JCM
Journals.ASM.org

Genetic Diversity and Dynamic Distribution of *Mycobacterium tuberculosis* Isolates Causing Pulmonary and Extrapulmonary Tuberculosis in Thailand

Prapaporn Srilohasin,[a] Angkana Chaiprasert,[a] Katsushi Tokunaga,[b] Nao Nishida,[c] Therdsak Prammananan,[d] Nat Smittipat,[d] Surakameth Mahasirimongkol,[e] Boonchai Chaiyasirinroje,[f] Hideki Yanai,[f,g] Prasit Palittapongarnpim[h]

Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand[a]; Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan[b]; Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Chiba, Japan[c]; National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency, Ministry of Science and Technology, Pathum Thani, Thailand[d]; Medical Genetics Section, National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand[e]; TB/HIV Research Project, Research Institute of Tuberculosis, Chiang Rai, Thailand[f]; Fukujuji Hospital, Japan Anti-Tuberculosis Association, Kiyose, Tokyo, Japan[g]; Department of Microbiology, Faculty of Science, Mahidol University, Bangkok, Thailand[h]

This study examined the genetic diversity and dynamicity of circulating *Mycobacterium tuberculosis* strains in Thailand using nearly neutral molecular markers. The single nucleotide polymorphism (SNP)-based genotypes of 1,414 culture-positive *M. tuberculosis* isolates from 1,282 pulmonary tuberculosis (PTB) and 132 extrapulmonary TB (EPTB) patients collected from 1995 to 2011 were characterized. Among the eight SNP cluster groups (SCG), SCG2 (44.1%), which included the Beijing (BJ) genotype, and SCG1 (39.4%), an East African Indian genotype, were dominant. Comparisons between the genotypes of *M. tuberculosis* isolates causing PTB and EPTB in HIV-negative cases revealed similar prevalence trends although genetic diversity was higher in the PTB patients. The identification of 10 reported sequence types (STs) and three novel STs was hypothesized to indicate preferential expansion of the SCG2 genotype, especially the modern BJ ST10 (15.6%) and ancestral BJ ST19 (13.1%). An association between SCG2 and SCG1 genotypes and particular patient age groups implies the existence of different genetic advantages among the bacterial populations. The results revealed that increasing numbers of young patients were infected with *M. tuberculosis* SCGs 2 and 5, which contrasts with the reduction of the SCG1 genotype. Our results indicate the selection and dissemination of potent *M. tuberculosis* genotypes in this population. The determination of heterogeneity and dynamic population changes of circulating *M. tuberculosis* strains in countries using the *Mycobacterium bovis* BCG (bacillus Calmette-Guérin) vaccine are beneficial for vaccine development and control strategies.

Tuberculosis (TB) remains one of the most infectious and deadly diseases worldwide and has significant medical, social, and economic impacts. The *Mycobacterium bovis* BCG (bacillus Calmette-Guérin) vaccine was first used in humans in 1921, while the first effective anti-TB drug, isoniazid, was developed in 1952 (1). The incidence of TB has declined with the worldwide distribution of the BCG vaccine and improving living conditions (2); however, the protective effect of the vaccine remains controversial. The likely factors influencing the degree of disease variation, including host and environmental factors, and/or mycobacterial traits are largely unknown (3). There are several lines of evidence that indicate that genetic variation in *Mycobacterium tuberculosis* contributes to the ambiguities concerning disease presentation, frequency of transmission, BCG vaccine response, and treatment outcome (4–6). Host factors, immune status, nutrition, and genetic polymorphisms are known to affect BCG vaccine efficacy (7, 8). The protective effect of the Glaxo freeze-dried BCG vaccine was more than 75% in the United Kingdom but displayed low efficiency in Malawi (8). Genetic evolution trees of humans and *M. tuberculosis* strongly indicate an adaptation of TB bacteria to the changing human population (3). Improving TB control requires a better understanding of the impact of the environment on the interaction between the pathogen and its host. Like other high-TB-burden countries, Thailand has officially included the BCG vaccine in the national vaccination program since 1977. However, the production of BCG vaccine was started and held in a small-scale setting in 1953 by the Queen Saovabha Memorial Institute (QSMI). In 1987, the BCG vaccine strain in Thailand was changed from a Danish strain to a Tokyo strain, and the dose was also changed from two doses to one dose at birth; 100% coverage of BCG vaccination in Thailand was achieved in 1990 (9, 10). Although the BCG vaccine appears to provide a high level of protection against tuberculous meningitis and disseminated TB in children, Beijing (BJ) strains have been reported as the most common genotype causing both pulmonary tuberculosis (PTB) and tuberculous meningitis in Thailand (6, 11, 12). Therefore, BJ strains might be postulated to represent one of the most common Thai-adapted genotypes in the mass vaccination environment.

Several different molecular typing methods have been used to study genetic diversity within *M. tuberculosis* complex (MTBC)

Received 23 May 2014 Returned for modification 26 June 2014
Accepted 19 September 2014
Published ahead of print 8 October 2014
Editor: K. C. Carroll
Address correspondence to Angkana Chaiprasert, angkana.cha@mahidol.ac.th.
This work is dedicated to Her Royal Highness Princess Galyanivadhana KromLaung Narathiwas Rajnakarindth on her 90th birthday.

doi:10.1128/JCM.01467-14

TABLE 1 Incidence of PTB and EPTB cases in Thailand from 1995 to 2011 by age group and HIV status

Age group (yr)	No. of PTB patients (%)			No. of EPTB patients (%)			Total no. of patients (%)
	HIV^{-a}	HIV^{+a}	Unknown	HIV^{-}	HIV^{+}	Unknown	
0–10	4 (0.4)	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.8)	11 (11.8)	16 (1.1)
11–20	61 (5.7)	1 (0.5)	1 (5.6)	1 (33.3)	1 (2.8)	4 (4.3)	69 (4.9)
21–30	142 (13.4)	54 (26.6)	2 (11.1)	0 (0.0)	14 (38.9)	22 (23.7)	234 (16.5)
31–40	171 (16.1)	98 (48.3)	0 (0.0)	1 (33.3)	13 (36.1)	31 (33.3)	314 (22.2)
41–50	227 (21.4)	35 (17.2)	7 (38.9)	1 (33.3)	4 (11.1)	12 (12.9)	286 (20.2)
51–60	182 (17.2)	14 (6.9)	1 (5.6)	0 (0.0)	2 (5.6)	7 (7.5)	206 (14.6)
61–70	146 (13.8)	1 (0.5)	2 (11.1)	0 (0.0)	0 (0.0)	3 (3.2)	152 (10.7)
71–80	105 (9.9)	0 (0.0)	4 (22.2)	0 (0.0)	1 (2.8)	2 (2.2)	112 (7.9)
>80	23 (2.2)	0 (0.0)	1 (5.6)	0 (0.0)	0 (0.0)	1 (1.1)	25 (1.8)
Total	1,061 (82.8)	203 (15.8)	18 (1.4)	3 (2.3)	36 (27.3)	93 (70.5)	1,414 (100.0)

