小された。

移民を受け入れている国の多くでは，スクリーニングが制度化されている。実施時期は入国前・入国時・入国後，検査法として胸部X線検査及びツ反またはIGRA，対象の選定方法は出身国の罹患率あるいは長期滞在者など，国・地域によって様々な方法が取られている。発見率は国レベルのスクリーニングでは0.05%から0.22%，systematic reviewの中央値で0.18%，メタアナリシスで0.27%等であった。入国前スクリーニングの長所は，①受入国の費用負担がないこと，②受診率が100%であること，③罹患率減少に直接的効果が証明されていることであるが，健診の精度管理が必要である。

接触者健診は患者発見方法として重要性が高いが，地域・施設によって実施状況に差があり，改善が必要である。

医療機関における患者発見は2013年の新登録患者の83%を占めており，早期発見のために最も重要である。2000年以降サーベイランスにおける発見の遅れは少なくなる傾向にあるが，集団感染事件は減少していないことから，結核に接する機会の少ない一般の医療機関等での診断の遅れは著しくなっている可能性がある。活動性結核及びLTBI治療終了後の再発・発病が一定数発生することは避けられない。LTBI治療後の発病については発病リスクが低い者も治療対象になっているために全体として低くなるが，3%程度の発病リスクを持つ者が含まれている可能性がある

4. インターフェロンγ遊離試験陽性率に関する研究

依頼状を送付した490保健所中，回答があったのは236施設(48.2%)であった。該当事例なしが32施設(13.6%)あり，204保健所(86.4%)から735事例のデータを収集した。不適格と判断されたもの84例を除いた分析対象事例数は651例であった。

一般人口のQFT-3G陽性率は40歳代以上では年齢と共に上昇し，推定既感染率と比較すると概ね3-4割程度であったが，80歳以上ではIGRA陽性率は70歳代以上に上昇しなかった。T-SPOTの陽性率はQFT-3Gと大きな違いはなかったが，80歳代でも比較的高い陽性率を示した。

D. 考察

1. 施設内感染対策に関する研究

近年の院内(施設内)における集団感染は，高齢者施設や精神病院（認知症等の高齢者が多い）での発生が目立っている。リスクのあるにも関わらず，結核に対する関心が低いために患者発見の大幅な遅れになり集団感染発生に至っていると思われる。研修会への参加者は，結核対策に関心もっていて集団感染事件を起こす可能性が低い施設が多いものと推定されることから，本手引きの活用にあたって，周知の方法には工夫が必要である。

2. 結核感染・罹患時の基礎疾患への影響について

原疾患の増悪を予防し患者の不利益とならないよう，結核を専門とする医師も免疫低下を来す疾患を診療する医療機関の医師と薬剤の相互作用や副作用対策，治療期間の決定など，治療計画に広く協力していく必要がある。

回答した多くの医療機関は高度な医療水準が保たれているため，感染症および院内感染対策に配慮した良好な治療環境が維持されたと推測する。潜在性結核感染症治療の考え方については，さらなる情報発信が必須である。

3. 低まん延化に向けての患者発見対策に関する研究

スクリーニングとして定期健診，外国出生者健診，接触者健診，治療終了後の経過観察を患者発見方法として整合性に配慮しながら，強化と効率化の観点から整理する必要がある。

定期健診については対象者の絞り込みと受診率の確保等の強化が必要である。外国出生者に対するスクリーニングの対象・実施時期及び方法を検討の上，早期に実施が望まれる。接触者健診については実施状況の評価法の開発とそれに基づく技術支援が必要と考えられる。医療機関における患者発見のためには情報提供，特に一般の医療機関や国民の情報伝達方法の検討が必要と思われる。治療終了後の経過観察については再発・発病及び患者発見の観点から対象の選定・実施方法の検討が必要と考えられる。

4. インターフェロンγ遊離試験陽性率に関する研究

今回の結果を低まん延地域における接触者健康診断における全対象者のQFT-3G陽性率と比較すると，20歳代，30歳代で若干低い数値になっているが，本研究では当該接触者健診で感染を受けた人を除外しているためと思われた。40歳代以上では大きな違いがなかったのは，今回は全国の一般人口を対象にしているため既感染者が含まれていることと感染者が除外されていることが相殺されて，同様の数値になった可能性がある。ただし，本調査は感染リスクが低い対象とはいえ，接触者健診のデータを使っていることから，感染を受けた者が混入している可能性は否定できない。

2010年の推定既感染率との比較では，高年齢層のIGRA陽性率は推定既感染率の概ね3-4割程度であった。

本調査結果は高齢者が多く含まれる接触者健診において，集団感染の有無の判断に有用なデータとなることが期待される。

E. 結論

低まん延実現に向けての課題となっている，院内感染対策，免疫低下状態の患者への対応の調査，患者発見対策，インターフェロンγ遊離試験に関する研究を実施した。成果としての「結核院内（施設内）感染対策の手引き」はホームページでの公開，刊行物，研修会等の資料として広く活用されている。免疫低下状態の患者の診療を担っている医療施設での結核患者の対応の課題が明らかになった。患者発見の課題・論点の整理は予防指針の改定時に役立てられることを期待したい。IGRAの陽性率は高齢者が多く含まれる接触者健診において，集団感染の有無の判断等に活用されるものと思われる。

以上のように本研究は対策の強化に直接的に裨益する成果が得られたと考えられる。

F. 健康危険情報

なし

G. 研究発表

3. 書籍

1) 結核院内（施設内）感染対策の手引き（平成26年版）：結核研究所ホームページ（http://www.jata.or.jp/tp_detail.php?id=39）で公開

2) 加藤誠也編．結核院内（施設内）感染対策の手引き　実際に役立つQ&A．公益財団法人結核予防会．東京，2014年

H. 知的財産権の出願・登録状況

なし

Ⅲ　研究成果の刊行に関する一覧表

別紙４

書籍

書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年
森亨,加藤誠也,御手洗聡	結核院内（施設内）感染対策の手引き　実際に役立つQ&A	公財）結核予防会	東京	2014

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Iwamoto T, Grandjean L, Arikawa K, Nakanishi N, Caviedes L, Coronel J, Sheen P, Wada T, Taype CA, Shaw MA, Moore D.A.J, and Robert H. Gilman R.H.	Genetic diversity and transmission characteristics of Beijing family strains of *Mycobacterium tuberculosis* in Peru.	PLoS ONE	7	e49651	2012
Nakanishi N, Wada T, Arikawa K, Millet J, Rastogi N, and Iwamoto T.	Evolutionary robust SNPs reveal the misclassification of *Mycobacterium tuberculosis* Beijing family strains into sublineages.	Infect. Genet. Evol.	16	174-177	2013
Nakajima C, Tamaru A, Rahim Z, Poudel A, Maharjian B, Aye KS, Ling H, Hattori T, Iwamoto T, Fukushima Y, Suzuku H, Suzuki Y and Matsuba T.	Simple multiplex PCR assay for identification of Beijing family *Mycobacterium tuberculosis* isolates with a lineage-specific mutation in Rv0679c.	J. Clin. Microbiol.	51	2025-2032	2013
Iwamoto T, Arikawa K, Nakajima C, Nakanishi N, Nichiuchi Y, Yoshida S, Tamaru A, Tamura Y, Hoshino H, Yo o H, Park YK, Saito H and Suzuki Y.	Intra-subspecies sequence variability of the MACPPE12 gene in *Mycobacterium avium* subsp. *hominissuis*.	Infect. Genet. Evol.	21	479-83	2014加

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Wada T, Maeda S.	Multiplex agarose gel electrophoresis system for variable number of tandem repeats genotyping: Analysis example using Mycobacterium tuberculosis.	Electrophoresis	34(8)	1171-1174	2013
Tateishi Y, Tamaru A, Ogura Y, Niki M, Wada T,Yamamoto T, Hirata K,Hayashi T, Matsumoto S.	Whole-Genome Sequence of the Potentially Hypertransmissible Multidrug-Resistant Mycobacterium tuberculosis Beijing Strain OM-V02_005.	Genome Announcement	1(4)	e00608-e00613.	2013
Yamada H, Bhatt A, Danev R, Fujiwara N, Maeda S, Mitarai S, Chikamatsu K, Aono A, Nitta K, Jacobs WR Jr, Nagayama K.	Non-acid-fastness in *Mycobacterium tuberculosis* Δ*kasB* mutant correlates with the cell envelope electron density.	Tuberculosis	92	351-357	2012
Mitarai S, Karinaga R, Yamada H, Mizuno K, Chikamatsu K, Aono A, Sugamoto T, Hatano T.	TRICORE, a novel bead-based specimen concentration method for the culturing of *Mycobacterium tuberculosis*.	Journal of Microbiological Methods	90	152-155	2012
Chikamatsu K, Aono A, Yamada H, Mitarai S.	Evaluation of Genotype MTBDRsl for testing resistance of Mycobacterium tuberculosis isolates to fluoroquinlone, aminoglycoside, and ethambutol.	Kekkaku	87	641-647	2012
青野昭男、近松絹代、山田博之、村田正太、結城 篤、三澤成毅、小栗豊子、御手洗 聡	抗酸菌塗抹検査外部精度評価の試み	日本臨床微生物学雑誌	22	279-283	2012
Yamada H, Chikamatsu K, Aono A, Mitarai S.	Pre-fixation of virulent *Mycobacterium tuberculosis* with glutaraldehyde preserves exquisite ultrastructure on transmission electron microscopy through cryofixation and freeze-substitution with osmium-acetone at ultralow temperature.	Journal of Microbiological Methods.	96	50-55	2014

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
小林典子	特集② 地域連携体制の整備に向けた取り組み その概要	保健師・看護師の結核展望	99、前期	39-41	2012
小林典子	日本の結核対策における保健看護活動 50 年の歩み	保健師・看護師の結核展望	100、後期	25-30	2012
Hoshino A, Hanada S, Yamada H, Mii S, Takahashi M, Mitarai S, Yamamoto K, Manome Y.	Mycobacterium tuberculosis escapes from the phagosomes of infected human osteoclasts reprograms osteoclast development via dysregulation of cytokines and chemokines.	Pathogens and disease	70	28-39	2014

Ⅳ　研究成果の刊行物・別刷(一部)

PLOS ONE

Genetic Diversity and Transmission Characteristics of Beijing Family Strains of *Mycobacterium tuberculosis* in Peru

Tomotada Iwamoto[1]*, Louis Grandjean[2,3,4], Kentaro Arikawa[1], Noriko Nakanishi[1], Luz Caviedes[4], Jorge Coronel[4], Patricia Sheen[4], Takayuki Wada[5], Carmen A. Taype[6], Marie-Anne Shaw[6], David A. J. Moore[3,4], Robert H. Gilman[4,7]

