenic trees. All lineages and two sub-genotypes in Lineage 2 were identified as clusters of isolates with over 0.7 bootstrap support. *Two isolates in (a) and (b) that could be categorized as of the atypical Beijing sub-lineage, but Rv0069-0075 was deleted rather than RD105. "F" and "J" in (b) indicate clusters comprising isolates specific to foreign- or Japan-born patients, respectively. Scale bar indicates nucleotide substitutions per site.

distribution of isolates from East Asia-, non-East Asia-, and Japan-born patients among lineages was detected between 2001–2008 and 2009–2012 (Supplementary Fig. 3), although the proportion of isolates from East Asia- and Japan-born patients belonging to typical Beijing sub-genotype in Lineage 2 decreased, and that from Japan-born patients belonging to atypical Beijing sub-genotype in Lineage 2 increased, from 2001–2008 to 2009–2012.

TABLE 2. Distribution of isolates from foreign- and Japan-born patients among *Mycobacterium tuberculosis* lineages

	Isolates from patients				
	Foreign born		Japan born		
Lineage	*n*	%	*n*	%	p[a]
Lineage 2	43	47.3	139	82.7	5.80E-09
Typical Beijing	29	67.4	52	37.2	1.01E-03
Atypical Beijing	14	32.6	87	62.8	1.01E-03
Lineage 4	23	25.3	23	13.7	0.030891
Lineage 1	19	20.9	4	2.4	1.87E-06
Lineage 3	6	6.6	2	1.2	0.043065

[a]Statistical analysis done using Fisher's exact test.

Foreign-born patients infected with typical Beijing sub-genotype isolates in Lineage 2 were from eight countries: China (*n* = 14), Korea (*n* = 4), Mongolia (*n* = 3), Vietnam (*n* = 3), Burma (Myanmar) (*n* = 2), Indonesia (*n* = 1), India (*n* = 1), and France (*n* = 1), whereas foreign-born patients infected with typical Beijing sub-genotype isolates in Lineage 2 were from four countries: China (*n* = 5), Korea (*n* = 7), Iran (*n* = 1), and the USA (*n* = 1). Isolates belonging to Lineage 1 (Indo-Oceanic, *n* = 23) were much more common in foreign-born patients; of the 19 patients infected with these isolates, eight were from the Philippines, six from Burma (Myanmar), two from Nepal, and one each from China, Taiwan, and Vietnam.

Drug susceptibility

Among the 259 clinical isolates, 36 (13.9%) were resistant to at least one drug: isoniazid (n = 6), isoniazid + streptomycin (n = 4), isoniazid + streptomycin + ethambutol (n = 3), isoniazid + streptomycin + ethambutol + ethionamide (n = 2), streptomycin (n = 14), rifampicin + streptomycin (n = 1), rifampicin + levofloxacin (n = 1), ethionamide (n = 1), levofloxacin (n = 3), and para-amino salicylic acid (n = 1). The majority of these drug-resistant isolates had a mutation(s) associated with isoniazid-resistance; i.e. all 15 isoniazid-resistant isolates had a mutation including −15c→t upstream of *mabA*, S315T in *katG*, and g609a in *mabA*; one of two rifampicin-resistant isolates had S450L in *rpoB*, but another had no mutation in *rpoB*; and two of four levofloxacin-resistant isolates had a mutation in A90V in *gyrA*, but the remaining two had no mutation in *gyrA*. There was no multi-drug-resistant isolates (resistant to isoniazid + rifampicin). Most drug-resistant isolates' drug resistance was greater for the strains in Lineage 2 (East Asian or Beijing n = 31, 17.0%) than the non-Lineage 2 (n = 5, 6.5%) (p <0.05, Fisher's exact test). There was no significant difference in the proportions of typical (10 of 81, 12.3%) and atypical Beijing sub-genotype in Lineage 2 (21 of 101, 20.8%) isolates resistant to drug. Of the isolates from foreign- and Japan-born patients, 8 (8.8%) and 28 (16.7%), respectively, were resistant to at least one drug, a difference that was not statistically significant.

Whole genome SNP analysis

A total of 359 240 SNPs and indels were called across the 259 isolates, using manual setting (coverage >11 and frequency >95%). Of these SNPs and indels, 315 166 (87.7%) were coding region SNPs and indels. Of these 315 166 coding region SNPs and indels, 202 999 non-synonymous substitutions (64.4%) were identified. The numbers of SNPs and indels detected in individual isolates ranged from 338 to 2321. Little difference (p >0.05, Mann-Whitney U test) was observed between the numbers of the SNPs and indels of the isolates from foreign-born (median 1494, range 608–2278) and Japan-born (median 1479.5, range 338–2321) patients.