[a] HIV status is indicated as negative (−) or positive (+).

clinical isolates. Because MTBC isolates contain less genetic diversity than many other bacteria, molecular markers that reveal individual and group differences of the strains under investigation have to be selected according to the study purpose. The unique rate of change of each marker is an important factor in marker selection. Repetitive sequences evolve more quickly than single nucleotide polymorphisms (SNPs) and large-sequence polymorphisms. Therefore, evolutionary studies do not commonly use repetitive sequences. The unique polymorphic properties of SNPs are a valuable measurement for categorizing *M. tuberculosis* clinical isolates into genetically related groups. Recently, we developed a high-throughput DNA chip that provides genotyping results congruent with corresponding phylogenies inferred from large-sequence polymorphisms, spoligotyping, mycobacterial interspersed repetitive-unit–variable-number tandem repeat analysis based on 12 loci, principal genetic groups based on *katG463-gyrA*95 polymorphisms, and another SNP set (13, 14). In particular, *M. tuberculosis* snip cluster group 1 (SCG1) and SCG2 strains were completely comparable to East African-Indian (EAI)/Indo-Oceanic and BJ/East Asia genotypes, respectively (14). Interestingly, neutral genetic variations across modern *M. tuberculosis* populations are categorized into different groups based on spoligotype, IS*6110* sequence, or mycobacterial interspersed repetitive-unit–variable-number tandem repeat typing, in which outlier genetic diversity is commonly caused by genetic convergence.

In Asian countries such as China (15), Myanmar (16), Japan (17), South Korea (18), Taiwan (19), Thailand (6, 11, 12), and Vietnam (20, 21), *M. tuberculosis* strains belonging to the BJ and EAI genotypes are predominant. In Europe and America, *M. tuberculosis* Haarlem (H), S, X, T, and Latin American and Mediterranean (LAM) genotypes, alternatively known as the Euro-American lineage, have been predominantly isolated from TB-positive patients. In addition, evolution and phylogeography studies indicate that EAI strains represent the ancestral *M. tuberculosis* genotypes and are linked to ancestral African predecessors (3). For BJ genotype strains, several human-mycobacterial coevolution studies have provided evidence of BCG vaccine impact on the selection of *M. tuberculosis* isolates (22). Differential protein expression of the BJ genotype contributes to heterogeneous immune responses (23). In countries where BCG vaccination is uncommon, such as Ethiopia, the genetic diversity of *M. tuberculosis* strains is higher than in countries where it is common, such as Tunisia. It is therefore likely that selection of *M. tuberculosis* strains resistant to BCG vaccine-induced immunity has occurred (24).

This study aimed to assess the genetic diversity and dynamicity of the *M. tuberculosis* population in different age groups of patients in Thailand using the aforementioned high-throughput DNA chip, which is based on the DigiTag2 platform (13). The results provide useful, up-to-date epidemiological data on tuberculosis and may be able to help predict future trends in the prevalence of certain genotypes.

MATERIALS AND METHODS

Study population and patient data. A total of 1,414 culture-positive *M. tuberculosis* isolates from 1,282 PTB and 132 extrapulmonary TB (EPTB) patients in Thailand were characterized in this study (Table 1). PTB is defined as a case of TB in which the patient has disease only in the lungs, whereas in EPTB cases, the disease involves at least one nonpulmonary site. All isolates causing PTB and two causing EPTB (one from a patient with lymph node infection and one from a patient with TB pleurisy) were obtained from 13 hospitals covering 13 of 18 districts in Chiang Rai Province, Thailand, from 1998 to 2011. To increase the number of isolates associated with cases of EPTB, which are extremely rare, an additional 130 previously collected *M. tuberculosis* isolates from cases of meningitis, collected from 1995 to 2008, were retrieved from the Molecular Mycology and Mycobacteriology Laboratory (Drug-Resistant Tuberculosis Research Fund), Department of Microbiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand, and included in this study (6, 11). Demographic data, including age and human immunodeficiency virus (HIV) status, were collected by physicians and trained medical staff. For determining the changing prevalence of *M. tuberculosis* genotypes across the patients from the pre- and post-BCG vaccine era, the genotype of 396 *M. tuberculosis* isolates from PTB patients who were born prior to 1953 and of 873 isolates from PTB patients born postintroduction of the BCG vaccine were recruited. Because approximately 70% of EPTB patients were of unknown HIV status, only the genotypes of the 1,269 isolates from PTB patients with HIV-negative status who were 15 to 65 years old were analyzed. Isolates from patients in this group who were born before and after 1953 were analyzed in parallel. The ethical and scientific committees of the Faculty of Medicine, Siriraj Hospital, Mahidol University, approved the study protocols (EC no. 759/2551 and EC no. 604/2555).

SNP genotyping. The sequence types (STs) and SCGs of all *M. tuberculosis* isolates were determined using the MTBC-targeting DigiTag2 assay, based on 51 highly informative SNPs (13, 14). The first four SNPs were used to screen MTBC species. The next 47 SNPs were chosen for genotyping *M. tuberculosis* isolates, with 2 SNP positions for determina-

tion of principal genetic groups and 45 SNPs for characterizing *M. tuberculosis* isolates into six SCGs (14, 25). Based on the DNA sequence of *M. tuberculosis* H37Rv (GenBank accession no. NC_000962), SNPs at either position 43943 or 3440542 were able to discriminate between ancestral (TbD1[+]) and modern (TbD1[−]) *M. tuberculosis* genotypes. In addition, the polymorphisms at nucleotide 4280708 are the same as the absence or presence of RD105 and were used to define the BJ genotype (13). The SNP at position 1477598 is the same as the *ogt12* polymorphism and was used to classify the ancestral and modern BJ genotypes (26). The perfect concordance of the evolutionary events between either SNP 797736 or SNP 2825581 and the RD181 marker was also used to subdivide the ancestral BJ genotype into early and late genotypes (27).