1 Department of Microbiology, Kobe Institute of Health, Kobe, Japan, **2** Wellcome Centre for Clinical Tropical Medicine, St. Mary's Campus, Imperial College, London, United Kingdom, **3** Clinical Research Department and TB Centre, London School of Hygiene and Tropical Medicine, London, United Kingdom, **4** Laboratorio de Investigación de Enfermedades Infecciosas, Universidad Peruana Cayetano Heredia, Lima, Peru, **5** Department of Microbiology, Osaka City Institute of Public Health and Environmental Sciences, Osaka, Japan, **6** Faculty of Biological Science, University of Leeds, Leeds, United Kingdom, **7** Department of International Health, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, Maryland, United States of America

Abstract

Beijing family strains of *Mycobacterium tuberculosis* have attracted worldwide attention because of their wide geographical distribution and global emergence. Peru, which has a historical relationship with East Asia, is considered to be a hotspot for Beijing family strains in South America. We aimed to unveil the genetic diversity and transmission characteristics of the Beijing strains in Peru. A total of 200 Beijing family strains were identified from 2140 *M. tuberculosis* isolates obtained in Lima, Peru, between December 2008 and January 2010. Of them, 198 strains were classified into sublineages, on the basis of 10 sets of single nucleotide polymorphisms (SNPs). They were also subjected to variable number tandem-repeat (VNTR) typing using an international standard set of 15 loci (15-MIRU-VNTR) plus 9 additional loci optimized for Beijing strains. An additional 70 Beijing family strains, isolated between 1999 and 2006 in Lima, were also analyzed in order to make a longitudinal comparison. The Beijing family was the third largest spoligotyping clade in Peru. Its population structure, by SNP typing, was characterized by a high frequency of Sequence Type 10 (ST10), which belongs to a modern subfamily of Beijing strains (178/198, 89.9%). Twelve strains belonged to the ancient subfamily (ST3 [n = 3], ST25 [n = 1], ST19 [n = 8]). Overall, the polymorphic information content for each of the 24 loci values was low. The 24 loci VNTR showed a high clustering rate (80.3%) and a high recent transmission index ($RTI_{n-1} = 0.707$). These strongly suggest the active and ongoing transmission of Beijing family strains in the survey area. Notably, 1 VNTR genotype was found to account for 43.9% of the strains. Comparisons with data from East Asia suggested the genotype emerged as a uniquely endemic clone in Peru. A longitudinal comparison revealed the genotype was present in Lima by 1999.

Citation: Iwamoto T, Grandjean L, Arikawa K, Nakanishi N, Caviedes L, et al. (2012) Genetic Diversity and Transmission Characteristics of Beijing Family Strains of *Mycobacterium tuberculosis* in Peru. PLoS ONE 7(11): e49651. doi:10.1371/journal.pone.0049651

Editor: Igor Mokrousov, St. Petersburg Pasteur Institute, Russian Federation

Received August 5, 2012; **Accepted** October 11, 2012; **Published** November 21, 2012

Funding: This work was supported by Health Science Research grants(H24-SHINKO-IPPAN-011) from Ministry of Health, Labour and Welfare of Japan, MEXT/JSPS KAKENHI Grant Number 24590845, R24 TW007988 NIH FIC, D43 TW006581 NIH FIC, RO1 HD059005 NIH NICHD, WT088559MA with the Wellcome Trust, Wellcome Trust grant 078067/Z/05, a research grant of the Institute for Fermentation, Osaka, Japan, and the US-Japan Cooperative Medical Science Program (TB and Leprosy panel). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: kx2t-iwmt@asahi-net.or.jp

Introduction

Strains of the Beijing family of *Mycobacterium tuberculosis* (*M. tuberculosis*), first described in 1995 [1], have attracted worldwide attention because of their wide geographical distribution and global emergence. These strains have also been shown to have an endemic prevalence in certain regions, including Asia, except for the Indian subcontinent; northern Eurasia; and South Africa [2,3,4,5,6,7]. The lineage has been reported to have caused major outbreaks worldwide, some involving drug-resistant variants [2,8,9]. These characteristics suggest that strains belonging to this family might have selective advantages (higher virulence or transmissibility) over other *M. tuberculosis* strains [10,11,12,13]. Therefore, a better understanding of the contribution of the Beijing family to the tuberculosis (TB) pandemic is vital to improve global TB control.

The Beijing family is reported to be phylogenetically divisible into 2 main subfamilies: the modern (typical) and ancient (atypical) subfamilies, [14,15]. The modern subfamily is highly prevalent in China, western Russia, South Africa, and Thailand [3,15,16,17], whereas the ancient subfamily, with a deleted region of difference (RD) RD181[-] (late ancient type), is endemic in Japan and the ancient subfamily with RD181[+] (early ancient type) is endemic in Korea [4,18,19,20,21,22,23,24]. Although the reasons for the phylogeographical differences remain elusive, the above trends can be used to assess the influence of the Beijing family strains from East Asian countries, where the prevalence is very high [2,3], on the prevalence of Beijing family strains in other regions. In fact, a

large number of Chinese and Japanese immigrants settled in Peru in 19th century. Korea also has a history of migration to Peru, but this migration has occurred more recently. This historical relationship with East Asia encouraged us to characterize the Beijing family strains in Peru.

Unlike in other South American countries, the proportion of Beijing family strains of *M. tuberculosis* in Peru is exceptionally high [25]. In Brazil, Colombia, Paraguay, Venezuela, Argentina, Chile, and Ecuador the prevalence of Beijing family strains was reported to be less than 1% [25,26,27,28,29]. However, the prevalence in Peru was 5.9% (11/185) in 1999 [25] and 9.3% (30/323) for samples obtained between 2004 and 2006 [30]. Ritacco et al. [25] speculated that the Beijing family strains were first introduced into Peru, and eventually into other South American countries, when Peru received a significant number of Chinese immigrants in the mid-19th century. This same study also showed considerable diversity in the insertion sequence IS*6110* restriction fragment length polymorphism (RFLP) patterns, supporting the concept of earlier introduction(s) of different ancestral strains during the past 150 years. In addition to the importation of Beijing family strains from Asia to Peru, there is also evidence that the Beijing strains were imported into Europe through a South American route, specifically through Peru [31,32]. Therefore, Peru can be considered as the South American country that has been most strongly affected by the introduction of Beijing family and is also most commonly associated with the spread of these strains to other South American and European countries.

An in-depth analysis of Beijing family strains in Peru may have a significant impact on the understanding of global epidemics involving the *M. tuberculosis* Beijing family strains. The aims of the current study were to unveil the genetic diversity and transmission characteristics of Beijing family strains of *M. tuberculosis* in Lima, Peru, and to elucidate the probable impact of past immigration from East Asian countries.

Materials and Methods

Ethics Statement

Prior to the start of the study, ethical approval was obtained from both Universidad Peruana Cayetano Heredia and Imperial College London and institutional approval was obtained from the Peruvian Ministry of Health. Samples for this study were anonymized.

Study Samples

The Beijing family strains used in this study were identified on the basis of deletion of the spacers 1–34 assessed by the spoligotyping assay [33]. In order to negate the inclusion of "Pseudo-Beijing strains" [34], all of the strains classified as the early ancient type of Beijing family (RD 181[+]) were subjected to the RD 105 analysis [35].

A total of 200 Beijing family strains were identified from 2140 *M. tuberculosis* isolates obtained through a population-level implementation of a new diagnostic test (MODS, Microscopic Observation Drug Susceptibility) [36,37], which was conducted in Callao and South Lima between December 2008 and January 2010. Strains from all culture-positive patients with respiratory symptoms in the study area were included and none of the samples were duplicated from a single patient. Of the 200 Beijing family strains, 2 were excluded from this study because of insufficient DNA samples. For the remaining 198 Beijing strains, detailed information is provided in Tables 1 and S1.

An additional 70 Beijing family strains isolated in Lima between 1999 and 2006 were also analyzed in order to make a longitudinal comparison. In detail, 26 were obtained from a previous study in North Lima between 2004 and 2006 [30] and 44 were obtained from 4 distinct areas in Lima (North, South, East, and Central Lima) through the hospital based studies between 1999 and 2004. The details of the strains are described in Tables 1 and S2.

VNTR Typing

Genotypic data for the 24 loci that comprised the international standard set of 15 loci of variable number of tandem repeat(s) of mycobacterial interspersed repetitive units (15-MIRU-VNTR) [38], and 9 additional loci (2074, 2372, 3155, 3336, 3232, 3820, 4120, QUB11a, QUB18) were analyzed. The 9 additional loci were selected as Beijing-type optimized loci because of their highly discriminatory values in different studies focused on Beijing family strains [4,24,39,40,41,42,43]. This 24-loci VNTR was called the 24_{Beijing}-VNTR. The polymerase chain reaction (PCR) primers and the number of repeats for each locus, based on the *M. tuberculosis* H37Rv strain, are described in Table S3; the PCR conditions were as described previously [4]. The amplicon samples were diluted 20-fold with ultrapure water and analyzed on an AB3500 genetic analyzer system (Applied Biosystems, Foster City, CA) at a constant room temperature of 25°C following the manufacturers' instruction, i.e., capillary temperature, 60°C; electrophoresis voltage, 8.5 kV; and separation time, 5800 s. A GeneScan 1200 LIZ Size Standard (Applied Biosystems) was used to provide internal size markers. Fragment sizes were measured using GeneMapper Ver. 4 (Applied Biosystems). The numbers of repeats at each locus were calculated using the offset values of the size, which correct differences in relative migration between the size standard and the amplicons depending on the locus. The reproducibility and accuracy of size calling and the size offsets were checked by including *M. tuberculosis* H37Rv and 1 Beijing family strain (reference for quality control) into every batch of the analysis (one 96-well plate was used as a batch). For the large alleles (specifically, for the one larger than 1000 bases) of a locus, we used stutter peak counting, as shown in Figure S1, to obtain unambiguous results with high reproducibility [4]. Moreover, we confirmed the reproducibility of our assay by blindly re-testing 22 selected samples (Table S1). The allelic diversity of each VNTR locus was evaluated using Nei's diversity index [44], i.e., the polymorphic information content (PIC) (Table S4). Genotypic discrimination of the 198 Beijing strains, based on 15-MIRU-VNTR and 24_{Beijing}-VNTR, were calculated using the Hunter-Gaston discriminatory index (HGDI) [45]. A recent transmission index (RTI_{n-1}) [46,47] was also calculated using the VNTR profiles. A minimum spanning tree (MST), based on VNTR types, was constructed using Bionumerics software (Bionumerics ver. 4.2; Applied Math., Sint-Martens-Latem, Belgium), as previously described [23].

Single Nucleotide Polymorphisms at 10 Loci

The sequence types (ST) were determined based on the 10 synonymous SNPs, which were sufficient to divide the Beijing strains obtained from the global population [48]. Each chromosomal position in the whole-genome sequence of H37Rv [49] was as follows: 797736, 909166, 1477596, 1548149, 1692069, 1892017, 2376135, 2532616, 2825581, and 4137829. The SNP at position 1477596 is the same as *ogt*12, which discriminates between the ancient and modern type Beijing strains [16]. The SNP at 2532616 is the same as *adhE2* (codon 124), which further discriminates the modern type [50]. SNPs at 797736 and 2825581

Table 1. Patient demographics for the Beijing family strains in this study.