Three SNPs (two coding regions and one intergenic region) and one indel against H37Rv were shared in common across all 259 isolates. These two coding regions SNPs included a synonymous substitution of a477g in Rv1514c and a non-synonymous substitution of a430c in Rv1321. The SNP in the intergenic region was an a-to-c substitution at H37Rv position 459 399. The shared indel across all the isolates was a frameshift mutation in Rv3655c (deletion at H37Rv position 4 095 002).

All 81 of the typical Beijing sub-genotype isolates in Lineage 2 had eight SNPs in their coding regions, whereas the remaining 178 isolates, belonging to the non-Beijing sub-genotype and the other lineages, did not. The eight SNPs specific to the typical Beijing isolates included five non-synonymous (Rv0245 (c308t), Rv1027c (g178a), Rv1160 (*mutT2*) (g172c), Rv2180c (g747a), and Rv3758c (a250g)) and three synonymous (Rv0044c (a525g), Rv0237 (c1075t), and Rv0512 (g549a)) SNPs. However, no SNP specific to atypical Beijing isolates was identified.

Of the 259 isolates, 244 had "strain-specific SNPs and indels," with each unique to one strain. The numbers of these strain-specific SNPs and indels in individual isolates ranged from 1 to 520. The numbers of strain-specific SNPs and indels differed significantly (p 3.66E-15 Mann-Whitney test) between isolates from foreign-born (median 89, range 0–520) (Fig. 2(a)) and Japan-born (median 23, range 0–415) patients (Fig. 2(b)). These between-group differences were observed when comparing isolates belonging to the typical Beijing (p = 1.05E-4), atypical Beijing (p = 0.02), and Lineage 4 (Euro-American) (p = 1.73E-6) (Fig. 2(c)) lineages.

Discussion

To our knowledge, this study is the first to show an impact of foreign-born TB patients on genetic diversity and bacteriological properties of *M. tuberculosis* isolates in a middle TB burden country such as Japan. The genomic features of *M. tuberculosis* isolates differed between foreign-born and Japan-born patients in Japan, due to differences between the two groups in their lineage distributions and the numbers of strain-specific SNPs and indels. These results indicate that *M. tuberculosis* isolates from foreign-born patients were not transmitted extensively to Japan-born patients, although isolates mainly belonging to clusters in Lineage 2 and Lineage 4 would transmit between these two groups (Fig. 1(b)) and these isolates had the potential risk of spreading to others in Japan. The larger numbers of strain-specific SNPs and indels in the isolates from foreign-born patients also indicated that the isolates originated from many countries, not only Japan, although this would be expected even if it had not been reported.

Several reports have described epidemiological differences between foreign-born and native-born TB patients, although these studies were performed in countries with low TB prevalence. Most of these studies found little transmission of TB from foreign-born to native-born residents. *M. tuberculosis* isolates from foreign-born patients were much less likely to be clustered, with extensive transmission due mainly to autochthonous strains in Madrid [19]. Although immigrants did not transmit TB more than indigenous patients in Catalonia, Spain [20], the most frequent infectious disease in immigrants to Spain was latent TB (32.6%) [21]. US-born TB patients generated more secondary cases than TB-positive immigrants in San Francisco (http://www.ingentaconnect.com/content/iuatld/ijtld/2000/00000004/00000004/art00002). Although foreign-born