Genomic DNA from all *M. tuberculosis* isolates was extracted using the cetyltrimethylammonium bromide (CTAB) method (28). PCRs (10 μl) contained 20 to 40 ng of bacterial DNA, 25 nM each primer, and 1× KAPA2G Robust HotStart Ready Mix with dye (Kapa Biosystems, Woburn, MA, USA). Two-step multiplex PCR for increasing target SNPs was performed for 15 min at 95°C, followed by 40 cycles of 95°C for 90 s and 68°C for 2 min using a TGradient thermal cycler (Biometra, Göttingen, Germany). In the ligation reaction, the perfect-match 3′ query-cD1s probes were first phosphorylated at 37°C for 30 min and then incubated at 95°C for 3 min in a phosphorylation mixture containing 1× protruding-end kinase buffer, 30 mM ATP, 40 U of polynucleotide kinase, and 1 μM each 3′ query probe (Kination kit, Toyobo, Osaka, Japan) using a TGradient thermal cycler (Biometra). Next, 100 nM mismatch-induced 5′ query cEDs probes were mixed with 100 nM phosphorylated perfect-match 3′ query cD1s probes for the encoding reaction. The encoding mixture, containing 1 μl of multiplex PCR product, 50 nM each 5′ query cED and 3′ query cD1s probe mixture, 1× *Taq* DNA ligase reaction buffer, and 10 U of T4 DNA ligase (New England BioLabs, Beverly, MA, USA), was incubated at 95°C for 5 min and at 59°C for 15 min and then held at 10°C using a TGradient thermal cycler (Biometra). For the labeling reaction, the mixture (12-μl total volume) contained 6 μl of ligation product, 3 nM each D1s primer, 0.5 μM (each)Alexa 647-ED-2 and Alexa 555-ED-1 primer, 1.5× KAPA2G Fast buffer, 4.5 mM Mg^{2+}, 0.2 mM (each) deoxynucleoside triphosphate (dNTP), and 0.4 U of KAPA2G Fast HotStart DNA polymerase (Kapa Biosystems). The reaction mixture was cycled for 15 min at 95°C for the initial denaturation step, followed by 30 cycles of 95°C for 1 min, 55°C for 6 min, and 72°C for 30 s, before being held at 10°C using a TGradient thermal cycler (Biometra) (13). To define the SNP sequence type of each position on a DNA microarray, fluorescent signal intensities were evaluated using SNPStar software (version 0.0.0.8; Olympus, Tokyo, Japan). The SNP profiles indicating either the ST or SCG were compared with previous data (14, 26, 29). When necessary, the novel STs were confirmed by direct sequencing along with previous spoligotyping results (N. Smittipat, personal communication).

Statistical analysis. Descriptive statistics of patient clinical and demographic data, including patient age, HIV status, and disease phenotypes, are presented as the number, percentage, or median, as appropriate. The association between bacterial genotypes among the PTB and EPTB groups was evaluated using a chi-square test. A P value of ≤0.05 was considered statistically significant. Associations between the bacterial genotypes and pre- and post-BCG vaccination eras among all TB cases were determined by calculating the odds ratio with 95% confidence intervals (CI).

RESULTS

Genetic diversity of the studied *M. tuberculosis* PTB- and EPTB-causing strains based on SNP typing. The age of the PTB and EPTB patients ranged from 9 to 104 years (median, 45 years) and from 3 months to 83 years (median, 33 years), respectively. The HIV status was known in most of the PTB cases (n = 1,264). Unfortunately, for approximately 70% of the EPTB cases (n = 93), the HIV status was unknown (Table 1). Based on comparison with the global collection of *M. tuberculosis* strains, the 1,414 clinical *M. tuberculosis* isolates were classified into eight SCGs: including SCG1 (39.4%), -2 (44.1%), -3a (1.6%), -3b (1.6%), -3c (0.1%), -4 (0.1%), -5 (11.8%), and -6a (0.6%). Among this collection, we found 28 STs, including nine novel STs (named STTh1 to STTh9) (Table 2). Interestingly, the incidence of SCG3a (1.7%) was higher than that of SCG3b (1.5%) among PTB patients while SCG3a was not found among the EPTB patients, and SCG3b was detected in a higher percentage of samples (3.0%). The population size may affect this finding, as well as the greater variation observed in *M. tuberculosis* genotypes of the PTB patients (27 STs) than in the EPTB patients (13 STs). STF was not identified in PTB patients but was found in EPTB patients. Conversely, STTh1 to STTh9 and STs 3, 9, 11, 21, 25, and 41 were detected only in PTB patients. However, the prevalence trend of *M. tuberculosis* infections causing PTB and EPTB in this study displayed no significant association between the site of disease and the SCG of the infecting bacterial isolate (P = 0.07).

Interestingly, our results revealed that in Thailand, *M. tuberculosis* SCG2 accounted for a large proportion of the strains and also displayed the highest genetic diversity (13 STs). There was a preferential expansion of modern ST10 (15.6%) and ancestral ST19 (13.1%) strains of the BJ genotype, which indicates either good adaptation or high efficiency of transmission. The few predominant *M. tuberculosis* subtypes in this population indicate that the Thai-adapted *M. tuberculosis* genotypes have a selective advantage. Moreover, 13 isolates from PTB patients were heterozygous mixed-strain *M. tuberculosis* infections that included ST15 plus ST3 (1 isolate), ST15 plus ST10 (2 isolates), ST15 plus ST19 (1 isolate), ST15 plus ST18 (1 isolate), ST15 plus ST24 (2 isolates), ST10 plus ST24 (1 isolate), ST19 plus ST24 (1 isolate), ST25 plus ST24 (2 isolates), and ST18 plus ST19 (2 isolates). These findings were very similar to the spoligotyping results (data not shown).

Dynamic change in *M. tuberculosis* genotypes between different patient age groups. To clarify whether *M. tuberculosis* co-adapted with its host to cause disease, the distribution of bacterial genotypes within the different patient age groups was determined. Patient ages were obtained from hospital registration records. We divided the subjects into eight age groups in 10-year intervals to determine the dominant *M. tuberculosis* genotype for each age group. Mixed-genotype strains were excluded because of ambiguity between true mixed infection and DNA contamination during processing. The results revealed that all *M. tuberculosis* genotypes were distributed evenly within both patient groups (PTB and EPTB) but that the predominant strains within these populations differed between the age groups (Fig. 1). Among the 1,269 PTB patients, *M. tuberculosis* SCG1 was predominant in elderly patients, including patients aged 51 to 60 (46.4%), 61 to 70 (56.1%), 71 to 80 (53.7%), and >80 years (78.3%). Conversely, the most common *M. tuberculosis* genotype from samples collected from younger patients, including patients aged <10 years (75.0%), 11 to 20 (65.1%), 21 to 30 (57.1%), 31 to 40 (43.9%), and 41 to 50 years (42.7%), was SCG2. In addition, *M. tuberculosis* SCG5 was significantly more common in young patients than in elderly patients. Although there were fewer cases of EPTB (n = 132), the prevalence still clearly displayed an infection trend of *M. tuberculosis* SCG1 in elderly patients and SCG2 in younger patient groups.