Characteristics	No. (%) of isolates	
	Between 2008 and 2010	Between 1999 and 2006
Total	198 (100)	70 (100)
Sex		
Male	133 (67)	45 (64)
Female	63 (32)	25 (36)
Unknown	2 (1)	0 (0)
Age group		
<25	85 (43)	19 (27)
25–34	45 (23)	17 (24)
35–44	32 (16)	8 (11)
45–54	10 (5)	3 (4)
55–64	5 (3)	0 (0)
65+	8 (4)	2 (3)
Unknown	13 (7)	21 (30)
Previous TB		
Yes	55 (28)	ND
No	141 (71)	ND
Unknown	2 (1)	70 (100)
HIV status		
Positive	9 (5)	9 (13)
Negative	188 (95)	35 (50)
Unknown	1 (1)	26 (37)
M. tuberculosis resistance		
MDR	17 (9)	10 (14)
not MDR	177 (89)	47 (67)
Unknown	4 (2)	13 (19)

doi:10.1371/journal.pone.0049651.t001

can discriminate between the early ancient (RD181[+]) and late ancient (RD181[-]) Beijing strains [19]. Polymorphic nucleotides of the respective isolates were determined according to Hanekom et al. [17]. STs were designated according to Filliol et al. [48] and Iwamoto et al. [19].

Data Retrieved from Previous Reports

15-MIRU-VNTR data and the results of subfamily classifications were retrieved from previous publications for the isolates in China [50], Japan [18], and Korea [20]. In order to compare the population structure, based on Filliol's STs, further sublineage classifications for the Chinese and Korean isolates were determined according to the phylogenetically informative VNTR loci data proposed by Wada and Iwamoto [51]. The VNTR data from Peru and these 3 East Asian countries were compared by constructing MST, based on 15-VNTR-MIRU, as previously described [23].

Statistical Analysis

Fisher's exact test (SPSS 17.0; IBM, NY, USA) was used to determine the association of large cluster-forming isolates with patient gender, history of TB treatment, HIV status, and strain drug susceptibility; the association with the median age of patients was evaluated by Mann-Whitney *U*-test using PASW statistics 18 (IBM).

Results

Proportion of Beijing Family Strains in Peru

The Beijing family strain, with a prevalence of 9.3% (200/2140), was the third largest spoligotyping clade after the H3 and T1 clades in the study population sampled between 2008 and 2010. This ratio is the same as for the previous study (9.3%, 30/323) conducted in 2004 and 2006 [30] but higher than in 1999 (5.9%, 11/185) [25]. In another study setting, which collected Beijing strains between 1999 and 2004 from 4 distinct areas in Lima, the ratio was 5% (46/912, 2 Beijing strains were excluded in this study because of the lack of DNA). This would imply an increasing trend of Beijing family strains in Lima, Peru. We negated the existence of "Pseudo-Beijing strains" [34] in our sample set by the detection of RD 105 [-] [35] for all of the strains classified as early ancient type by SNP typing.

Classification of Beijing Family Strains by SNP Typing and MIRU-VNTR

Further subdivision of the Beijing family strains in Peru, according to the 10 loci SNP panel [17,19,48], revealed similar population structures between the 2 sample sets (2008–2010 and 1999–2006) (Table 2). The high prevalence of isolates associated with the modern subfamily in Peru is consistent with the worldwide trend, except in the case of Japan and Korea where

the ancient subfamily predominates [20,50,51,52,53,54]. When we compared the population structure of the Beijing family strains in Peru with China, Japan, and Korea (Table 2), the characteristics of the Peruvian population structure highlighted a high frequency of the ST10 sublineage. ST19/ST25, which evolved before the modern subfamily [16,51], predominated in the ancient subfamily, as has been seen in China. Recently, Wada et al. proposed the use of SNP at 1576481 instead of 909166 due to the probable homoplastic behavior of SNP 909166 [55]. This is especially critical for discriminating between STK and ST3, which are highly prevalent in Japan. We, therefore, applied it for 4 ST3 stains in our sample set and confirmed them as true ST3 sublineage. Interestingly, the phylogenetic informativity of certain VNTR loci reported in Asian strains [16,51] was also retained for the strains in Peru (Table 3). Specifically, VNTR 4156, 1955, and 3155 demonstrated high sensitivity and specificity for the sublineages classification.

The 15-MIRU-VNTR data from the modern subfamily of Beijing strains in China, Korea, and Japan were retrieved from previous reports [18,20,50] and compared with the data of the population-based study in Peru (2008–2010) (Fig. 1). Two extremely large clusters were found in the MST, one (n = 103) composed of almost exclusively Peruvian isolates (101 Peruvian isolates and 2 Japanese isolates) (C1 in Fig. 1) and the other (n = 88) composed of isolates from all of the 4 countries (C2 in Fig. 1). None of the large clusters were composed exclusively of Japanese, Chinese, or Korean isolates. The biggest cluster, C1, did not show a star-like network with the others; rather, it had a nearly terminal topological position. In addition, the VNTR profile could not be found in the MIRU-VNTR*plus* database [56] or in other reports from Russia and China [57,58]. Together, these observations suggested that this genotype emerged as a uniquely endemic clone in Peru. The other large cluster (C2) was found at the core position and was connected with many isolates, regardless of their geographical origin. Because the Peruvian isolates belonging to the C2 cluster were further subdivided into many different genotypes by the 24_{Beijing}-VNTR analysis (data not shown), it is clear that the large cluster was formed because of the convergence of VNTR profiles. Other than these 2 large clusters, all of the other isolates dispersed equally (Fig. 1), i.e., no other branches consisted solely of Peruvian isolates.

Allelic Diversity of 24 VNTR Loci in Peru

The allelic diversity of each of the 24 VNTR loci is listed in Table S4, including the data from a previous report for Russia, Japan, and China [43]. Overall, the allelic diversity of the VNTR loci in Peru was much lower than in China or Japan. This low diversity was similar to that observed in Russia, where the Beijing family strains are considered to have recently emerged [57,59]. Only 11 out of the 24 loci had a PIC value greater than 0.1. Of these loci, 7 were in the 9 additional VNTR loci described in the Methods section. The PIC for the 3 hypervariable loci [4] were high: VNTR 3232 (0.710), VNTR 3820 (0.592), and VNTR 4120 (0.495). These results clearly supported the necessity of the use of the 9 additional loci for improving the discriminating power of VNTR genotyping for these strains. As expected, the use of 24_{Beijing}-VNTR improved discrimination compared to that of 15-MIRU-VNTR (Table 4).

Transmission Characteristics of Beijing Family Strains

Although 24_{Beijing}-VNTR improved the power of strain discrimination, the results showed high levels of clustering (80.3%) and RTI_{n-1} (0.707) (Table 4). This strongly suggests the active and on-going transmission of Beijing family strains in the survey area. Notably, a large size clustering (43.9%, 87/198) was identified, which was named Peru Cluster Type 001 (PCT001) (Fig. 2). The PCT001 strains belong to the largest cluster (C1) in 15 MIRU-VNTR (Fig. 1), therefore, they are considered a singular genotype in Peru. In the MST (Fig. 2), it formed a star-like VNTR-based network with the probable derivatives from it, mainly resulting from single locus changes. The results suggest the continuous evolution of the clone through active transmission and thus being a "currently successful clone" of the Beijing family strains in Peru. The second and third largest clusters were much smaller than PCT001:15 isolates (7.6%) and 10 isolates (5.1%), respectively. Although the number of multidrug-resistant *M. tuberculosis* (MDR-TB) belonging to the PCT001 cluster is small at the moment (n = 6), this should be an addressed as a potential public health threat. The comparison between patients harboring the PCT001 strains and those of other patients harboring Beijing strains revealed that only gender is significantly different ($P = 0.02$) (Table 5).

To ascertain whether the PCT001 genotype strains were present before the surveillance periods (in 2008–2010), an MST

Table 2. Distribution of Beijing sublineage strains.

SNP type	No. (%) of isolates					Definition[2]
	Peru		Japan	China[1]	Korea[1]	
	2008–2010 (n = 198)	1999–2006 (n = 70)	ref [18] (n = 714)	ref [50] (n = 187)	ref [20] (n = 62)	
ST11	0 (0)	0 (0)	3 (0.4)	9 (4.8)	29 (46.8)	Early ancient
ST26	0 (0)	4 (5.7)	50 (7.0)			
STK	0 (0)	0 (0)	111 (15.5)	0	2 (3.2)	Late ancient
ST3	3 (1.5)	1 (1.4)	182 (25.5)	3 (1.6)	3 (4.8)	
ST25	1 (0.5)	0 (0)	6 (0.8)	52 (27.8)	10 (16.1)	
ST19	8 (4.0)	5 (7.1)	195 (27.3)			
ST10	178 (89.9)	53 (75.7)	135 (18.9)	93 (49.7)	18 (29.0)	Modern
ST22	8 (4.0)	7 (10.0)	32 (4.5)	30 (16.0)		

[1]Two isolates in Korea and 4 isolates in China could not be assigned sublineages and excluded from the analysis.
[2]Early ancient, RD181 [+]; Late ancient, RD181 [−].
doi:10.1371/journal.pone.0049651.t002

Table 3. Specific VNTR allele for Beijing sublineages *M. tuberculosis* in Peru (n = 268).

VNTR locus	Specific allele	Corresponding sublineage(s)	No. of isolates	Sensitivity (%)	Specificity (%)
4156	3	Modern	246	246/246 (100)	246/248 (99.2)
	4	Early Ancient	4	4/4 (100)	4/4 (100)
	5	Late Ancient	16	16/16 (100)	16/18 (88.9)
1955	4	Modern	246	235/246 (95.5)	235/235 (100)
3155	2	Early Ancient	4	4/4 (100)	4/4 (100)

doi:10.1371/journal.pone.0049651.t003

was constructed based on the $24_{Beijing}$-VNTR profiles of the 268 Beijing family strains, comprised of the 198 strains (2008–2010) and the additional 70 strains isolated between 1999 and 2006 (Fig. 3). Twelve strains, which formed the second largest cluster among the 70 Beijing family strains, belonged to PCT001 (1 in 1999, 2 in 2000, 4 in 2001, 2 in 2004, and 3 in 2005), indicating that the genotype had existed since at least 1999. A striking difference between the 2 sample sets was found in the ratio of PCT001 genotype strains to total Beijing strains, i.e., a larger cluster (13 strains) than PCT001 was found in the 70 Beijing strains (PCT002 in Fig. 3). When we just focused on the strains isolated in South Lima (24/70 [34.3%]), the overlapping area of these 2 sample sets, PCT002 was still identified as the largest cluster (7 strains) followed by PCT001 (6 strains). These results suggested that the prevalence of PCT001 in the survey area increased over a short period.