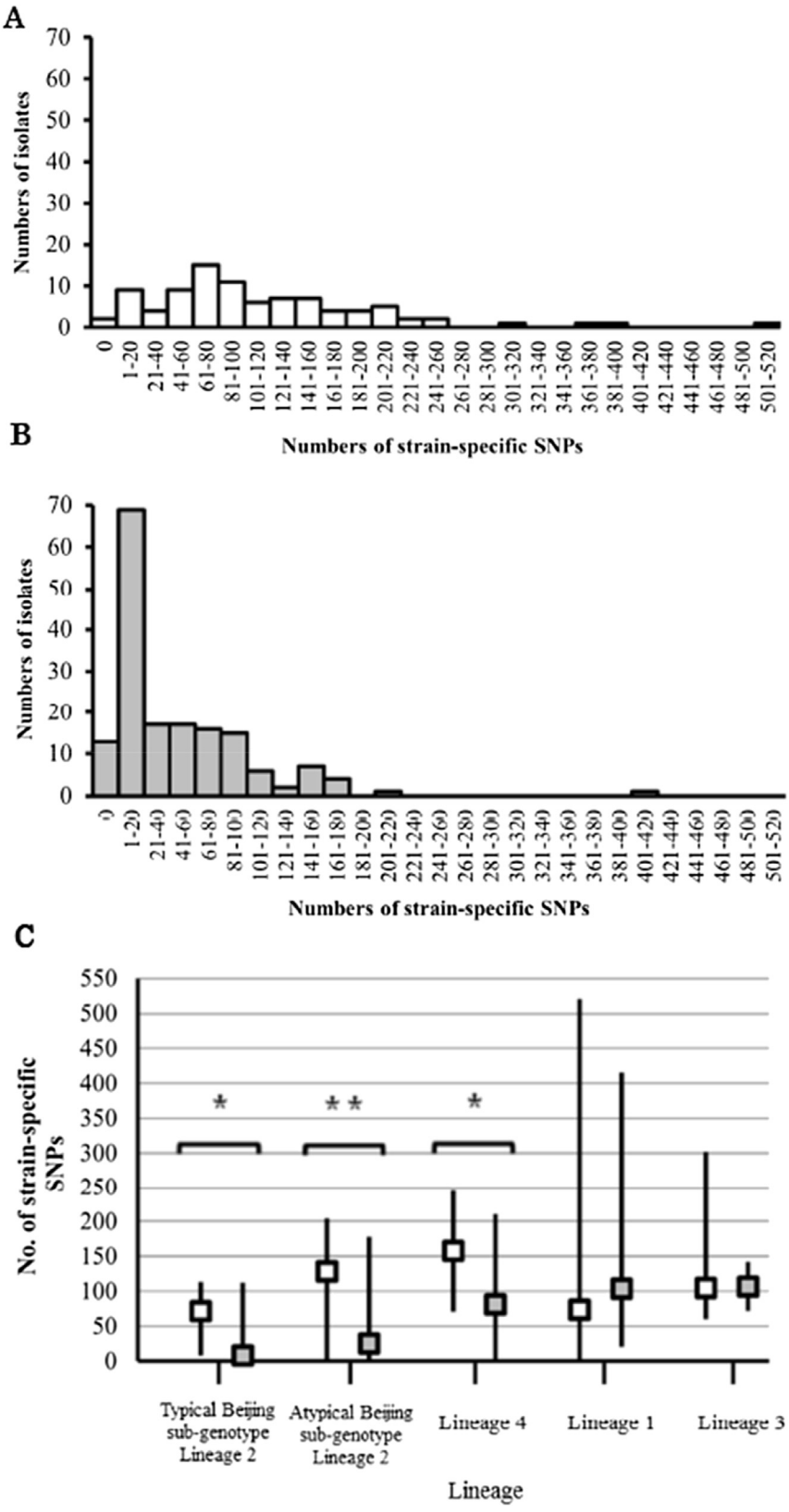

FIG. 2. Numbers of strain-specific single-nucleotide polymorphisms (SNPs) and insertions and deletions (indels) in *Mycobacterium tuberculosis* isolates from (A) foreign-born and (B) Japan-born patients. Of the 259 isolates, 244 had strain-specific SNPs and indels. (C) Minimum to maximum number of strain-specific SNPs and indels of isolates in Lineage 1 (Indo-Oceanic), Typical/Atypical Beijing sub-genotype of Lineage 2 (East Asian or Beijing), Lineage 3 (East African-Indian), and Lineage 4 (Euro-American) are represented by black lines. Median numbers of strain-specific SNPs and indels from foreign- and Japan-born patients are represented as white and grey squares, respectively. Significant differences among foreign-/Japan-born groupings were present in isolates belonging to the typical/atypical Beijing sub-genotype in Lineage 2 and the Lineage 4. *p <0.01; **p <0.05.

TB patients constituted the majority of these patients in Canada and Switzerland, there was little transmission to native-born individuals [22,23]. TB also showed a low level of transmission from immigrants to Germany with a high prevalence of disease to the autochthonous population in Germany, with a low prevalence of disease [24].