There is some evidence of disruption of sympatric host-pathogen coevolution by HIV infection (30). The dynamic distribution of *M. tuberculosis* genotypes among different age groups of the HIV-negative PTB patients indicated the presence of clonal selection among patient age groups (Fig. 1b). HIV infection has fre-

TABLE 2 Diversity of SNP genotyping patterns between *M. tuberculosis* strains causing PTB and EPTB

SCG	ST	No. of PTB cases (%)	No. of EPTB cases (%)	SNP profile[a]
1	15	506 (39.9)	47 (35.6)	GGGCTGCCTTCCCTCCGACGTCGGAAGATTCAGGGCCTGCCCGGG
	Th1	1 (0.1)	0 (0.0)	...T.....
	Th2	3 (0.2)	0 (0.0)	.....G.....
2	3	20 (1.6)	0 (0.0)	.A....T.....GG.....A.....T.T.A
	8	4 (0.3)	2 (1.5)	.A....TT.....T.G..C...G.AGG.....A.....TTT.A
	10	179 (14.1)	42 (31.8)	.A....TT.....T....C.....GG.....A.....TTT.A
	11	1 (0.1)	0 (0.0)	.A.....G.....A.....T.T.A
	19	176 (13.9)	9 (6.8)	.A....TT.....C.....GG.....A.....TTT.A
	22	70 (5.5)	14 (10.6)	.A....TT.....T....C...G.AGG.....A.....TTT.A
	25	39 (3.1)	0 (0.0)	.A....T.....C.....GG.....A.....TTT.A
	26	10 (0.8)	2 (1.5)	.A.....T.....G.....A.....T.T.A
	F	0 (0.0)	1 (0.8)	.A....T.....T....C.....GG.....A.....TTT.A
	K	50 (3.9)	1 (0.8)	.A....TT.....GG.....A.....T.T.A
	Th3	1 (0.1)	0 (0.0)	.A....TT.....GG.....A.....TTT.A
	Th4	1 (0.1)	0 (0.0)	.A....TT....C.....C.....GG.....A.....TTT.A
	Th5	2 (0.2)	0 (0.0)	.A....TT.....T....C.....AGG.....A.....TTT.A
3a	21	22 (1.7)	0 (0.0)	.A.....G.....A.....T.T..
3b	20	12 (0.9)	2 (1.5)	.A.....A.....T....
	34	7 (0.6)	2 (1.5)	.A.T.....A.....A..CA.T....
3c	Th6	1 (0.1)	0 (0.0)	.A.T....CA.....TA.....A..CA.T....
4	9	1 (0.1)	0 (0.0)	.A.T....CA.....A.TA.....A..CA.T....
5	7	47 (3.7)	3 (2.3)	.A.....ATG...T....
	24	103 (8.1)	6 (4.5)	.A.....G.....ATG...T....
	41	1 (0.1)	0 (0.0)	.A.....C.....G.....ATG...T....
	Th7	1 (0.1)	0 (0.0)	.A.....T.....ATG...T....
	Th8	1 (0.1)	0 (0.0)	.A.....G.G.....ATG...T....
	Th9	3 (0.2)	0 (0.0)	.A.....G.....ATG.....
6a	18	7 (0.6)	1 (0.8)	CAT.....G.....A.G..TT..A.

[a] The dot indicates that the nucleotide is identical to that in the reference sequence.

quently been reported to be a risk factor for manifestation of EPTB (31). Our data strongly agreed with these previous findings, with 36/39 EPTB patients being HIV positive (Table 1). Unfortunately, the bacterial genotype data across different age groups of HIV-negative EPTB patients could not be determined in this study (Fig. 1b).

Frequency distribution of circulating *M. tuberculosis* strains in Thailand. The most prevalent genotypes were SCG1, -2, and -5 among PTB cases (Table 3). While the rare *M. tuberculosis* genotypes SCG3a, -3b, -3c, -4, and -6a were found among isolates examined in the current study, they were statistically analyzed as "other" genotypes for the main trait. The frequency of distribution of bacterial genotypes of strains isolated from different age groups was significantly different between these populations ($P < 0.01$). Interestingly, a decrease in the prevalence of SCG1 was significantly associated with patients who were born in the post-BCG vaccination environment. On the other hand, statistical analysis revealed an increasing number of patients harboring *M. tuberculosis* SCG2 and -5 in the postvaccination population.

We attempted to reduce the possible risk factors of human hosts by adjusting the results of *M. tuberculosis* genotyping by the selective force of BCG vaccination for adult PTB patients who were 15 to 65 years old and HIV negative. Within the PTB patient group, there were changing dynamics in the prevalence of *M. tuberculosis* genotypes SCG1, -2, and -5 between the two groups (pre-BCG vaccine, $n = 181$ isolates; and post-BCG vaccine, $n = 662$ isolates) (Table 3). There was a negative correlation between patients who were born in the post-BCG vaccine era and *M. tuberculosis* SCG1 though there was a positive association with SCG2 and SCG5.

DISCUSSION

Our results provide up-to-date information on the genotypes of circulating *M. tuberculosis* strains after the first introduction of the BCG vaccine in Thailand in 1953. Several previous studies indicated the possible association between genotype and clinical expression; however, the small sample sizes of these studies meant that the results were unreliable (12, 32, 33). In this study, SNP-based genetic variation of circulating *M. tuberculosis* isolates was investigated by examining a large number of *M. tuberculosis* isolates from patients with PTB and EPTB in Thailand. All isolates were collected from individual patients, and probable outbreak cases were ruled out. A previous study on TB meningitis by our group provided data similar to the data from the current study in

