

Brief communication (Original)

Genetic characteristics of *Mycobacterium bovis* BCG Thai Red Cross Society strain

Wichit Thaveekarn¹, Sunchai Payungporn², Narumol Pakmanee¹, Sunutcha Suntrarachun¹, Suchitra Khunsap¹, Suthidee Petsong³

¹Queen Saovabha Memorial Institute, Thai Red Cross Society, Bangkok 10330, Thailand

²Department of Medical Biochemistry, Chulalongkorn University, Bangkok 10330, Thailand

³Department of Medical Microbiology, Chulalongkorn University, Bangkok 10330, Thailand

Background: Bacillus Calmette–Guérin (BCG) Tokyo 172-1 strain has been used as a parent strain since 1988 by Queen Saovabha Memorial Institute, Thai Red Cross Society to manufacture BCG vaccine in Thailand. In general, the BCG Tokyo strain includes 2 BCG type populations: types I and II. Several studies have found that the type I population is higher than type II. Polymerase chain reaction has shown the presence of 2 types of subpopulations in the BCG Thai Red Cross Society (TRCS) strain. However, the characteristics of all genes from BCG TRCS have not been studied in depth, including whether or not the original features of the BCG Tokyo 172-1 strain are still retained.

Objectives: To compare BCG TRCS and BCG Tokyo 172-1 for differences between their genomes.

Methods: The whole genome of BCG TRCS was analyzed by next-generation sequencing.

Results: BCG TRCS contained 23 different points and was 4 bases less in length than BCG Tokyo 172-1. However, its main features remain similar to those of the original BCG Tokyo 172-1 strain.

Conclusions: The changes in the genome may be caused by a variety of factors including the strain of Tokyo 172, vaccine processing, storage, and natural changes in the genome.

Keywords: BCG, BCG Tokyo, *Mycobacterium bovis*, next-generation sequencing

Introduction

Tuberculosis (TB) is a major public health problem. One-third of the global population is infected with *Mycobacterium tuberculosis*. Since 1988, Thailand has become one among 22 countries with a high TB incidence according to World Health Organization (WHO) records. The number of new TB patients in these 22 countries accounts for about 80 percent of patients worldwide. The country with the highest incidence is India, followed by China with more than 1 million new cases per year (in the year 2012). The WHO estimated that Thailand has around 80,000 new TB cases annually and the prevalence is 119 per 100,000 population, around 30 times higher than in some developed countries [1, 2].

Bacillus Calmette–Guérin (BCG) is a live strain of *M. bovis* developed by Calmette and Guerin for use as an attenuated vaccine to prevent TB and other

mycobacterial infections [3]. BCG vaccine has been used in the prevention of TB since the 1920s. It provides good tuberculosis protection during childhood. The major daughter strains of BCG used for vaccine production are BCG Danish, Glaxo, Pasteur, Moreau, Tokyo, and Russia. To prevent continuing genetic changes in the strains, the WHO recommended in 1987 that vaccine should not be prepared from any culture that had gone more than 12 passages starting from a defined freeze-dried seed lot [4]. The Thai Red Cross Society (TRCS) Queen Saovabha Memorial Institute (QSMI) brought Tokyo 172-1 BCG vaccine strains from Japan to produce the vaccine since 1988 until now. In 2009, the complete whole genomic sequence of *M. bovis* BCG Tokyo 172-1 determined by Seki et al. had a length of 4,371,711 bp and contained 4033 genes, including 3950 genes coding for proteins or coding DNA sequences (CDS) [5]. However, the complete whole genomic sequence of the Tokyo 172-1 BCG strain (*M. bovis* BCG TRCS strain) used in Thailand has never been determined. Bedwell et al. described a multiplex polymerase chain

Correspondence to: Wichit Thaveekarn, Queen Saovabha Memorial Institute, Thai Red Cross Society, Bangkok 10330, Thailand. E-mail: wichit.t@redcross.or.th

reaction (PCR) in which one region of difference (RD)16 product was present in BCG Russia, Japan, Connaught, Tice, Glaxo, Pasteur, and Birkhaug. Exceptionally, BCG Japan gave a 379 bp product (22 bp missing) and 401 bp product. The absence of RD16 BCG Moreau from Brazil was investigated [6]. Shibayama et al. identified to 2 BCG genotype subpopulations, types I and II, by multiplex PCR using RD16 region criteria. Type I had a 22-bp deletion in RD16, but type II had no such deletion [7]. Two genotypes and colonies have been found in common BCG Tokyo vaccines and the major population was type I [8]. In addition to RD16, the gene sequence for phenolic glycolipid (PGL) *ppsA* can distinguish the 2 types of BCG. Type II has a single base insertion and 2 single base substitutions. PGL plays critical roles in the killing activity of host macrophages [9, 10]. Here we report the complete whole genomic sequence and the genetic characteristics of the BCG TRCS strain by using next-generation sequencing technology.