This is the first report describing the whole genome sequence-based molecular epidemiology of *M. tuberculosis* isolates obtained from foreign-born patients in Japan, although there has been a comprehensive study on *M. tuberculosis* in an urban area in Japan [25]. The lineage proportion of isolates from our Japan-born patients was similar to that of previously reported studies [26–30]; most belonged to the Beijing lineage, especially the atypical Beijing sub-genotype. The proportion of the atypical Beijing sub-genotype isolates is unique and endemic to Japan and Korea, whereas the typical Beijing sub-genotype is more common among other Asian countries [27,31].

We detected eight SNPs present in all 81 typical Beijing sub-genotype isolates in Lineage 2 tested, but not in any of the other 178 isolates (Supplementary Table 5). Of these eight SNPs, three—Rv0044c (a525g), Rv3758c (a250g), and Rv1160 (*mutT2*) (g172c)—were not included in the set of 51 SNPs unique to typical Beijing strains [32]. Differences in detected SNPs specific to typical Beijing sub-genotype isolates in Lineage 2 will be due to the size/setting of isolates examined in these studies, e.g. the whole genome sequences of three typical Beijing sub-genotype isolates were compared with those of three atypical Beijing sub-genotype isolates in a previous study [32]. Further studies will be necessary to determine whether these SNPs are unique to typical Beijing sub-genotype strains.

The limitation of this study is that we did not get information about foreign-born patients, including social activities in Japan. We, therefore, could not demonstrate whether or not some strains were transmitted between foreign- and Japan-born patients in Japan.

In conclusion, we showed that the proportions of genetic lineages of *M. tuberculosis* isolates differed in foreign-born and Japan-born patients in Tokyo, with the genetic diversity of the former being greater, indicating that foreign-born TB patients in Tokyo carry many different strains reflecting the global distribution of *M. tuberculosis* diversity. There were no prevalent isolates from foreign-born TB patients in Tokyo. However, globally disseminated isolates may be introduced into Japan by foreign-born TB patients, especially those from countries with high TB burdens.

Transparency declaration

This work was supported by a grant for International Health Research (24S-103 and 26A-103) and Health Sciences Research grants (H23-SHINKO-IPPAN-002 and H26-SHINKO-IPPAN-004) from the Ministry of Health, Labour, and Welfare of Japan. The authors report no conflicts of interest relevant to this article.

Appendix A. Supplementary data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.cmi.2014.09.013.

References

[1] World Health Organization. Global tuberculosis report 2013. Geneva: World Health Organization; 2013.
[2] Comas I, Chakravartti J, Small PM, Galagan J, Niemann S, Kremer K, et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. Nat Genet 2010;42:498–503.
[3] Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. Nat Genet 2013;45:1176–82.
[4] Sekiguchi J-I, Nakamura T, Miyoshi-Akiyama T, Kirikae F, Kobayashi I, Augustynowicz-Kopec E, et al. Development and evaluation of a line probe assay for rapid identification of *pncA* mutations in pyrazinamide-resistant *Mycobacterium tuberculosis* strains. J Clin Microbiol 2007;45:2802–7.
[5] Miyoshi-Akiyama T, Takeshita N, Ohmagari N, Kirikae T. Complete genome sequence of *Helicobacter cinaedi* type strain ATCC BAA-847. J Bacteriol 2012;194:5692.
[6] Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol 2004;5:R12.
[7] Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30:3059–66.
[8] Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 2010;59:307–21.
[9] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9.
[10] Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, et al. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 2006;103:2869–73.
[11] Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis 2007;7:328–37.
[12] Kurepina NE, Sreevatsan S, Plikaytis BB, Bifani PJ, Connell ND, Donnelly RJ, et al. Characterization of the phylogenetic distribution and chromosomal insertion sites of five IS6110 elements in *Mycobacterium tuberculosis*: non-random integration in the *dnaA-dnaN* region. Tuber Lung Dis 1998;79:31–42.
[13] Kremer K, Glynn JR, Lillebaek T, Niemann S, Kurepina NE, Kreiswirth BN, et al. Definition of the Beijing/W lineage of *Mycobacterium tuberculosis* on the basis of genetic markers. J Clin Microbiol 2004;42:4040–9.

[14] Plikaytis BB, Marden JL, Crawford JT, Woodley CL, Butler WR, Shinnick TM, et al. Multiplex PCR assay specific for the multidrug-resistant strain W of *Mycobacterium tuberculosis*. J Clin Microbiol 1994;32:1542–6.