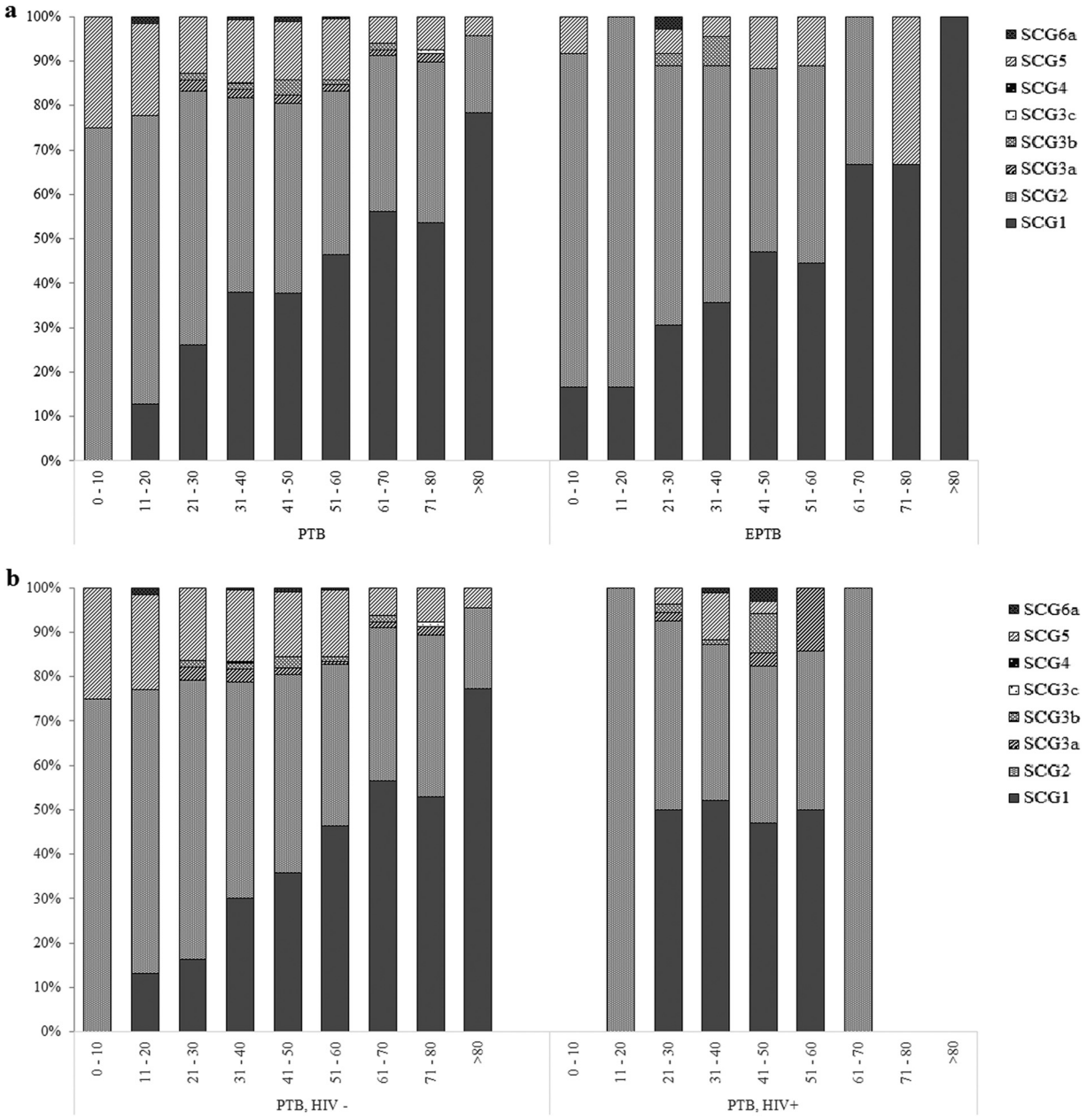

FIG 1 Trend of *M. tuberculosis* genotypes for the different patient age groups. The SCG genotypes in the PTB and EPTB groups (a) and HIV-negative and -positive PTB groups (b) are shown as percentages. Age groups (years) are indicated along the *x* axes.

regard to the most prevalent *M. tuberculosis* genotypes (BJ strains, 56%; EAI, 31%; T, 10%; and H, 3%) (6, 11). Even with the addition of the PTB cases, both the BJ and EAI genotypes, which are classified as SCG2 and SCG1, respectively, were still found to be the endemic genotypes causing TB in Thailand (6, 11, 12, 34–36). The prevalences of each genotype did not differ significantly between the PTB and EPTB cases in our study. Studies from Madagascar (37) and Brazil (38) using both PTB and EPTB cases also implied that host or environmental factors, rather than bacterial lineages, have an impact on disease phenotypes. In our study, there was no significant difference in the association between the genotypes of *M. tuberculosis* strains and the locations from which the strains were collected. However, the high prevalence of the EAI genotype in PTB patients (48.4%, $n = 150$) might be caused by the migratory habits of people living in Chiang Rai Province, where there is a border between Thailand and Myanmar (16). The presence of the largest cluster, *M. tuberculosis* ST15 belonging to SCG1, could be caused by the low discriminatory power of the selected markers or by special characteristics of this ST when it is the cause of PTB in Thailand. As described previously by spoligotyping analysis, SCG1 strains are subdivided into many different genotypes (14). To better classify SCG1 strains, other highly informative SNPs or markers should be developed. In addition, in this study, it was not clear if the mixed *M. tuberculosis* infections evident from DNA samples were caused by contamination, mixed infection, or coevolution in the host. Our results based on the

TABLE 3 Prevalence comparison of SCGs between *M. tuberculosis* strains causing PTB

SCG	Prevalence by patient age and group (no. of patients [%])[a]					
	PTB patients (n = 1,269)			HIV-negative adult PTB patients (n = 843)		
	Born after 1953 (n = 873)	Born before 1953 (n = 396)	Odds ratio (95% CI)	Born after 1953 (n = 662)	Born before 1953 (n = 181)	Odds ratio (95% CI)
1	292 (33.4)	218 (55.1)	0.4 (0.3–0.5)	191 (28.9)	89 (49.2)	0.4 (0.3–0.6)
2	415 (47.5)	138 (34.8)	1.7 (1.3–2.2)	331 (50.0)	70 (38.7)	1.6 (1.1–2.2)
5	128 (14.7)	28 (7.1)	2.3 (1.5–3.5)	112 (16.9)	16 (8.8)	2.1 (1.2–3.6)
Other[b]	38 (4.4)	12 (3.0)	1.5 (0.8–2.8)	28 (4.2)	6 (3.3)	1.3 (0.5–3.2)

[a] Prevalence was determined by the number of patients (%) positive for the SCG.
[b] Other, SCGs 3a, 3b, 3c, 4, and 6a.

developed DNA chip revealed evidence of mixed DNA in an individual sample by displaying heterogeneous results.