Material and methods

Bacterial strain

BCG TRCS vaccine (lot No. FB03012), obtained from QSMI, was used for whole genomic sequencing. A working seed lot (BCG TRCS strain), which is used for production of BCG vaccine in Thailand, was characterized genetically.

Preparation of genomic DNA, library construction, and DNA sequencing

Genomic DNA from BCG-TRCS vaccine (lot No. FB03012) was extracted using a genomic extraction kit (QIAamp DNA Mini Kit, Qiagen). Construction of the library and sequencing were performed by the Biochemistry Department, Medicine Faculty, Chulalongkorn University, Thailand. The library was prepared using a NEBNext Ultra DNA Library Prep Kit for Illumina. Whole genome sequencing and assembly were analyzed using a HiSeq 2000 Illumina system (New England Biolabs).

PCR and DNA sequencing of the RD16 region (*Rv3405c*)

Genomic DNA from the working seed BCG (lot No. FB03012) was extracted using a genomic extraction kit (QIAamp DNA Mini Kit, Qiagen). PCR were performed with 2 primers RD16-F (5'-GGC TGG TGTT TCG TCA CTT C-3'), RD16-R (5'-ACA TTG GGAAAT CGC TGT TG-3') [8]. PCR products were cloned using a pTG19-T vector (Vivantis,

Malaysia) and then nucleotide sequences were determined by the 1st Base DNA sequencing service (Seri Kembangan, Malaysia).

Comparative genomic analysis

The complete genomic DNA of the BCG TRCS strain was submitted to GenBank and compared with the reference genomic DNA using the basic BLAST alignment tool (National Center for Biotechnology Information (NCBI) website at <https://blast.ncbi.nlm.nih.gov/Blast.cgi#1051534617>; U.S. National Institutes of Health). The genomic analysis tool was CLC Genomics Workbench 9 (CLC bio software, Qiagen). The complete genome of the BCG TRCS strain was compared to reference complete genome of *M. bovis* BCG Tokyo 172, GenBank: AP010918.

Results and discussion

Detection of 2 subpopulations by using RD16 region (*Rv3405c*) criteria

The working seed BCG TRCS strain was used as subject for conventional PCR amplification of the RD16 region with the primer set RD16-F and RD16-R. The PCR products showed 2 bands with different sizes (**Figure 1**).

RD16 alignment

Sequencing and alignment of the RD16 region in 2 different types of BCG were determined using CLC Genomics Workbench 9 (CLC bio software, Qiagen) and showed a 22 bp difference (**Figure 2**).

The results from sequencing and alignment of the RD16 region in 3 BCG strains are shown in **Figure 3**.

Comparison of BCG 172 Tokyo and BCG TRCS strain

The genome of the BCG TRCS strain is 4,371,707 bp. The genomic sequences of the BCG TRCS strain and BCG Tokyo 172 strain were found to be homologous. The results from identification of 4,076 genes indicated 3,963 CDS; 3 genes for rRNA, 45 genes for tRNA, and 62 pseudogenes; genome sequence details were deposited in GenBank (accession no. CP014566). Comparison of whole genome sequence between the BCG TRCS strain (CP014566) and BCG Tokyo 172 (AP010918) by next-generation sequencing analysis found 23 different points in amino acids (**Table 1**).

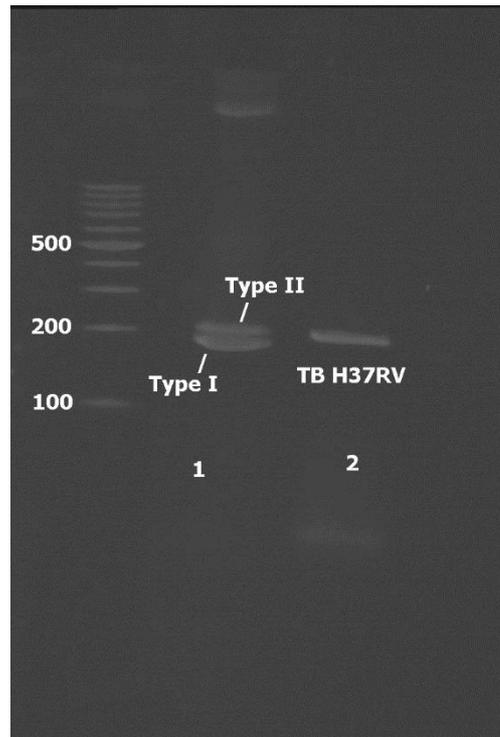


Figure 1. Electrophoresis of PCR products of the RD16 region in a 12.5% polyacrylamide gel. Lane 1: BCG TRCS strain working seed lot. BCG TRCS had 2 types of band, the upper is type II (with the 22-bp deletion) and the lower is type I (BCG TRCS is composed of type I and II subpopulations). Lane 2: TB H₃₇Rv was used as a positive control.