Discussion

The high prevalence of Beijing family strains in Peru (5.9%, 11/185) compared to other South American countries was first reported as the result of a survey of 7 countries on the continent [25]. Very recently, Taype et al. [30] reported that the proportion of TB patients with the Beijing family strains in Peru was 9.3% (30/323). The report nicely described the genetic diversity of all the *M. tuberculosis* in Peru. However, the population size was too

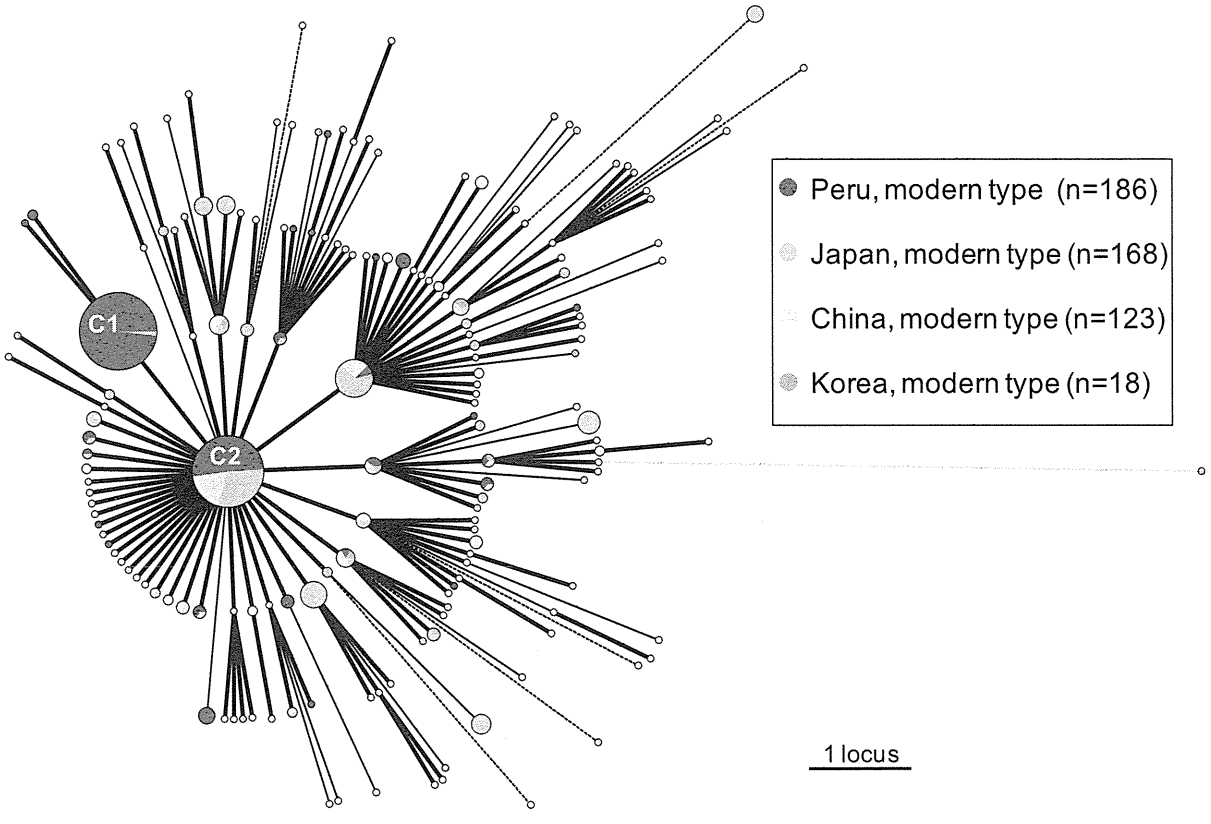

Figure 1. A minimum spanning tree based on 15 loci of a variable number of tandem repeat(s) of mycobacterial interspersed repetitive units (15-MIRU-VNTR) genotyping of the modern subfamily of *M. tuberculosis* Beijing strains from Peru (n = 186), Japan (n = 168), China (n = 123), and Korea (n = 18). Circles correspond to the different types discriminated by 15-MIRU-VNTR genotypes. Their sizes are proportional to the numbers of isolates sharing an identical pattern. The origin of each isolate is represented by different colors. Heavy lines connecting 2 types denote single-locus variants; thin lines connect double-locus variants; and dotted lines (black), triple-locus variants. The gray dotted lines indicate the most likely connection between 2 types differing by more than 3 VNTR loci.
doi:10.1371/journal.pone.0049651.g001

Table 4. Clustering analysis of 198 Beijing family strains (2008–2010) in Peru.

	No. of patterns	No. of clusters	No. clustered isolates	Clustering rate (%)	RTI_{n-1}	HGDI
15 VNTR	31	14	181	91.4	0.843	0.688
24 VNTR	58	19	159	80.3	0.707	0.797

doi:10.1371/journal.pone.0049651.t004

small to investigate the Beijing family strains in detail. In the current study, with a much larger population size and a well-designed sample collection scheme between 2008 and 2010, the genetic diversity and transmission dynamics of Beijing strains in Peru was clarified. The results highlighted the following: 1) an increasing prevalence of Beijing family strains during a past decade in Peru, 2) high clonality of the Beijing family strains, suggesting active and ongoing transmission, 3) the successful clone "PCT001 genotype," which has existed since at least 1999 as a singular clone in Peru, and 4) China as the greatest contributor of imported Beijing family strains into Peru. This study is limited by being unable to determine whether the highly active, ongoing transmission of Beijing family strains is specific for the family or a common trend for all strains in Peru. However, we could find a similarly high clustering rate of Beijing family strains (24/30 [80%]), compared with that of the non-Beijing family strains (174/293 [59.4%]) in the previous study using 12-MIRU-VNTR and spoligotyping [30]. This higher clustering in Beijing family strains implies a high transmission of it in Peru (more successful than non-Beijing strains). This trend could be confirmed later on by further VNTR analysis of all of the 2140 isolates.

A longitudinal comparison with the 70-strain sample set collected between 1999 and 2006 suggested that the successful clone, PCT001, was already present in Lima, by 1999, but at a lower prevalence (n = 12 [17.1%]) (Fig. 3). This difference in the ratio of PCT001 genotype strains to the total Beijing strains between the 2 sample sets suggests that the increase in PCT001 prevalence occurred recently, over a relatively short period. HIV infection, one of the high-risk characteristics for a large-scale outbreak [2], did not explain this high prevalence of PCT001 strains in the survey area (Table 5). One of the possible explanations for the high prevalence of PCT001 strains could be that it is highly transmissible and/or has increased virulence. The highly prevalent strains from large clusters have been previously reported to be more virulent than the sporadic strains of lower prevalence [59,60]. An attractive hypothesis is that PCT001 strains gained a selective advantage that allowed them to have spread more easily between the 2 sample collection periods (1999–2006 and 2008–2010).

The 15-MIRU-VNTR data from the modern subfamily of Beijing strains in China, Japan, and Korea were retrieved from previous reports [18,20,50] for comparison with the Peruvian isolates. The topology of the MST, based on these data (Fig. 1),

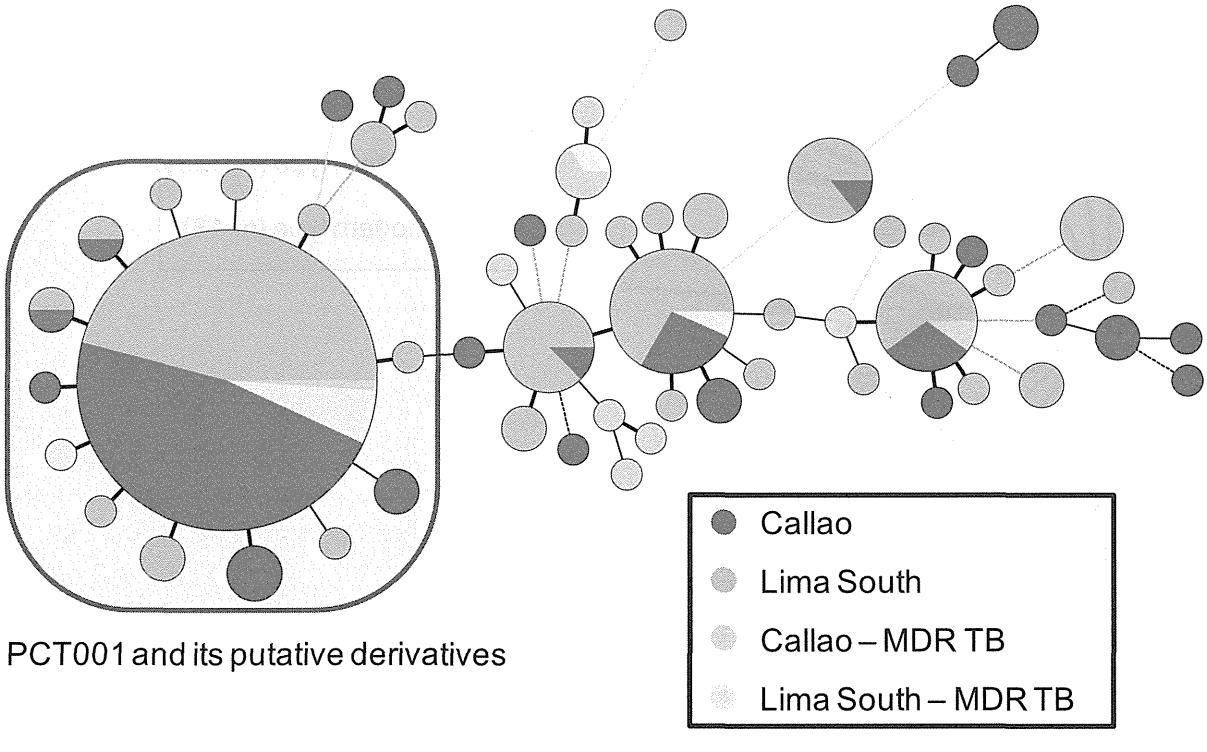

Figure 2. A minimum spanning tree of 198 Beijing family strains from Peru based on the 24-loci variable number of tandem repeats (VNTR). The colors of the circles represent the areas where the strain was isolated and its multidrug-resistant status. The designations for each circle and line in the tree are the same as in Fig. 1.
doi:10.1371/journal.pone.0049651.g002

Table 5. Demographics of successful clone "PCT001" and other strains.