The predominant genotype in the current populations was SCG2, followed by SCG1 and SCG5, which frequently display higher transmissibility and virulence than other genotypes (22, 29). In other high-BJ-burden countries, such as China, a high prevalence of modern BJ strains has also been reported; however, in other countries, such as Japan and South Korea, more strains from ancestral BJ genotypes (ST3 and ST19 in Japan and ST11 or ST26 in South Korea) have been found. Modern BJ genotypes (ST10 and ST22) are also predominant in Taiwan and Peru (17, 18, 29). The close historical relationship between China (39) and Japan (17) might be responsible for the BJ sublineages in Thailand. The most frequently identified strains in Taiwan were similar to predominant BJ subtypes in our populations, including ST10 (35.42%), ST19 (29.65%), and ST22 (13.46%) (40). Advanced means of global travel could affect the spread of infectious strains among countries, and bacterial adaptation raises concern over transnational transmission. BCG mass vaccination and the high prevalence of BJ genotypes in Southeast Asia support the hypothesis of BCG vaccine subtype selection being responsible for limiting the dynamic distribution of BJ genotypes (41).

Differences in the distribution of *M. tuberculosis* genotypes across patient generations were observed in our study. Using different molecular markers and population age groupings, previous observations from countries where vaccination is required, including Thailand (12, 35), Vietnam (20, 21, 42), Taiwan (19), and China (15), have also revealed a higher frequency of *M. tuberculosis* genotype SCG2 strains causing TB in younger patients than in elderly patients. The association trend between EAI lineage strains (SCG1) and older patient age has also been reported in neighboring countries, including Vietnam (42). These results imply that BJ strains take the place of the EAI genotype in the current population. The differences in the percentages of particular bacterial genotypes across patient generations make it difficult to specify the roles of adaptation factors in this distribution. It is becoming increasingly accepted that the differential protein expression of bacterial genotypes could contribute to heterogeneous immune responses. The unique protein expression profile of the BJ genotype that affects immune suppression or modulation of the host may result in differences in BCG vaccine protection ability (23). Age of the patient and level of immunity at the time of bacterial reactivation and the degree of transmissibility could be emphasized. Our study was inconclusive in regard to the periods of infected individuals. Age at onset of disease was also not significantly associated with the duration of TB infection. The impacts of genotype-to-genotype variation, host immune response, and host susceptibility on the onset of the disease are largely unknown. However, based on our findings in PTB patients, it can be presumed that HIV coinfection probably interferes with this sympatric relationship (Fig. 1b). Further studies using a larger sample size and including data on host social behavior and demographic factors are needed to confirm the results of this pilot study.

According to several human-mycobacterial coevolution studies, there is evidence of the vaccination having an impact on the selection of *M. tuberculosis* genotypes (22). Most of the elderly patients included in the current study were born prior to the introduction of the BCG vaccine in Thailand in 1953. The particular *M. tuberculosis* genotypes of strains isolated from the younger patients were dynamic in terms of dominance. We hypothesized that BCG vaccine distribution accidentally introduced selective pressure into the *M. tuberculosis* population in Thailand. We identified an association between *M. tuberculosis* SCG and the clinical phenotype in all patients, regardless of whether they had exposure to the prevaccine environment. The results of the current study highlight the genetic diversity of cultured *M. tuberculosis* strains in Thailand. The data revealed that the majority of isolates were classified as genotypes SCG2 and SCG1, belonging to the BJ and EAI genotypes. An increase in the prevalence of *M. tuberculosis* SCG2 and SCG5 isolates was observed in patients who were born post-BCG vaccine implementation. These results indicate an ongoing transmission of these genotypes in the younger generation. Conversely, SCG1 genotype strains tended to cause disease in patients who were born prior to 1953. The presence of transmission dynamics among patients who were born during the BCG vaccination era indicates that mass BCG vaccination might be one of the selective factors that have driven the spread of the *M. tuberculosis* BJ genotype in Thailand. These epidemiological data on the currently circulating strains of *M. tuberculosis* could assist in the selection of appropriate vaccine strains in the future. Whether these strains are associated with BCG vaccination requires further study. In addition, further extended studies on successful BJ strains in vaccination countries are needed.

ACKNOWLEDGMENTS

We thank the Drug Resistant Tuberculosis Fund Laboratory, the Siriraj Foundation, under the patronage of Her Royal Highness Princess Galyanivadhana KromLaung Narathiwas Rajnakarindth, and the Department of Microbiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, for the laboratory facilities and patient data for EPTB cases. The Ministry of Public Health and the TB/HIV Research Project, RIT, Chiang Rai, are acknowledged for providing clinical data and DNA samples. We

sincerely thank the Department of Human Genetics of the Graduate School of Medicine at the University of Tokyo for their contribution to this study.

This study was supported by grants from a collaborative research project of the Japan Science and Technology Agency (JST) and National Science and Technology Development Agency (NSTDA) (grant no. P-12-01777), the Mahidol University Research Fund, Siriraj Graduate Thesis Scholarship, and the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program (grant no. PHD/0315/2551) to P.S. (student) and A.C. (advisor). A.C. was also supported by a Chalermprakiat grant, Faculty of Medicine, Siriraj Hospital, Mahidol University. Sample collections in Chiang Rai were supported by various sources including funding from Ministry of Health, Labor and Welfare of Japan.

We declare that we have no conflicts of interest.