The far-left lane shows a marker ladder (VC 100bp DNA Ladder, Vivantis) in units of bp.

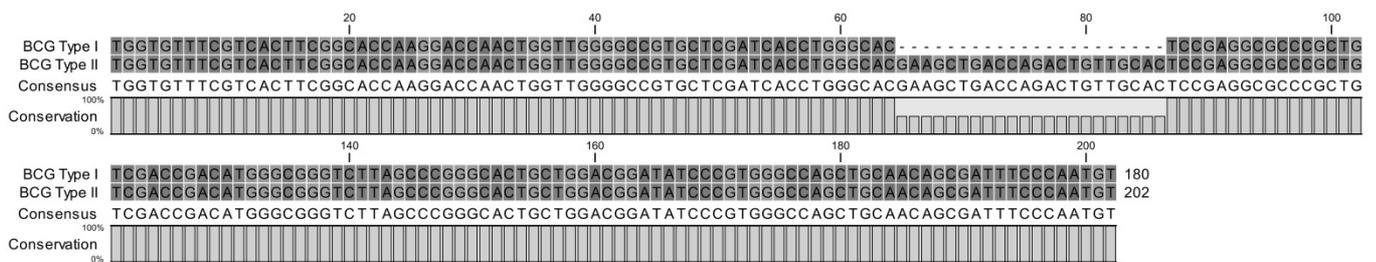


Figure 2. Sequencing and alignment of the RD16 region in the 2 different types of BCG using CLC Genomics Workbench 9 (CLC bio software, Qiagen).