Variable	No. of strains (%)		*P* value
	PCT001 (n = 87)	Others (n = 111)	
Gender			
Male	67 (77)	66 (59)	0.02
Female	20 (23)	43 (39)	
Unknown	0 (0)	2 (2)	
Median age (range)	23.5 (12–69)	27 (12–83)	0.07
Unknown	4 (5)	9 (8)	
Previous TB			1
Yes	24 (28)	31 (28)	
No	63 (72)	78 (70)	
Unknown	0 (0)	2 (2)	
HIV status			0.30
Positive	2 (2)	7 (6)	
Negative	85 (98)	103 (93)	
Unknown	0 (0)	1 (1)	
MDR TB			0.45
Yes	6 (7)	11 (10)	
No	81 (93)	96 (86)	
Unknown	0 (0)	4 (4)	

doi:10.1371/journal.pone.0049651.t005

suggested that the PCT001 genotype was a uniquely endemic clone in Peru. Except for the 2 largest clusters, the VNTR patterns from these countries were similarly diverse and none of the branches was composed exclusively of Peruvian isolates. This result would further support the earlier introduction of different ancestral strains into Peru from Asia during the past 150 years, an idea that was proposed on the basis of the diversity in IS*6110* RFLP patterns in a previous study [25].

The 10 loci SNPs could identify the modern Beijing lineage ST10 as the predominant sublineage in Peru. Very recently, this sublineage was reported as the most common in Taiwan and Thailand [16,61]. Allelic distribution of VNTR loci in each sequence type of Peruvian samples revealed the phylogenetic informativity of 3 MIRU-VNTR loci (4156, 1955, 3155) (Table 3). This is consistent with the results for East Asian strains [51] and could be an evidence for sharing the common ancestors of the Beijing family strains in Peru with those in East Asian countries. With the tremendous increase in whole-genome sequencing data, an increasing number of SNP typing systems have been developed [54,62,63]. The 10 loci SNPs typing is useful for a classification of Beijing family strains into sublineages level but apparently needed more discriminatory power to compensate for the VNTR homoplasy effect (VNTR-based clusters with mixed SNP sublineages), which was demonstrated with optical sets of 8 SNPs in Shanghai [50]. Moreover, inclusion of the lineage specific SNPs for the clades other than Beijing family [62,63] would expand the potential for phylogenetic studies. Further elaboration and optimization of the SNP sets would facilitate future molecular epidemiology and phylogenetic studies on *M. tuberculosis* in Peru.

In conclusion, the current results revealed the predominance of the modern subfamily and active transmission of Beijing family strains within Peru. Moreover, the emergence of the highly prevalent strains with the PCT001 genotype was also detected. The importation of Beijing family strains into European countries from Peru has already been reported [31,32], raising concern over transnational transmission. Future trends regarding the prevalence of PCT001 strains and the changes in population structure need to be carefully monitored from both the local and global epidemiological standpoints.

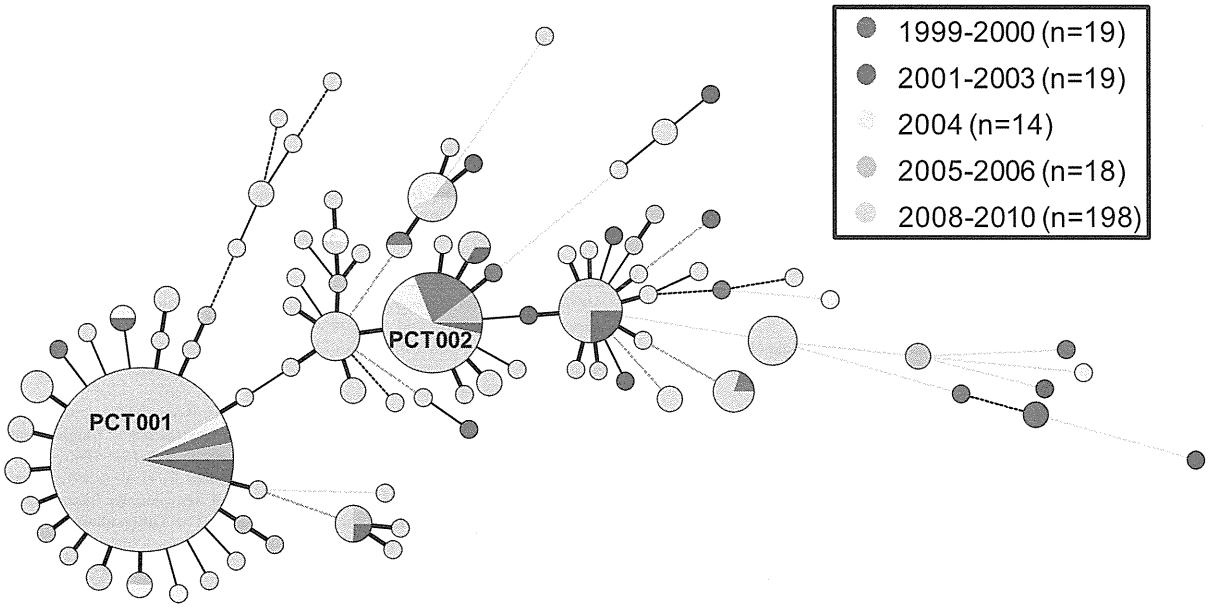

Figure 3. A minimum spanning tree of 268 Beijing family strains comprised of the 198 strains from the population-based study between 2008 and 2010 and an additional 70 strains isolated between 1999 and 2006. The colors of the circles represent the years of isolation. The designations for the circles and lines in the tree are the same as in Fig. 1.
doi:10.1371/journal.pone.0049651.g003

Supporting Information

Figure S1 Quality control approach for VNTR.
(PDF)

Table S1 Description of *M. tuberculosis* Beijing family strains obtained between December 2008 and January 2010 in Lima.
(XLS)

Table S2 Description of the additional 70 *M. tuberculosis* Beijing family strains in Lima, Peru.
(XLS)

Table S3 Locus designations and PCR primer sequences of the VNTR locus.
(XLS)

Table S4 Allelic diversity of VNTR loci.
(XLS)

Acknowledgments

We thank D. Sara, J.B. Phu, P. Fuentes, W. Solano, and T. Ichijo for excellent technical assistance.

Author Contributions

Conceived and designed the experiments: TI LG DM RG. Performed the experiments: TI LG KA NN JC LC PS. Analyzed the data: TI LG KA NN RG. Contributed reagents/materials/analysis tools: JC LC PS TW CT MS DM RG. Wrote the paper: TI LG CT MS RG.

References

1. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, et al. (1995) Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia. J Clin Microbiol 33: 3234–3238.
2. Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN (2002) Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. Trends Microbiol 10: 45–52.
3. Glynn JR, Whiteley J, Bifani PJ, Kremer K, van Soolingen D (2002) Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. Emerg Infect Dis 8: 843–849.
4. Iwamoto T, Yoshida S, Suzuki K, Tomita M, Fujiyama R, et al. (2007) Hypervariable loci that enhance the discriminatory ability of newly proposed 15-loci and 24-loci variable-number tandem repeat typing method on *Mycobacterium tuberculosis* strains predominated by the Beijing family. FEMS Microbiol Lett 270: 67–74.
5. Kremer K, Glynn JR, Lillebaek T, Niemann S, Kurepina NE, et al. (2004) Definition of the Beijing/W lineage of *Mycobacterium tuberculosis* on the basis of genetic markers. J Clin Microbiol 42: 4040–4049.
6. Mokrousov I (2008) Genetic geography of *Mycobacterium tuberculosis* Beijing genotype: a multifacet mirror of human history? Infect Genet Evol 8: 777–785.
7. Wan K, Liu J, Hauck Y, Zhang Y, Zhao X, et al. (2011) Investigation on *Mycobacterium tuberculosis* diversity in China and the origin of the Beijing clade. PLoS One 6: e29190.
8. Devaux I, Kremer K, Heersma H, Van Soolingen D (2009) Clusters of multidrug-resistant *Mycobacterium tuberculosis* cases, Europe. Emerg Infect Dis 15: 1052–1060.
9. Affolabi D, Faihun F, Sanoussi N, Anyo G, Shamputa IC, et al. (2009) Possible outbreak of streptomycin-resistant *Mycobacterium tuberculosis* Beijing in Benin. Emerg Infect Dis 15: 1123–1125.
10. 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.
11. Parwati I, Alisjahbana B, Apriani L, Soetikno RD, Ottenhoff TH, et al. (2010) *Mycobacterium tuberculosis* Beijing genotype is an independent risk factor for tuberculosis treatment failure in Indonesia. J Infect Dis 201: 553–557.
12. Gagneux S, Small PM (2007) Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis 7: 328–337.
13. Reed MB, Gagneux S, Deriemer K, Small PM, Barry CE 3rd (2007) The W-Beijing lineage of *Mycobacterium tuberculosis* overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated. J Bacteriol 189: 2583–2589.
14. Mokrousov I, Ly HM, Otten T, Lan NN, Vyshnevskyi B, et al. (2005) Origin and primary dispersal of the *Mycobacterium tuberculosis* Beijing genotype: clues from human phylogeography. Genome Res 15: 1357–1364.
15. Mokrousov I, Narvskaya O, Otten T, Vyazovaya A, Limeschenko E, et al. (2002) Phylogenetic reconstruction within *Mycobacterium tuberculosis* Beijing genotype in northwestern Russia. Res Microbiol 153: 629–637.
16. Faksri K, Drobniewski F, Nikolayevskyy V, Brown T, Prammananan T, et al. (2011) Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on IS6110, SNP, LSP and VNTR profiles from Thailand. Infect Genet Evol 11: 1142–1149.
17. Hanekom M, van der Spuy GD, Streicher E, Ndabambi SL, McEvoy CR, et al. (2007) 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 45: 1483–1490.
18. Iwamoto T, Fujiyama R, Yoshida S, Wada T, Shirai C, et al. (2009) Population structure dynamics of *Mycobacterium tuberculosis* Beijing strains during past decades in Japan. J Clin Microbiol 47: 3340–3343.
19. 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.
20. Kang HY, Wada T, Iwamoto T, Maeda S, Murase Y, et al. (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.
21. Maeda S, Wada T, Iwamoto T, Murase Y, Mitarai S, et al. (2010) Beijing family *Mycobacterium tuberculosis* isolated from throughout Japan: phylogeny and genetic features. Int J Tuberc Lung Dis 14: 1201–1204.
22. Millet J, Miyagi-Shiohira C, Yamane N, Mokrousov I, Rastogi N (2012) High-resolution MIRU-VNTRs typing reveals the unique nature of *Mycobacterium tuberculosis* Beijing genotype in Okinawa, Japan. Infect Genet Evol 12: 637–641.
23. Wada T, Iwamoto T, Maeda S (2009) Genetic diversity of the *Mycobacterium tuberculosis* Beijing family in East Asia revealed through refined population structure analysis. FEMS Microbiol Lett 291: 35–43.
24. Yokoyama E, Hachisu Y, Hashimoto R, Kishida K (2012) Population genetic analysis of *Mycobacterium tuberculosis* Beijing subgroup strains. Infect Genet Evol 12: 630–636.
25. Ritacco V, Lopez B, Cafrune PI, Ferrazoli L, Suffys PN, et al. (2008) *Mycobacterium tuberculosis* strains of the Beijing genotype are rarely observed in tuberculosis patients in South America. Mem Inst Oswaldo Cruz 103: 489–492.
26. Gomes HM, Elias AR, Oelemann MA, Pereira MA, Montes FF, et al. (2012) Spoligotypes of *Mycobacterium tuberculosis* complex isolates from patients residents of 11 states of Brazil. Infect Genet Evol 12: 649–656.
27. Cerezo I, Jimenez Y, Hernandez J, Zozio T, Murcia MI, et al. (2012) A first insight on the population structure of *Mycobacterium tuberculosis* complex as studied by spoligotyping and MIRU-VNTRs in Bogota, Colombia. Infect Genet Evol 12: 657–663.
28. Candia N, Lopez B, Zozio T, Carrivale M, Diaz C, et al. (2007) First insight into *Mycobacterium tuberculosis* genetic diversity in Paraguay. BMC Microbiol 7: 75.
29. Abadia E, Sequera M, Ortega D, Mendez MV, Escalona A, et al. (2009) *Mycobacterium tuberculosis* ecology in Venezuela: epidemiologic correlates of common spoligotypes and a large clonal cluster defined by MIRU-VNTR-24. BMC Infect Dis 9: 122.
30. Taype CA, Agapito JC, Accinelli RA, Espinoza JR, Godreuil S, et al. (2012) Genetic diversity, population structure and drug resistance of *Mycobacterium tuberculosis* in Peru. Infect Genet Evol 12: 577–585.
31. Garcia de Viedma D, Chaves F, Inigo J (2006) New route of importation of *Mycobacterium tuberculosis* Beijing genotype. Emerg Infect Dis 12: 169–170.
32. Lari N, Rindi L, Bonanni D, Rastogi N, Sola C, et al. (2007) Three-year longitudinal study of genotypes of *Mycobacterium tuberculosis* isolates in Tuscany, Italy. J Clin Microbiol 45: 1851–1857.
33. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, et al. (1997) Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol 35: 907–914.
34. Fenner L, Malla B, Ninet B, Dubuis O, Stucki D, et al. (2011) "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of *Mycobacterium tuberculosis*. PLoS One 6: e24737.
35. Tsolaki AG, Hirsh AE, DeRiemer K, Enciso JA, Wong MZ, et al. (2004) Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. Proc Natl Acad Sci U S A 101: 4865–4870.
36. Caviedes L, Lee TS, Gilman RH, Sheen P, Spellman E, et al. (2000) Rapid, efficient detection and drug susceptibility testing of *Mycobacterium tuberculosis* in sputum by microscopic observation of broth cultures. The Tuberculosis Working Group in Peru. J Clin Microbiol 38: 1203–1208.
37. Moore DA, Mendoza D, Gilman RH, Evans CA, Hollm Delgado MG, et al. (2004) Microscopic observation drug susceptibility assay, a rapid, reliable diagnostic test for multidrug-resistant tuberculosis suitable for use in resource-poor settings. J Clin Microbiol 42: 4432–4437.
38. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, et al. (2006) Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J Clin Microbiol 44: 4498–4510.