[15] Faksri K, Drobniewski F, Nikolayevskyy V, Brown T, Prammananan T, Palittapongarnpim P, et al. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on IS6110, SNP, LSP and VNTR profiles from Thailand. Infect Genet Evol 2011;11:1142–9.

[16] Rindi L, Lari N, Cuccu B, Garzelli C. Evolutionary pathway of the Beijing lineage of *Mycobacterium tuberculosis* based on genomic deletions and *mutT* genes polymorphisms. Infect Genet Evol 2009;9:48–53.

[17] Ebrahimi-Rad M, Bifani P, Martin C, Kremer K, Samper S, Rauzier J, et al. Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family. Emerg Infect Dis 2003;9:838–45.

[18] Hanekom M, van der Spuy GD, Streicher E, Ndabambi SL, McEvoy CR, Kidd M, et al. A recently evolved sublineage of the *Mycobacterium tuberculosis* Beijing strain family is associated with an increased ability to spread and cause disease. J Clin Microbiol 2007;45:1483–90.

[19] Iñigo J, García de Viedma D, Arce A, Palenque E, Alonso Rodríguez N, Rodríguez E, et al. Analysis of changes in recent tuberculosis transmission patterns after a sharp increase in immigration. J Clin Microbiol 2007;45:63–9.

[20] Godoy P, Caylà JA, Carmona G, Camps N, Álvarez J, Rodés A, et al. Immigrants do not transmit tuberculosis more than indigenous patients in Catalonia (Spain). Tuberculosis 2013;93:456–60.

[21] Monge-Maillo B, Jiménez BC, Pérez-Molina JA, Norman F, Navarro M, Pérez-Ayala A, et al. Imported infectious diseases in mobile populations. Spain. Emerg Infect Dis 2009;15:1745–52.

[22] Kunimoto D, Sutherland K, Wooldrage K, Fanning A, Chui L, Manfreda J, et al. Transmission characteristics of tuberculosis in the foreign-born and the Canadian-born populations of Alberta, Canada. Int J Tuberc Lung Dis 2004;8:1213–20.

[23] Fenner L, Gagneux S, Helbling P, Battegay M, Rieder HL, Pfyffer GE, et al. *Mycobacterium tuberculosis* transmission in a country with low tuberculosis incidence: role of immigration and HIV infection. J Clin Microbiol 2012;50:388–95.

[24] Barniol J, Niemann S, Louis VR, Brodhun B, Dreweck C, Richter E, et al. Transmission dynamics of pulmonary tuberculosis between autochthonous and immigrant sub-populations. BMC Infect Dis 2009;9:197.

[25] Ohkado A, Nagamine M, Murase Y, Uchimura K, Kaguraoka S, Tatsumi Y, et al. Molecular epidemiology of *Mycobacterium tuberculosis* in an urban area in Japan, 2002–2006. Int J Tuberc Lung Dis 2008;12:548–54.

[26] Tamaru A, Nakajima C, Wada T, Wang Y, Inoue M, Kawahara R, et al. Dominant incidence of multidrug and extensively drug-resistant specific *Mycobacterium tuberculosis* clones in Osaka Prefecture, Japan. PLoS One 2012;7:e42505.

[27] Wada T, Fujihara S, Shimouchi A, Harada M, Ogura H, Matsumoto S, et al. High transmissibility of the modern Beijing *Mycobacterium tuberculosis* in homeless patients of Japan. Tuberculosis (Edinb) 2009;89:252–5.

[28] Wada T, Iwamoto T, Maeda S. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family in East Asia revealed through refined population structure analysis. FEMS Microbiol Lett 2009;291:35–43.

[29] Wada T, Maeda S, Hase A, Kobayashi K. Evaluation of variable numbers of tandem repeat as molecular epidemiological markers of *Mycobacterium tuberculosis* in Japan. J Med Microbiol 2007;56:1052–7.

[30] Glynn JR, Kremer K, Borgdorff MW, Rodriguez MP, van Soolingen D. Beijing/W genotype *Mycobacterium tuberculosis* and drug resistance. Emerg Infect Dis 2006;12:736–43.

[31] Chen YY, Chang JR, Huang WF, Kuo SC, Su IJ, Sun JR, et al. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on SNP and VNTR typing profiles in Asian countries. PLoS One 2012;7:e39792.

[32] Schürch AC, Kremer K, Warren RM, Hung NV, Zhao Y, Wan K, et al. Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the *Mycobacterium tuberculosis* Beijing genotype. Infect Genet Evol 2011;11:587–97.