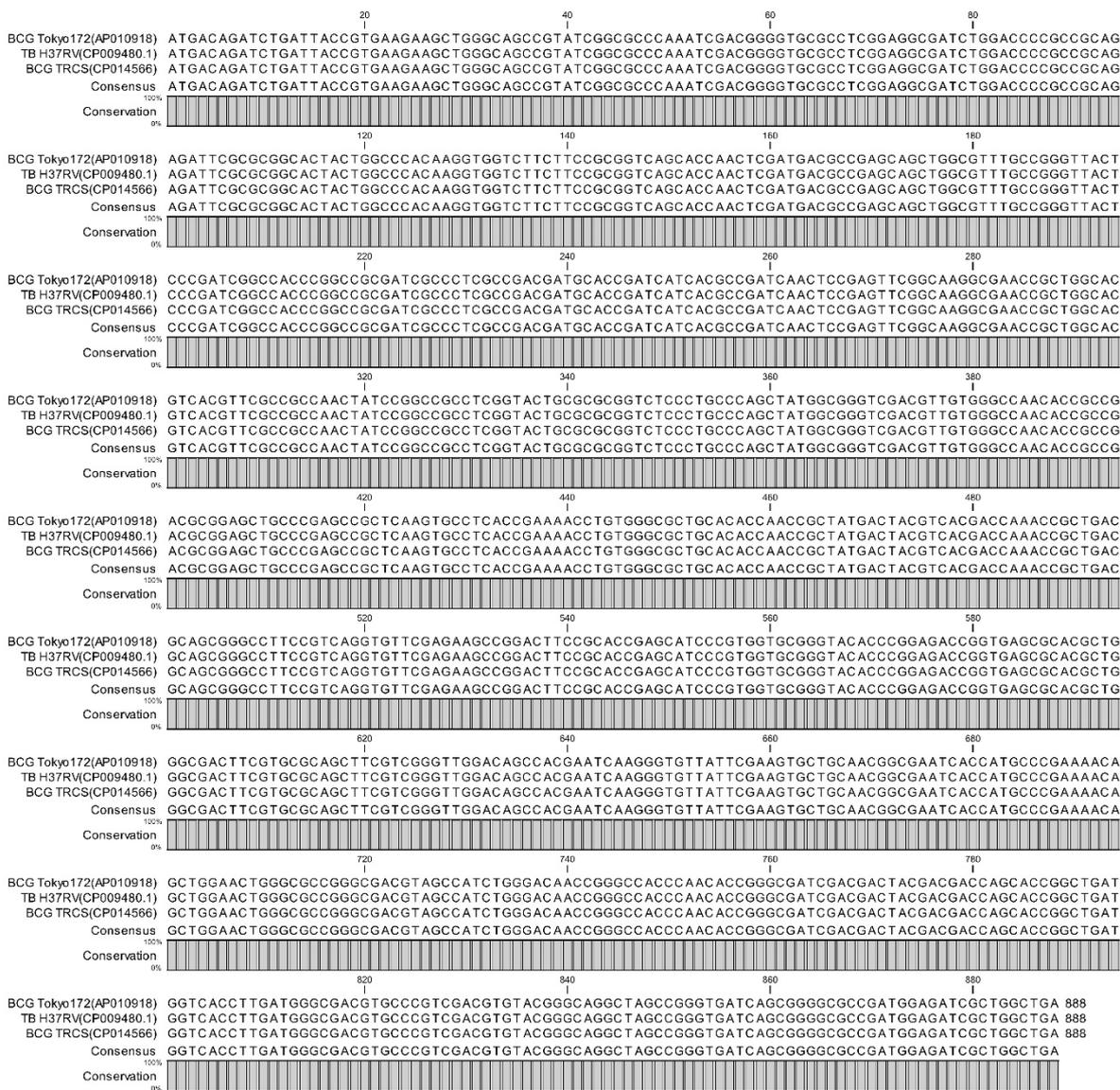


Figure 3. RD16 region alignment of BCG Tokyo172 (type I), TB *H₃₇R_v*, and BCG TRCS (type I) using CLC Genomics Workbench 9 (CLC bio software, Qiagen).

16S rRNA alignment

The DNA alignment of the 16S rRNA gene showed several different points between BCG TRCS and BCG Tokyo 172 strains (**Figure 4**).

Single nucleotide polymorphisms

Single nucleotide polymorphisms uniquely represent the BCG Tokyo species and those of the BCG TRCS strain remain the same as for the Japanese species (**Table 2**).

ppsA gene analysis

Figure 5 shows the results of *ppsA* gene alignment of BCG TRCS strain and BCG Tokyo

type II (AB665170). The nucleotides at the position 275, 379, and 2415 are different.

Variable-number tandem repeats

Variable-number tandem repeat (VNTR) sequences are valuable markers for genotyping bacterial species. In the present study, 7 regions of difference (RDs) at the VNTR locus, which have shown polymorphism in the *M. tuberculosis* complex were used to compare the identities of the BCG species. The number from repeats from 7 RDs in the VNTR locus of the BCG TRCS strain was shown to be the same as that of BCG Tokyo 172 (**Table 3**).