REFERENCES

1. **Bonah C.** 2005. The "experimental stable" of the BCG vaccine: safety, efficacy, proof, and standards, 1921–1933. Stud. Hist. Philos. Biol. Biomed. Sci. **36:**696–721. http://dx.doi.org/10.1016/j.shpsc.2005.09.003.
2. **Brewer TF.** 2000. Preventing tuberculosis with bacillus Calmette-Guerin vaccine: a meta-analysis of the literature. Clin. Infect. Dis. **31**(Suppl 3)**:** S64–S67. http://dx.doi.org/10.1086/314072.
3. **Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, Parkhill J, Malla B, Berg S, Thwaites G, Yeboah-Manu D, Bothamley G, Mei J, Wei L, Bentley S, Harris SR, Niemann S, Diel R, Aseffa A, Gao Q, Young D, Gagneux S.** 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. Nat. Genet. **45:**1176–1182. http://dx.doi.org/10.1038/ng.2744.
4. **Caws M, Thwaites G, Dunstan S, Hawn TR, Lan NT, Thuong NT, Stepniewska K, Huyen MN, Bang ND, Loc TH, Gagneux S, van Soolingen D, Kremer K, van der Sande M, Small P, Anh PT, Chinh NT, Quy HT, Duyen NT, Tho DQ, Hieu NT, Torok E, Hien TT, Dung NH, Nhu NT, Duy PM, van Vinh Chau N, Farrar J.** 2008. The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*. PLoS Pathog. **4:**e1000034. http://dx.doi.org/10.1371/journal.ppat.1000034.
5. **Drobniewski F, Balabanova Y, Nikolayevsky V, Ruddy M, Kuznetzov S, Zakharova S, Melentyev A, Fedorin I.** 2005. Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia. JAMA **293:**2726–2731. http://dx.doi.org/10.1001/jama.293.22.2726.
6. **Faksri K, Drobniewski F, Nikolayevskyy V, Brown T, Prammananan T, Palittapongarnpim P, Prayoonwiwat N, Chaiprasert A.** 2011. Epidemiological trends and clinical comparisons of *Mycobacterium tuberculosis* lineages in Thai TB meningitis. Tuberculosis (Edinb.) **91:**594–600. http://dx.doi.org/10.1016/j.tube.2011.08.005.
7. **Batieha A, Abo-Khadhra K, Rawashdah M, Jaddou H, Sheyyab M, Abo-Romman K, Daoud A.** 1998. The effectiveness of BCG vaccination: the Jordanian experience. J. Trop. Pediatr. **44:**288–290. http://dx.doi.org/10.1093/tropej/44.5.288.
8. **Abebe F, Bjune G.** 2006. The emergence of Beijing family genotypes of *Mycobacterium tuberculosis* and low-level protection by bacille Calmette-Guerin (BCG) vaccines: is there a link? Clin. Exp. Immunol. **145:**389–397. http://dx.doi.org/10.1111/j.1365-2249.2006.03162.x.
9. **Zwerling A, Behr MA, Verma A, Brewer TF, Menzies D, Pai M.** 2011. The BCG World Atlas: a database of global BCG vaccination policies and practices. PLoS Med. **8:**e1001012. http://dx.doi.org/10.1371/journal.pmed.1001012.
10. **Payanandana V, Kladhuang B, Somsong W, Jittimanee S.** 1999. Battle against TB national tuberculosis programme Thailand, 1999. Ministry of Public Health, Bangkok, Thailand.
11. **Yorsangsukkamol J, Chaiprasert A, Prammananan T, Palittapongarnpim P, Limsoontarakul S, Prayoonwiwat N.** 2009. Molecular analysis of *Mycobacterium tuberculosis* from tuberculous meningitis patients in Thailand. Tuberculosis (Edinb.) **89:**304–309. http://dx.doi.org/10.1016/j.tube.2009.05.001.
12. **Rienthong D, Ajawatanawong P, Rienthong S, Smithtikarn S, Akarasewi P, Chaiprasert A, Palittapongarnpim P.** 2005. Restriction fragment length polymorphism study of nationwide samples of *Mycobacterium tuberculosis* in Thailand, 1997–1998. Int. J. Tuberc. Lung Dis. **9:**576–581.
13. **Srilohasin P, Chaiprasert A, Tokunaga K, Nao N, Prammananan T.** 2014. Novel DNA chip based on modified DigiTag2 assay for high-throughput species identification and genotyping of *Mycobacterium tuberculosis* complex isolates. J. Clin. Microbiol. **52:**1962–1968. http://dx.doi.org/10.1128/JCM.00153-14.
14. **Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, Bobadilla del Valle M, Fyfe J, Garcia-Garcia L, Rastogi N, Sola C, Zozio T, Guerrero MI, Leon CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendon A, Sifuentes-Osornio J, Ponce de Leon A, Cave MD, Fleischmann R, Whittam TS, Alland D.** 2006. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J. Bacteriol. **188:**759–772. http://dx.doi.org/10.1128/JB.188.2.759-772.2006.
15. **Pang Y, Song Y, Xia H, Zhou Y, Zhao B, Zhao Y.** 2012. Risk factors and clinical phenotypes of Beijing genotype strains in tuberculosis patients in China. BMC Infect. Dis. **12:**354. http://dx.doi.org/10.1186/1471-2334-12-354.
16. **Phyu S, Stavrum R, Lwin T, Svendsen OS, Ti T, Grewal HM.** 2009. Predominance of *Mycobacterium tuberculosis* EAI and Beijing lineages in Yangon, Myanmar. J. Clin. Microbiol. **47:**335–344. http://dx.doi.org/10.1128/JCM.01812-08.
17. **Iwamoto T, Fujiyama R, Yoshida S, Wada T, Shirai C, Kawakami Y.** 2009. Population structure dynamics of *Mycobacterium tuberculosis* Beijing strains during past decades in Japan. J. Clin. Microbiol. **47:**3340–3343. http://dx.doi.org/10.1128/JCM.01061-09.
18. **Kang HY, Wada T, Iwamoto T, Maeda S, Murase Y, Kato S, Kim HJ, Park YK.** 2010. Phylogeographical particularity of the *Mycobacterium tuberculosis* Beijing family in South Korea based on international comparison with surrounding countries. J. Med. Microbiol. **59:**1191–1197. http://dx.doi.org/10.1099/jmm.0.022103-0.
19. **Dou HY, Tseng FC, Lin CW, Chang JR, Sun JR, Tsai WS, Lee SY, Su IJ, Lu JJ.** 2008. Molecular epidemiology and evolutionary genetics of *Mycobacterium tuberculosis* in Taipei. BMC Infect. Dis. **8:**170. http://dx.doi.org/10.1186/1471-2334-8-170.
20. **Buu TN, Huyen MN, Lan NT, Quy HT, Hen NV, Zignol M, Borgdorff MW, Cobelens FG, van Soolingen D.** 2009. The Beijing genotype is associated with young age and multidrug-resistant tuberculosis in rural Vietnam. Int. J. Tuberc. Lung Dis. **13:**900–906.
21. **Anh DD, Borgdorff MW, Van LN, Lan NT, van Gorkom T, Kremer K, van Soolingen D.** 2000. *Mycobacterium tuberculosis* Beijing genotype emerging in Vietnam. Emerg. Infect. Dis. **6:**302–305. http://dx.doi.org/10.3201/eid0603.000312.
22. **Parwati I, van Crevel R, van Soolingen D.** 2010. Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains. Lancet Infect. Dis. **10:**103–111. http://dx.doi.org/10.1016/S1473-3099(09)70330-5.
23. **Pheiffer C, Betts JC, Flynn HR, Lukey PT, van Helden P.** 2005. Protein expression by a Beijing strain differs from that of another clinical isolate and *Mycobacterium tuberculosis* H37Rv. Microbiology **151:**1139–1150. http://dx.doi.org/10.1099/mic.0.27518-0.
24. **Hermans PW, Messadi F, Guebrexabher H, van Soolingen D, de Haas PE, Heersma H, de Neeling H, Ayoub A, Portaels F, Frommel D, Zribi M, van Embden JD.** 1995. Analysis of the population structure of *Mycobacterium tuberculosis* in Ethiopia, Tunisia, and The Netherlands: usefulness of DNA typing for global tuberculosis epidemiology. J. Infect. Dis. **171:**1504–1513. http://dx.doi.org/10.1093/infdis/171.6.1504.
25. **Ford C, Yusim K, Ioerger T, Feng S, Chase M, Greene M, Korber B, Fortune S.** 2012. *Mycobacterium tuberculosis*—heterogeneity revealed through whole genome sequencing. Tuberculosis (Edinb.) **92:**194–201. http://dx.doi.org/10.1016/j.tube.2011.11.003.
26. **Faksri K, Drobniewski F, Nikolayevskyy V, Brown T, Prammananan T, Palittapongarnpim P, Prayoonwiwat N, Chaiprasert A.** 2011. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on IS*6110*, SNP, LSP and VNTR profiles from Thailand. Infect. Genet. Evol. **11:**1142–1149. http://dx.doi.org/10.1016/j.meegid.2011.04.007.
27. **Iwamoto T, Yoshida S, Suzuki K, Wada T.** 2008. Population structure analysis of the *Mycobacterium tuberculosis* Beijing family indicates an association between certain sublineages and multidrug resistance. Antimicrob. Agents Chemother. **52:**3805–3809. http://dx.doi.org/10.1128/AAC.00579-08.
28. **Hill EB, Wayne LG, Gross WM.** 1972. Purification of mycobacterial deoxyribonucleic acid. J. Bacteriol. **112:**1033–1039.