39. Murase Y, Mitarai S, Sugawara I, Kato S, Maeda S (2008) Promising loci of variable numbers of tandem repeats for typing Beijing family *Mycobacterium tuberculosis*. J Med Microbiol 57: 873–880.
40. Zhang L, Chen J, Shen X, Gui X, Mei J, et al. (2008) Highly polymorphic variable-number tandem repeats loci for differentiating Beijing genotype strains of *Mycobacterium tuberculosis* in Shanghai, China. FEMS Microbiol Lett 282: 22–31.
41. Velji P, Nikolayevskyy V, Brown T, Drobniewski F (2009) Discriminatory ability of hypervariable variable number tandem repeat loci in population-based analysis of *Mycobacterium tuberculosis* strains, London, UK. Emerg Infect Dis 15: 1609–1616.
42. Smittipat N, Billamas P, Palittapongarnpim M, Thong-On A, Temu MM, et al. (2005) Polymorphism of variable-number tandem repeats at multiple loci in *Mycobacterium tuberculosis*. J Clin Microbiol 43: 5034–5043.
43. Wang J, Liu Y, Zhang CL, Ji BY, Zhang LZ, et al. (2011) Genotypes and characteristics of clustering and drug susceptibility of *Mycobacterium tuberculosis* isolates collected in Heilongjiang Province, China. J Clin Microbiol 49: 1354–1362.
44. Keim P, Price LB, Klevytska AM, Smith KL, Schupp JM, et al. (2000) Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J Bacteriol 182: 2928–2936.
45. Hunter PR, Gaston MA (1988) Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol 26: 2465–2466.
46. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, et al. (1994) The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. N Engl J Med 330: 1703–1709.
47. Durmaz R, Zozio T, Gunal S, Allix C, Fauville-Dufaux M, et al. (2007) Population-based molecular epidemiological study of tuberculosis in Malatya, Turkey. J Clin Microbiol 45: 4027–4035.
48. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, et al. (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.
49. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393: 537–544.
50. Luo T, Yang C, Gagneux S, Gicquel B, Mei J, et al. (2012) Combination of single nucleotide polymorphism and variable-number tandem repeats for genotyping a homogenous population of *Mycobacterium tuberculosis* Beijing strains in China. J Clin Microbiol 50: 633–639.
51. Wada T, Iwamoto T (2009) Allelic diversity of variable number of tandem repeats provides phylogenetic clues regarding the *Mycobacterium tuberculosis* Beijing family. Infect Genet Evol 9: 921–926.
52. Mokrousov I, Jiao WW, Sun GZ, Liu JW, Valcheva V, et al. (2006) Evolution of drug resistance in different sublineages of *Mycobacterium tuberculosis* Beijing genotype. Antimicrob Agents Chemother 50: 2820–2823.
53. Dou HY, Tseng FC, Lu JJ, Jou R, Tsai SF, et al. (2008) Associations of *Mycobacterium tuberculosis* genotypes with different ethnic and migratory populations in Taiwan. Infect Genet Evol 8: 323–330.
54. Mestre O, Luo T, Dos Vultos T, Kremer K, Murray A, et al. (2011) Phylogeny of *Mycobacterium tuberculosis* Beijing strains constructed from polymorphisms in genes involved in DNA replication, recombination and repair. PLoS One 6: e16020.
55. Wada T, Iwamoto T, Hase A, Maeda S (2012) Scanning of genetic diversity of evolutionarily sequential *Mycobacterium tuberculosis* Beijing family strains based on genome wide analysis. Infect Genet Evol 12: 1392–1396.
56. Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D (2010) MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic Acids Res 38: W326–331.
57. Mokrousov I, Narvskaya O, Vyazovaya A, Millet J, Otten T, et al. (2008) *Mycobacterium tuberculosis* Beijing genotype in Russia: in search of informative variable-number tandem-repeat loci. J Clin Microbiol 46: 3576–3584.
58. Jiao WW, Mokrousov I, Sun GZ, Guo YJ, Vyazovaya A, et al. (2008) Evaluation of new variable-number tandem-repeat systems for typing *Mycobacterium tuberculosis* with Beijing genotype isolates from Beijing, China. J Clin Microbiol 46: 1045–1049.
59. Lasunskaia E, Ribeiro SC, Manicheva O, Gomes LL, Suffys PN, et al. (2010) Emerging multidrug resistant *Mycobacterium tuberculosis* strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence. Microbes Infect 12: 467–475.
60. Hernandez-Pando R, Marquina-Castillo B, Barrios-Payan J, Mata-Espinosa D (2012) Use of mouse models to study the variability in virulence associated with specific genotypic lineages of *Mycobacterium tuberculosis*. Infect Genet Evol 12: 725–731.
61. Chen YY, Chang JR, Huang WF, Kuo SC, Su IJ, et al. (2012) Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on SNP and VNTR typing profiles in Asian countries. PLoS One 7: e39792.
62. Homolka S, Projahn M, Feuerriegel S, Ubben T, Diel R, et al. (2012) High resolution discrimination of clinical *Mycobacterium tuberculosis* complex strains based on single nucleotide polymorphisms. PLoS One 7: e39855.
63. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, et al. (2012) Two New Rapid SNP-Typing Methods for Classifying *Mycobacterium tuberculosis* Complex into the Main Phylogenetic Lineages. PLoS One 7: e41253.

Infection, Genetics and Evolution 21 (2014) 479–483

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Intra-subspecies sequence variability of the MACPPE12 gene in *Mycobacterium avium* subsp. *hominissuis*

Tomotada Iwamoto [a,*], Kentaro Arikawa [a], Chie Nakajima [b], Noriko Nakanishi [a], Yukiko Nishiuchi [c], Shiomi Yoshida [d], Aki Tamaru [e], Yutaka Tamura [f], Yoshihiko Hoshino [g], Heekyung Yoo [h], Young Kil Park [h], Hajime Saito [i], Yasuhiko Suzuki [b]

[a] *Department of Infectious Diseases, Kobe Institute of Health, 4-6 Minatojima-nakamachi, Chuo-ku, Kobe 650-0046, Japan*
[b] *Division of Global Epidemiology, Hokkaido University Research Center for Zoonosis Control, Kita-20, Nishi-10, Kita-ku, Sapporo 001-0020, Japan*
[c] *Toneyama Institute for Tuberculosis Research, Osaka City University Medical School, Osaka 560-8552, Japan*
[d] *Clinical Research Center, National Hospital Organization Kinki-chuo Chest Medical Center, 1180 Nagasone-cho, Sakai, Osaka 591-8555, Japan*
[e] *Department of Bacteriology, Osaka Prefectural Institute of Public Health, 3-69, Nakamichi 1-chome, Higashinari-ku, Osaka 537-0025, Japan*
[f] *School of Veterinary Medicine, Rakuno Gakuen University, 583 Midorimachi Bunkodai, Ebetsu, Hokkaido 069-8501, Japan*
[g] *Leprosy Research Center, National Institute of Infectious Diseases, Tokyo 189-0002, Japan*
[h] *Research and Development, Korean Institute of Tuberculosis, Korean National Tuberculosis Association, 482, Mansu-ri, Kangwoi-myun, Cheongwon-gun, Chungbuk 363-954, South Korea*
[i] *Yoshijima Hospital, 2-33, Yoshijima-higashi 3 chome, Naka-ku, Hiroshima 730-0822, Japan*

ARTICLE INFO

Article history:
Received 10 June 2013
Received in revised form 11 August 2013
Accepted 12 August 2013
Available online 20 September 2013

Keywords:
Mycobacterium avium subsp. *hominissuis*
PE/PPE gene family
Variable numbers of tandem repeats
SNPs
Genetic diversity

ABSTRACT

The PE (Pro-Glu) and PPE (Pro-Pro-Glu) multigene families are unique to mycobacteria, and are highly expanded in the pathogenic members of this genus. We determined the intra-subspecies genetic variability of the MACPPE12 gene, which is a specific PPE gene in *Mycobacterium avium* subsp. *hominissuis* (MAH), using 334 MAH isolates obtained from different isolation sources (222 human isolates, 145 Japanese and 77 Korean; 37 bathroom isolates; and 75 pig isolates). In total, 31 single-nucleotide polymorphisms (SNPs), which consisted of 16 synonymous SNPs and 15 nonsynonymous SNPs, were determined through comparison with the MACPPE12 gene sequence of MAH strain 104 as a reference. As the result, the 334 MAH isolates were classified into 19 and 13 different sequevars at the nucleic acid level (NA types) and amino acid level (AA types), respectively. Among the 13 AA types, only one type, the AA02 type, presented various NA types (7 different types) with synonymous SNPs, whereas all other AA types had a one-to-one correspondence with the NA types. This finding suggests that AA02 is a longer discernible lineage than the other AA types. Therefore, AA02 was classified as an ancestral type of the MACPPE12 gene, whereas the other AA types were classified as modern types. The ubiquitous presence of AA02 in all of the isolation sources and all different sequevars classified by the *hsp65* genotype further supports this classification. In contrast to the ancestral type, the modern types showed remarkable differences in distribution between human isolates and pig isolates, and between Japanese isolates and Korean isolates. Divergence of the MACPPE12 gene may thus be a good indicator to characterize MAH strains in certain areas and/or hosts.