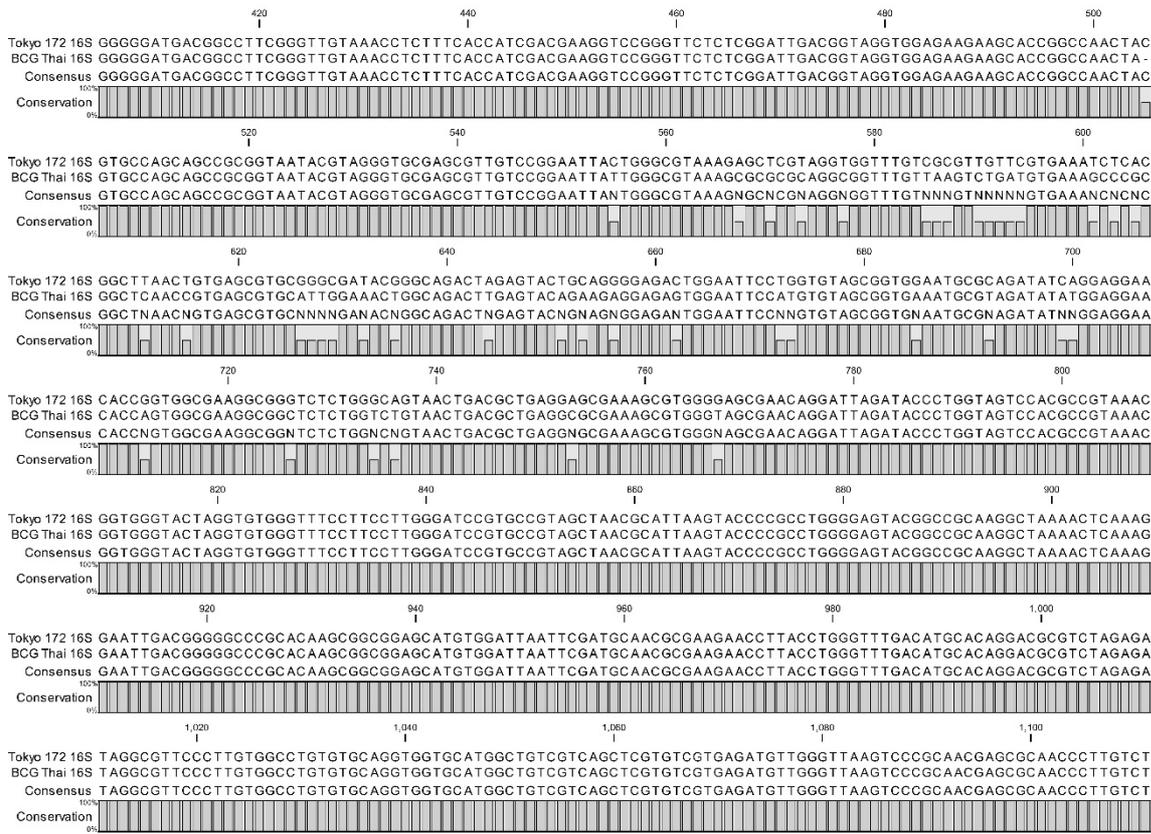


Figure 4. 16S rRNA sequence alignment analysis of the BCG TRCS strain (CP014566) and BCG Tokyo 172 (AP010918) using CLC Genomics Workbench 9 (CLC bio software, Qiagen).

Table 2. Comparison of single nucleotide polymorphisms between BCG TRCS strain (CP014566) and BCG Tokyo 172 (AP010918).

Gene	Position	BCG 172 Tokyo	
		Thai GenBank: CP014566	Japanese GenBank: AP010918
<i>pckA</i>	253,186	A	A
<i>JTY_0567</i>	644,562	T	T
<i>rpoC</i>	765,342	C	C
Intergenic	2,717,585	C	C
<i>JTY_3278</i>	3,606,131	G	G
<i>JTY_3735</i>	4,087,391	G	G

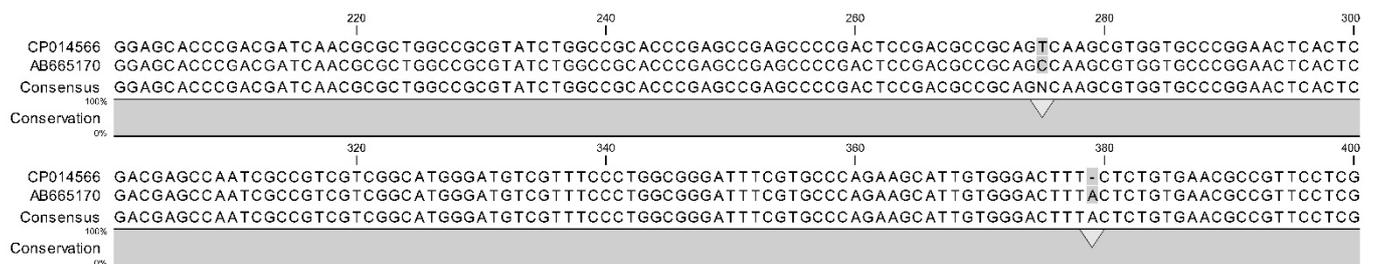


Figure 5. Analysis of *ppsA* gene alignment between the BCG TRCS strain (CP014566) and BCG Tokyo type II (AB665170) using CLC Genomics Workbench 9 (CLC bio software, Qiagen).

Table 3. Seven regions of difference (RDs) in the variable-number tandem repeat (VNTR) locus of the BCG TRCS strain.

RD in VNTR locus	Sequence from 5'→3'	Number of repeats
0557/intergenic (<i>Rv0487–Rv0488</i>) (ETR-C)	GTCGAGCCCGACGACGATGCAGAGCGCGCAGCGCGATGAGAA GGAGTTGGGCGGTTAG (58 bp)	3
0580/intergenic (<i>Rv0490–Rv0491</i>) (<i>senX3-regX3</i>)	TGCGCCGACGACGATGCAGAGCGTAGCGATGAGGTGGGGGCA CCACCCGCTTGCGGGGGAGAGTGGCGCTGATGACC (77 bp)	3
3155/intergenic (<i>Rv2847c–Rv2848c</i>)	CGACCCGCGGGCGCCCGGTCCCCGCGCTTGCGATCGCCACTGGC CCTGATGGTGG (54 bp)	4
3336/intergenic(<i>Rv2980–Rv2981c</i>)	TGCGCCGGGCGCGGGTTCGGCACCATCGGGCTAAGTGCCGA TCGCAAGCGCGGCGCT (177 bp= 59 bp×3)	6
3820/intergenic (<i>Rv3401–Rv3402c</i>)	CGATGCGGGCCGCGTAGCGGCCCGAGGAGGAGCCGGGCAATCC AGCCTGAGCCCGGTGA (59 bp)	3
4052 (<i>Rv3611</i>)	CCATCAGCCCCGTGGCGATCGCAAACCCCGCGCCTGGCGACAAT GCGGCCCGAAAACGGGCCGAGGAGGAGCCAGGCAATCACC CCAGAGCCGGGTGCAGCGGGTCGCCA (111 bp)	4
4155(<i>Rv3710:leuA</i>)	TCGCGAGCCCGCGCAGCCGGGCGAAGCGGGTTCGGCACGCGAT CGGACCCCGTGACGA (57 bp)	11