29. **Iwamoto T, Grandjean L, Arikawa K, Nakanishi N, Caviedes L, Coronel J, Sheen P, Wada T, Taype CA, Shaw MA, Moore DA, Gilman RH.** 2012. Genetic diversity and transmission characteristics of Beijing family strains of *Mycobacterium tuberculosis* in Peru. PLoS One **7:**e49651. http://dx.doi.org/10.1371/journal.pone.0049651.
30. **Fenner L, Egger M, Bodmer T, Furrer H, Ballif M, Battegay M, Helbling P, Fehr J, Gsponer T, Rieder HL, Zwahlen M, Hoffmann M, Bernasconi E, Cavassini M, Calmy A, Dolina M, Frei R, Janssens JP, Borrell S, Stucki D, Schrenzel J, Bottger EC, Gagneux S, Swiss HIV Cohort and Molecular Epidemiology of Tuberculosis Study Groups.** 2013. HIV infection disrupts the sympatric host-pathogen relationship in human tuberculosis. PLoS Genet. **9:**e1003318. http://dx.doi.org/10.1371/journal.pgen.1003318.
31. **Yang Z, Kong Y, Wilson F, Foxman B, Fowler AH, Marrs CF, Cave MD, Bates JH.** 2004. Identification of risk factors for extrapulmonary tuberculosis. Clin. Infect. Dis. **38:**199–205. http://dx.doi.org/10.1086/380644.
32. **Palittapongarnpim P, Rienthong S, Panbangred W.** 1993. Comparison of restriction fragment length polymorphism of *Mycobacterium tuberculosis* isolated from cerebrospinal fluid and sputum: a preliminary report. Tuber. Lung Dis. **74:**204–207. http://dx.doi.org/10.1016/0962-8479(93)90013-N.
33. **Lari N, Rindi L, Cristofani R, Rastogi N, Tortoli E, Garzelli C.** 2009. Association of *Mycobacterium tuberculosis* complex isolates of BOVIS and Central Asian (CAS) genotypic lineages with extrapulmonary disease. Clin. Microbiol. Infect. **15:**538–543. http://dx.doi.org/10.1111/j.1469-0691.2009.02712.x.
34. **Cheunoy W, Haile M, Chaiprasert A, Prammananan T, Cristea-FernstrÖM M, Vondracek M, Chryssanthou E, Hoffner S, Petrini B.** 2009. Drug resistance and genotypic analysis of *Mycobacterium tuberculosis* strains from Thai tuberculosis patients. APMIS **117:**286–290. http://dx.doi.org/10.1111/j.1600-0463.2009.02438.x.
35. **Prodinger WM, Bunyaratvej P, Prachaktam R, Pavlic M.** 2001. *Mycobacterium tuberculosis* isolates of Beijing genotype in Thailand. Emerg. Infect. Dis. **7:**483–484. http://dx.doi.org/10.3201/eid0703.017330.
36. **Palittapongarnpim P, Luangsook P, Tansuphaswadikul S, Chuchottaworn C, Prachaktam R, Sathapatayavongs B.** 1997. Restriction fragment length polymorphism study of *Mycobacterium tuberculosis* in Thailand using IS*6110* as probe. Int. J. Tuberc. Lung Dis. **1:**370–376.
37. **Rasolofo Razanamparany V, Menard D, Auregan G, Gicquel B, Chanteau S.** 2002. Extrapulmonary and pulmonary tuberculosis in Antananarivo (Madagascar): high clustering rate in female patients. J. Clin. Microbiol. **40:**3964–3969. http://dx.doi.org/10.1128/JCM.40.11.3964-3969.2002.
38. **Gomes T, Vinhas SA, Reis-Santos B, Palaci M, Peres RL, Aguiar PP, Ribeiro FK, Marques HS, Dettoni Vdo V, Johnson JL, Riley LW, Maciel EL.** 2013. Extrapulmonary tuberculosis: *Mycobacterium tuberculosis* strains and host risk factors in a large urban setting in Brazil. PLoS One **8:**e74517. http://dx.doi.org/10.1371/journal.pone.0074517.
39. **Lauro D.** 1988. Chinese migration and settlement in a Thai village. Int. Migr. **26:**301–310. http://dx.doi.org/10.1111/j.1468-2435.1988.tb00651.x.
40. **Chen YY, Chang JR, Huang WF, Kuo SC, Su IJ, Sun JR, Chiueh TS, Huang TS, Chen YS, Dou HY.** 2012. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on SNP and VNTR typing profiles in Asian countries. PLoS One **7:**e39792. http://dx.doi.org/10.1371/journal.pone.0039792.
41. **van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, Qing HZ, Enkhsaikan D, Nymadawa P, van Embden JD.** 1995. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia. J. Clin. Microbiol. **33:**3234–3238.
42. **Nguyen VA, Choisy M, Nguyen DH, Tran TH, Pham KL, Thi Dinh PT, Philippe J, Nguyen TS, Ho ML, Van Tran S, Banuls AL, Dang DA.** 2012. High prevalence of Beijing and EAI4-VNM genotypes among *M. tuberculosis* isolates in northern Vietnam: sampling effect, rural and urban disparities. PLoS One **7:**e45553. http://dx.doi.org/10.1371/journal.pone.0045553.