1. Introduction

Mycobacterial infections caused by strains of the *Mycobacterium avium* complex (MAC) are becoming increasingly prevalent in animals and humans (Falkinham, 2010; Turenne et al., 2006; Winthrop, 2010). In particular, *Mycobacterium avium* subsp. *hominissuis* (MAH) is a frequent agent of human and pig mycobacteriosis (Mijs et al., 2002). Although MAH is typically considered to be an opportunistic bacterium for immunocompromised persons, it also frequently occurs in immunocompetent individuals and generally manifests as a slowly progressive, often debilitating lung disease. Recently, middle-aged and elderly females without any predisposing conditions have been suggested to bear the brunt of this disease (Inagaki et al., 2009). Therefore, it has been speculated that MAH-associated mycobacteriosis is caused not only by host characteristics but also by bacterial factors (Ichikawa et al., 2009).

The PE (Pro-Glu) and PPE (Pro-Pro-Glu) multigene families are unique to mycobacteria and are suspected to be involved in immunostimulation and virulence (Gey van Pittius et al., 2006; Mackenzie et al., 2009; Sampson, 2011). The PE and PPE gene families are highly expanded in the pathogenic species of this genus (Gey van Pittius et al., 2006). Recently, Mackenzie et al., 2009 identified 12 PE and 49 PPE orthologs in the major groups of the MAC; *Mycobac-*

* Corresponding author. Tel.: +81 78 302 6251; fax: +81 78 302 0894.
E-mail address: tomotada_iwamoto@office.city.kobe.lg.jp (T. Iwamoto).

1567-1348/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2013.08.010

terium avium subsp. *paratuberculosis*, *M. avium* subsp. *hominissuis*, *Mycobacterium avium* subsp. *avium*, and *Mycobacterium intracellulare*. A genomic comparison among them identified the subspecies-specific PE/PPE genes and the missing PE/PPE genes from one subspecies but present in at least two members of the MAC (Mackenzie et al., 2009). The former are likely to emerge and/or be acquired after divergence into the certain subspecies, whereas the latter are likely to correspond to earlier deletions in the certain subspecies. Two PPE genes, i.e., Mav 0790c and Mav 2006, which are now denoted as MACPPE4 and MACPPE12, respectively, under the newly proposed uniform PE and PPE locus names for all members of the MAC, were specific for MAH strain 104. The corresponding gene products could be used to identify immune responses against this *M. avium* subspecies, and misinterpretations caused by cross-reactivity in current diagnostics for Johne's disease would thus be avoided (Mackenzie et al., 2009).

In this study, we first confirmed that MACPPE12 is ubiquitous in this subspecies, whereas MACPPE4 is not widely distributed in strains other than MAH strain 104. To determine the intra-subspecies genetic variations of the MACPPE12 gene, we sequenced the full length of the gene (1341 bp) using 334 isolates that were obtained from different sources, i.e., 222 human isolates (145 Japanese and 77 Korean), 37 bathroom isolates, and 75 pig isolates. We also determined whether the genetic variation was associated with the isolation source.

2. Material and methods

2.1. Bacterial isolates

We used crude DNA extracted from 334 isolates for this study. Of these isolates, 257 overlapped with 258 isolates that were previously identified as MAH by *hsp65* sequencing analyses and used to analyze genetic diversity (Iwamoto et al., 2012). One isolate from our previous study bank was excluded because of a lack of volume. We newly added 77 isolates from 77 human patients that were obtained from 7 different cities in Korea between 2010 and 2011. They were originally identified as *M. avium* through sequencing of the 16S rRNA gene (Devulder et al., 2005) at the Korean Institute of Tuberculosis. The *hsp65* sequencing analyses that were performed in this study confirmed that all of the isolates belonged to MAH. To compare the genetic diversity of the 77 MAH isolates with the previously obtained data for the other 257 isolates, the same genetic markers used in the previous study (Iwamoto et al., 2012), i.e., the 3′ portion of the *hsp65* gene sequence, presence of IS*Mav6*, and genotypes of the 19-locus variable number of tandem repeat (VNTR) sequence, were analyzed for the 77 isolates. The datasets used in this study consisted solely of sequence data and no personal data were disclosed at any point.

2.2. PCR and sequencing of MACPPE12

The MACPPE12 gene, the locus name of which in MAH strain 104 (accession number in GenBank, NC_008595) is Mav_2006, was amplified using the primer sets MAV2006F (5′-TGC GTG GTA ACA AAA GCA AC) and MAV2006R (5′-CTT GCT GCG TAA TGC GAT AA). The PCR reaction consisted of 94 °C for 3 min followed by 30 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min, with a final extension at 72 °C for 7 min. PCR was performed using Ex Taq Hot Start Version (TaKaRa Bio Inc., Shiga, Japan) with GC buffer I (TaKaRa Bio Inc., Shiga, Japan), and the PCR products were subjected to sequence analysis using an AB3500 genetic analyzer system (Applied Biosystems, Foster City, CA). The same primers used for PCR were also used for the sequencing of forward and reverse fragments. In addition, the interim primers MAV2006F634 (AAC GCG CTG CAG AAT CTC) and MAV2006R824 (TCC GTC ATC TTG TGT TCA GC) were used for the sequencing of forward and reverse fragments, respectively. Detailed information regarding VNTR genotypes, *hsp65* code types, presence of IS*Mav6*, and MACPPE12 sequevars of the 334 isolates in this study are summarized in Supplemental Table 1 (Table S1).

2.3. Phylogenetic analysis

The split-network phylogeny of the complete MACPPE12 gene sequence (1341 bp) was computed by NeighborNet analysis in SplitTree Version 4.8 (Huson and Kloepper, 2005). Recombination events in the MACPPE12 gene within the 334-isolate set were evaluated using DnaSP 4.10 (Rozas et al., 2003).

2.4. Nucleotide accession numbers

Sequences of the complete MACPPE12 gene representing each sequevar recognized in this study (NA types 2 to 19) were deposited in GenBank under accession Nos. AB820302 to AB820319.

3. Results

3.1. Presence of MACPPE4 and MACPPE12 in MAH

In a preliminary study, we first evaluated the ubiquitous presence of two previously reported MAH-specific MACPPE genes, i.e., MACPPE4 and MACPPE12, in MAH by using 16 randomly selected MAH isolates obtained from humans ($n = 6$), bathroom samples ($n = 2$), and pigs ($n = 8$). We attempted to detect the MACPPE4 gene using two PCR primer sets, one targeting the outside regions of MACPPE4, which can amplify the whole MACPPE4 gene with its flanking region, and the other targeting the inside sequences of MACPPE4, which can amplify partial regions of the gene. These primer sets produced expected sizes of PCR products from MAH strain 104 but the amplicons were not obtained from 16 other strains (data not shown). We therefore assumed that the MACPPE4 gene is not universally present in this subspecies, MAH. On the other hand, MACPPE12 was detected from all of the 16 isolates using the primer set targeting the outside regions of the gene, which can amplify whole MACPPE12 gene with its flanking region. Our expanding analysis for all of 334 samples could detect MACPPE12 from all of them. Therefore, it is highly likely that MACPPE12 is a ubiquitous gene in MAH.

3.2. Sequence variation of the MACPPE12 gene

First, we assured our sample set consisted of reasonably high heterogeneous isolates for the evaluation of the genetic variability and distribution of the MACPPE12 gene in MAH by 19-locus VNTR analysis. Actually, we retrieved the data from our previous study (Iwamoto et al., 2012) for 257 isolates and added newly analyzed data for 77 Korean isolates. The data demonstrated a reasonably high degree of genetic diversity in this sample set (Table 1 and Table S1). In brief, 99 genotypes in 145 Japanese isolates, 49 genotypes in 77 Korean isolates, 27 genotypes in 37 bathroom isolates, and 38 genotypes in 75 pig isolates.

The sequence analysis of the full length of the MACPPE12 gene for 334 MAH isolates identified in total 31 SNPs, which formed 19 different MACPPE12 sequevars at the nucleic acid (NA) level (NA type) through comparison with the MACPPE12 gene sequence of MAH strain 104 as a reference (Table 1). Of the 31 SNP positions, 15 positions were nonsynonymous SNPs (nsSNPs) that caused amino acid substitutions. This relatively high ratio of nsSNPs resulted in the formation of 13 different sequevars at the amino acid (AA)

Table 1
SNPs among the *M. avium* subsp. *hominissuis* in comparison with prototype strain *M. avium* subsp. *hominissuis* strain 104 (MAH strain 104).