BCG vaccine was originally developed from one strain of *M. bovis*. Different methods of passage and storage have led to many substrains showing the phenotype. Such natural diversification, e.g. heterogeneity on duplication of the DU2 region in the BCG Danish substrains, commonly occurs in BCG substrains and can be increased by continuous passage without appropriate cloning because of natural mutations and selective pressure in culture media [11, 12]. BCG Tokyo 172 is one of 4 WHO vaccine reference strains [13, 14]. The Japanese BCG strain exhibits good protective efficacy with a low rate of unfavorable side effects [15]. Two types of variant strains, subpopulations I and II, are commonly found in the BCG Tokyo 172 vaccine preparation. They continue to coexist in subsequently produced lots with a decrease in the population of variant strain type II. BCG Tokyo 172 subpopulation types I and II are different, with a 22-bp deletion in the RD16 region. To our knowledge, there is no information on vaccine efficacy of the 2 different types [8].

The results of the RD16 region genotypes of the BCG TRCS strain from a working seed lot have shown a difference in the 2 PCR band product sizes, type I amplified 180 bp and type II 202 bp. PCR could identify

type I with a 22-bp deletion and type II with the complete sequence in **Figure 1** (lane 1). There were 22 bp differences between the 2 types (GAA GCT GAC CAG ACT GTT GCA C) as seen in **Figure 2**. These findings indicated that the BCG Tokyo 172-derived TRCS strain contained 2 types. PCR products of the RD16 BCG Tokyo type I subpopulation showed a characteristic 22-bp deletion, but type II was complete in this region. BCG Connaught and BCG Pasteur strain had a complete sequence in this region [8], while RD16 was missing in BCG Moreau [6, 8]. The TB H₃₇Rv PCR product in the RD16 region showed a band between BCG TRCS types I and II. By alignment of the RD16 region, we found that TB H₃₇Rv was similar to BCG Tokyo 172 type I and BCG TRCS type I. Therefore, TB H₃₇Rv also did not contain the 22 bp (**Figure 3**). A limitation of the present study is that we did not have BCG Tokyo strain, BCG Danish strain, or BCG Pasteur strains available to provide additional controls for PCR products.

Next-generation sequencing has indicated that the genome of BCG TRCS strain was 4,371,707 bp. DNA homology with BCG Tokyo 172 (AP010918) was 4,371,711 bp. Only 4 bp of DNA sequence and 23

4,371,711 bp. Only 4 bp of DNA sequence and 23 points of whole genome sequence were different between these BCG strains. VNTR sequences have emerged as valuable markers for the genotyping of several bacterial species, especially for genetically homogenous pathogens such as *M. tuberculosis* complexes [16]. Based upon the number of repeat polymorphisms within these tandemly arranged repetitive DNA sequences, many of these tandem loci display hypervariability, enabling their exploitation for strain typing in numerous bacterial species [17]. From **Table 3**, BCG TRCS strain matched BCG Tokyo 172 [5] showing 7 RDs retained their original sequence. PGL is antigenic and involved in host responses by acting as a cell wall component [8]. PGL can be found only in type I, but not in type II because of a *ppsA* gene mutation. BCG Tokyo type II (AB665170) had a single base insertion at nucleotide 379 (A base) and 2 single-base substitutions at nucleotides 275 (T→C) and 2415 (T→C). The length of *ppsA* gene CP014566 (BCG TRCS strain) was 5631 and AB665170 was 5632 (data not shown). The result in **Figure 5** shows no mutation in the *ppsA* gene of BCG TRCS strain CP014566, indicating type I.

The difference between BCG TRCS strain (CP014566) and BCG Tokyo 172 (AP010918) might result from genetic change in PE-polymorphic GC-rich-repetitive sequence (PGRS) family protein, hypothetical protein, putative oxidoreductase, 16S rRNA, putative resuscitation promoting factor, putative ATP-dependent helicase, and *sdhB* (**Table 1**). The greatest genetic change position was for PE-PGRS family protein. PE-PGRS is a large family of typical proteins of pathogenic mycobacteria whose members are characterized by an N-terminal PE domain followed by a large Gly-Ala repeat-rich C-terminal domain [18]. The genes of the PE-PGRS family proteins are most often clustered in a region of the genome, often as overlapping genes. The proline-glutamic acid (PE) domain is responsible for the cellular localization of these proteins on bacterial cells [19]. The hypothetical proteins found 3 positions shown in **Table 1** might be involved in virulence, detoxification, or adaptation; and proteins involved in intermediary metabolism and respiration [20]. The 16S rRNA gene is often used to study bacterial phylogeny and taxonomy. **Table 1** shows there were the differences in positions 500–767 observed on the 16S rRNA DNA alignment between genes of the BCG Tokyo 172 (AP010918) and BCG TRCS strains (CP014566).

The result of 16S rRNA DNA alignment indicated some differences (**Figure 4**). The present study used next-generation sequencing, in which Q score >30 and genome coverage = 766×. This suggests the probability of an incorrect base is 1 in 1,000. The 16S rRNA region between BCG Tokyo 172 (AP010918) and BCG TRCS strain (CP014566) were apparently different. Similar occurrences were found in these genes, PE_PGRS16 (1093960–1095282), *rpfE*, PE_PGRS43b, JTY_3396 and PE_PGRS57 (data not shown). We found that there are apparently many nucleotide differences between BCG-TRCS and BCG-Tokyo-172 strains in the 16S rRNA region. The degree of difference appears high given the 30-year history of the parent strain, and probable minimized passages of the seed lot. It is possible that these differences are the result of sequencing error. Sometimes new-generation sequencing can result in errors that depend on sample preparation or other mechanical factors. To be clear whether the result obtained is true or not, Sanger sequencing should be performed on the regions inconsistent with those of comparison. However, we were not able to perform Sanger sequencing or repeat the 16S rRNA DNA new-generation sequencing for want of more sample from this lot (No. FB03012). Use of more samples is warranted to confirm the sequencing.

The BCG Tokyo 172 preparation is mainly composed of the type I genotype, with known characteristics of this preparation, such as high viability and good heat stability [21]. The type I subpopulation shows a growth advantage over the type II subpopulation both on culture media and in mice. Although type II was still present in every preparation including the seed lot, the proportion of type II subpopulation decreases during passaging [22]. Type I produces PGL, but type II cannot because of the *ppsA* mutation. PGL deficient mutants show a phenotype with low virulence [23], and this might be a reason for type I always being the major component of the Tokyo 172 preparations [22]. The specimens from patients receiving BCG vaccination or BCG therapy for bladder cancer have only shown the genotype with the deletion in RD16 (type I) [24]. The present report does not explain why there are 2 variants of the BCG type in the vaccine. It is possible that some conditions of lot preparation, such as nutrient components and temperature in culture, might affect the proportions of certain subpopulations. Propagation of seed lots, subculturing to mass product, and freeze-

drying techniques might generate mutations and give rise to the accumulation of genetic variants within the substrain [11].

The BCG vaccine of BCG TRCS containing 2 types (I and II) from working seed lots found the proportion of type I to be more than type II in the PCR step. As the result of whole genome sequencing type I was found as the main population. Overall genetics of the BCG TRCS remain the same as BCG Tokyo 172 including the 7 RDs in *ppsA*, VNTR locus, and single nucleotide polymorphisms. Other genes appear to have been changed, and are apparently no longer the same as in BCG Tokyo 172, probably because of the storage environment, shelf life, preparation process, freeze drying, culture medium or natural mutation of the BCG Tokyo 172 species itself. Sequencing of the entire genome sequence of the BCG TRCS type II strain is planned for future work.

Author contributions

WT and S Payungporn made substantial contributions to the concept and design of the project, WT, S Payungporn, SS, and SK substantially contributed to the acquisition of data, and WT, S Payungporn, and S Petsong analyzed and interpreted it. WT and NP drafted the manuscript and all authors critically revised it. All authors approved the final version submitted for publication and take responsibility for the statements made in this article.

Acknowledgments

We thank Miss Pornpimol Premchaiporn, Head of BCG vaccine production department, Queen Saovabha Memorial Institute, Thai Red Cross Society for providing the BCG vaccines. We thank Mr. Nibondh Udomsantisuk, Department of Microbiology, Faculty of Medicine, Chulalongkorn University for providing the TB H₃₇Rv (*M. tuberculosis*) control sample. This work was supported by a grant (for project No. QSMI5902) from the scientific committee of the Queen Saovabha Memorial Institute, Thai Red Cross Society, Bangkok, Thailand.

Conflict of interest statement

The authors declare that there is no conflict of interest in this research.