Amino acid type	Nucleic acid type	Nucleotide at the base pair position of strain MAH strain 104[a]																								Total number of:[b]			
		4*	66	88*	272*	468	558	571*	605*	715*	733	744	752*	765	822	829*	831	834	866*	867	868*	870	877*	924	978	1000	1005	1150*	1161	1163*	1192*	1330*	sSNP	nsSNP	Isolates (n = 334)	Genotypes
AA01	NA01	C	A	G	A	T	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	G	A			31		12
AA02	NA02	C	A	G	A	T	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	G	G	0	1	130		86
AA03	NA03	C	A	G	A	C	T	G	C	C	T	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	G	G	2	1	8		6
AA04	NA04	C	A	G	A	C	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	G	G	1	1	3		3
AA05	NA05	C	G	G	A	C	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	G	G	2	1	2		2
AA02	NA06	C	G	G	A	T	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	T	G	G	G	G	1	1	1		1
AA02	NA07	C	A	G	A	T	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	T	G	G	G	G	1	1	1		1
AA02	NA08	C	A	G	A	C	C	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	T	G	G	G	G	2	1	1		1
AA03	NA09	C	G	A	A	T	C	G	C	C	C	G	G	G	A	T	C	T	G	G	G	C	G	C	A	G	G	G	G	G	5	4	89		44
AA04	NA10	C	A	G	A	T	T	G	T	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	G	G	0	2	7		6
AA05	NA11	C	A	G	A	T	T	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	A	A	G	1	3	1		1
AA06	NA12	C	G	A	A	T	C	G	C	C	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	G	G	G	A	G	2	2	1		1
AA07	NA13	C	G	G	A	C	C	G	C	C	C	G	G	G	A	T	G	T	G	G	G	C	G	C	G	G	G	G	G	G	5	2	19		14
AA08	NA14	C	A	G	G	C	C	G	C	C	C	G	G	G	A	T	G	T	G	G	G	C	G	T	A	C	G	G	G	G	5	4	19		9
AA09	NA15	T	A	G	A	T	C	G	C	C	C	G	G	G	A	T	G	T	G	G	G	C	G	C	A	C	G	G	G	G	3	4	10		5
AA10	NA16	C	G	A	A	T	C	G	C	C	C	G	T	G	A	C	G	T	G	G	G	C	G	C	A	C	G	G	G	G	5	5	3		3
AA11	NA17	C	A	G	A	T	T	A	C	T	C	G	G	G	A	C	G	T	G	G	G	C	G	C	G	C	G	G	G	G	0	3	1		1
AA12	NA18	C	G	A	A	T	C	G	C	C	C	G	G	G	A	T	G	T	G	G	G	C	G	C	G	C	G	G	G	G	3	3	1		1
AA13	NA19	C	A	G	A	C	C	G	C	C	C	A	G	A	G	T	G	A	A	G	G	C	G	C	G	A	G	G	G	A	7	4	6		3

[a] Asterisks denote nonsynonymous position. AA01 and NA01 is the sequevars for MAH strain 104.
[b] sSNP, synonimous SNP; nsSNP, nonsynonymous SNP; genotypes were defined by 19-VNTR.

level (AA types) (Table 1). Among the 13 AA types, only AA02 showed variation in the nucleic acid sequence, i.e., it matched seven different NA types (NA02 to 08) that possessed various synonymous SNPs (sSNPs) (Table 1). All other AA types were associated in a one-to-one correspondence with NA types. These results suggest that AA02 has a longer history than the other AA types in MAH. We, therefore, classified AA02 as the ancestral type of MACPPE 12, whereas the other AA types were classified as modern types. When we used DnaSP, the minimum number of recombination events for the MACPPE12 gene in the sample set was estimated to be 3. The unrooted phylogeny for the gene sequences determined by SplitsTree4 demonstrated complex web-like topology (Fig. 1).

3.3. Comparison of the divergence of MACPPE12 in different isolation sources

The ancestral AA type of MACPPE12, AA02, was observed in all of the different isolation sources in this study (Table 2). In contrast, the modern types varied according to the isolation source. AA01 and AA07-09 were mostly present in pigs; AA03, in Japanese humans and bathroom samples; and AA13, in Korean humans. The topological positions of these AA types in the unrooted tree, except for that of AA01, were distinct from the position of the ancestral type (AA02) (Fig. 1).When we compared variation of MACPPE12 AA types with that of the *hsp65* gene sequevar, which consists of only sSNPs in this sample set, the AA02 type was observed in all 11 *hsp65* code types (Table 2). This would be natural due to the ancestral feature of AA02. Apart from the ancestral type, most of the modern types were distributed across more than two *hsp65* code types and not in a single lineage. This data strongly suggested that these two genes diverged independently, not in parallel to each other, during the history of MAH. It would be noteworthy that 57 pig isolates with *hsp65* code type 1, which was characterized as predominant code type of pig isolates in our previous study (Iwamoto et al., 2012), were subclassified into 5 AA types: AA01 (n = 26), AA02 (n = 4), AA07 (n = 12), AA08 (n = 11), and AA09 (n = 4) (Table 2, and Table S1).

A further finding in this study is the commonality and differences between Japanese and Korean human clinical isolates. A high prevalence of IS*Mav6*, which was reported in the genetic characterization of MAH in Japan (Ichikawa et al., 2009), was also observed in Korean isolates (42/77, 54.5%) as was reported by (Niimi et al., 2012). The characteristics of Korean isolates was demonstrated by their high prevalence of *hsp65* sequevar code 16 (26/77, 33.8%), which is a rare sequevar in other countries (Ichikawa et al., 2009; Iwamoto et al., 2012; Turenne et al., 2006). The majority of Korean isolates were ancestral type (AA02), whereas Japanese isolates predominated both AA02 and AA03. Moreover, other modern AA types showed different distributions between samples from these two countries.

4. Discussion

The precise functions of the PE and PPE families are still unknown except for a certain small number of members (Karboul et al., 2008; Mishra et al., 2008; Sassetti and Rubin, 2003), but these families are highly suspected to play key roles in the interaction between pathogens and their host (habitat) (Brennan et al., 2001; Sampson, 2011). A recent comparative study (Mackenzie et al., 2009) of the whole genomes of different MAC organisms revealed that two PPE paralogs, MACPPE4 and MACPPE12, were specifically found in MAH strain 104. Therefore, it would be a reasonable assumption that the two PPE genes reflect bacteriological characteristics of MAH in comparison with other MAC species.

Table 2
Characterization of MACPPE12 gene sequevars according to *hsp65* gene sequence and source of the 334 isolates.

MAC PPE 12		*hsp65* Code Type											Source		
AA type	NA type	C1	C2	C3	C7	C9	C15	C16	C17	N1	N2	N3	Human (Japan)	Human (Korea)	Bath-room	Pig
AA01	NA01	26	4	1	–	–	–	–	–	–	–	–	3	–	–	28
AA02	NA02-08	9	68	2	4	3	21	32	2	1	3	1	72	57	12	5
AA03	NA09	1	21	–	–	1	59	2	5	–	–	–	57	10	22	–
AA04	NA10	–	3	–	–	–	3	–	1	–	–	–	5	–	2	–
AA05	NA11	–	–	–	–	–	1	–		–	–	–	1	–	–	–
AA06	NA12	–	–	–	–	–	–	–	1	–	–	–	1	–	–	–
AA07	NA13	13	5	–	–	–	–	–	1	–	–	–	2	–	–	17
AA08	NA14	11	5	–	–	1	1	1	–	–	–	–	4	–	–	15
AA09	NA15	4	2	–	–	–	–	–	1	3	–	–	–	–	–	10
AA10	NA16	–	1	–	–	–	2	–	–	–	–	–	–	2	1	–
AA11	NA17	–	–	–	–	–	–	1	–	–	–	–	–	1	–	–
AA12	NA18	–	–	–	–	–	–	–	–	–	1	–	–	1	–	–
AA13	NA19	–	–	6	–	–	–	–	–	–	–	–	–	6	–	–

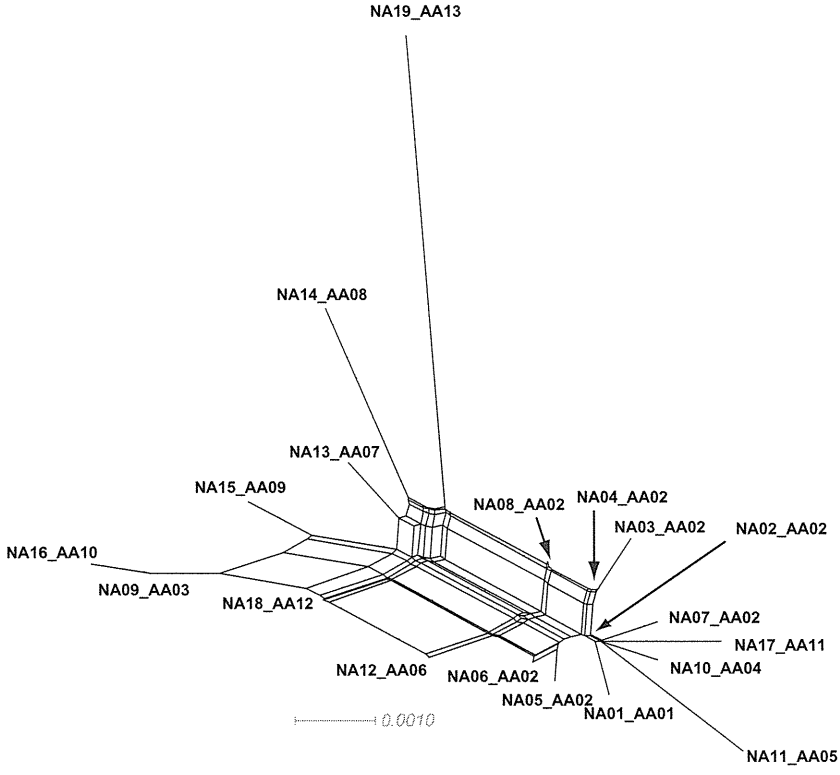

Fig. 1. Phylogenetic representation of each AA and NA type determined in this study and generated in SplitsTree4.

In this study, we evaluated the ubiquitous presence of these two PPE genes in MAH, and their genetic variability and association with the isolation sources and different genetic markers.

Our preliminary study using 16 MAH isolates suggested that MACPPE4 is not ubiquitous in MAH, although it is not certain that if this PPE gene is specific only for MAH strain104 or limited in subgroups of MAH. On the other hand, MACPPE12 was present in all of the 334 MAH isolates. Since our 334 isolates were a set of high heterogeneous isolates, the ubiquitous presence of the PPE gene in this sample set strongly supports the idea that this gene was present in the most recent common ancestor of MAH and universally retained in the subspecies. Because of the absence of the MACPPE12 gene in other members of the MAC and the ubiquitous presence in MAH isolates, MACPPE12 can be considered as a relatively new gene but ubiquitous in MAH.

By using the large number of isolates obtained from different sources, we demonstrated the variability of MACPPE12, which includes 19 different NA types and 13 AA types (Table 1). The web-like topology of the unrooted phylogeny (Fig. 1) and the estimated minimum number of recombination events suggest that genetic recombination plays a role in the divergence of this gene, although its mechanism is unknown. On the basis of the distribution of SNPs, isolation sources, and *hsp65* code types in the different MACPPE12 gene AA types, we classified the PPE gene into two groups; one is an ancestral type (AA02) and the other is a modern type. MAH is generally characterized as its ubiquitous host distribution and heterogeneous grouping (Turenne et al., 2008, 2007). However, when we look closely at the correlation between AA types and isolation sources, the distribution of the modern AA types well reflect their isolation sources, whereas the ancestral AA type (AA02) was observed in all of the different isolation sources. Thus, it can be hypothesized that emergence of the modern AA types somehow reflects an on-going evolution of MAH toward specialization (narrower range for host specificity and higher fitness to its habitat)