References

1. WHO report. Global tuberculosis control: surveillance,

- planning, financing. Geneva: Switzerland, 2007.
2. Ministry of Public Health, Thailand. National Tuberculosis Control Program Guidelines. Bangkok: Bureau of Tuberculosis, Ministry of Public Health, Thailand; 2013.
3. Olan P, Auchara T, Usa T, editors. Vaccines: 9th International Congress of Tropical Pediatrics. Bangkok, Thailand, October 18–20, 2011. Bangkok: Noppachai Printing; 2011.
4. WHO Expert Committee on Biological Standardization (36th Report). World Health Organization Technical Report Series 745. Geneva: WHO; 1987.
5. Seki M, Honda I, Fujita I, Yano I, Yamamoto S, Koyama A. Whole genome sequence analysis of *Mycobacterium bovis* bacillus Calmette–Guérin (BCG) Tokyo 172: A comparative study of BCG vaccine substrains. *Vaccine*. 2009; 27:1710-6.
6. Bedwell J, Kairo SK, Behr MA, Bygraves JA. [Identification of substrains of BCG vaccine using multiplex PCR.](#) *Vaccine*. 2001; 19:2146-51.
7. Shibayama K, Mochida K, Yagi T, Mori S, Arakawa Y, Yamamoto S. [Quantification of two variant strains contained in freeze-dried Japanese BCG vaccine preparation by real-time PCR.](#) *Biologicals*. 2007; 35: 139-43.
8. Naka T, Maeda S, Niki M, Ohara N, Yamamoto S, Yano I, et al. Lipid phenotype of two distinct subpopulations of *Mycobacterium bovis* Bacillus Calmette–Guérin Tokyo 172 substrain. *J Biol Chem*. 2011; 286:44153-61.
9. [Guenin-Mace L, Simeone R, Demangel C. Lipids of pathogenic Mycobacteria: contributions to virulence and host immune suppression.](#) *Transbound Emerg Dis*. 2009; 56:255-68.
10. Camacho LR, Constant P, Raynaud C, Laneelle MA, Triccas JA, Gicquel B, et al. Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. *J Biol Chem*. 2001; 276:19845-54.
11. Wada T, Maruyama F, Iwamoto T, Maeda S, Yamamoto T, Nakagawa I, et al. Deep sequencing analysis of the heterogeneity of seed and commercial lots of the bacillus Calmette–Guérin (BCG) tuberculosis vaccine substrain Tokyo-172. *Sci Rep*. 2015; 5:17827. doi:10.1038/srep17827.
12. Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, et al. [Genome plasticity of BCG and impact on vaccine efficacy.](#) *Proc Natl Acad Sci U S A*. 2007; 104:5596-601.
13. Ho MM, Markey K, Rigsby P, Hockley J, Corbel MJ.

- Report of an international collaborative study to establish the first WHO reference reagents for BCG vaccines of three different sub-strains. *Vaccine*. 2011; 29:512-8.
14. Dagg B, Hockley J, Rigsby P, Ho MM. [The establishment of sub-strain specific WHO reference reagents for BCG vaccine.](#) *Vaccine*. 2014; 32:6390-5.
 15. Milstien JB, Gibson JJ. Quality control of BCG vaccine by WHO: a review of factors that may influence vaccine effectiveness and safety. *Bull World Health Organ*. 1990; 68:93-108.
 16. Frothingham R, Meeker-O'Connell WA. [Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats.](#) *Microbiology*. 1998; 144:1189-96.
 17. van Belkum A, Scherer S, van Alphen L, Verbrugh H. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol Mol Biol Rev*. 1998; 62:275-93.
 18. Palucci I, Camassa S, Cascioferro A, Sali M, Anoosheh S, Zumbo A, et al. PE_PGRS33 contributes to *Mycobacterium tuberculosis* entry in macrophages through interaction with TLR2. *PLoS One*. 2016; 11:e0150800. doi:10.1371/journal.pone.0150800.
 19. Meena LS. An overview to understand the role of PE_PGRS family proteins in *Mycobacterium tuberculosis* H₃₇Rv and their potential as new drug targets. *Biotechnol Appl Biochem*. 2015; 62:145-53.
 20. Rodriguez-Alvarez M, Palomec-Nava ID, Mendoza-Hernandez G, Lopez-Vidal Y. [The secretome of a recombinant BCG substrain reveals differences in hypothetical proteins.](#) *Vaccine*. 2010; 28:3997-4001.
 21. Gheorghiu M, Lagrange PH. Viability, heat stability and immunogenicity of four BCG vaccines prepared from four different BCG strains. *Ann Immunol (Paris)*. 1983; 134:125-47.
 22. Honda I, Seki M, Ikeda N, Yamamoto S, Yano I, Koyama A, et al. Identification of two subpopulations of Bacillus Calmette-Guérin (BCG) Tokyo172 substrain with different RD16 regions. *Vaccine*. 2006; 24: 4969-74.
 23. Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, et al. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature*. 2004; 431:84-7.
 24. Seki M, Sato A, Honda I, Yamazaki T, Yano I, Koyama A, et al. Modified multiplex PCR for identification of Bacillus Calmette-Guérin substrain Tokyo among clinical isolates. *Vaccine*. 2005; 23:3099-102.