## Letter to the Editor

# The Epidemiological Characteristics of Beijing Lineage *Mycobacterium tuberculosis* from a National Referral Center in China*

LI Xiao Ying[1,†], LI Ying[7,†], ZHANG Yao[8,†], KANG Wan Li[3], ZHAO Li Ping[4,6], DING Peng Ju[4], DAI Wen Tao[5], HUANG Hai Rong[4,6,#], HUANG Yan Feng[2,#], and LI Wei Min[4,6,#]

**Our study was to investigate the epidemiological characteristics of *M.tuberculosis* from a national tuberculosis referral center in China. All strains isolated from TB patients, were genotyped by the RD105 deletion, 8 and 51 SNP loci and VNTR. The high differentiation SNPs of modern Beijing strains were analyzed for protein function and structure. 413 *M. tuberculosis* were included. Of 379 Beijing lineage *M. tuberculosis*, 'modern' and 'ancient' strains respectively represented 85.5% (324/379) and 14.5% (55/379). Rv2494 (V48A) and Rv0245 (S103F) were confirmed as high differentiation SNPs associated with modern strains. In a word, Modern Beijing lineage *M.tuberculosis* was dominant and the structural models suggested that modern sub-lineage may more easily survive in 'extreme' host condition.**

Tuberculosis (TB) still represents a serious public health problem not only in china, but also in the world[1]. Beijing genotype *Mycobacterium tuberculosis* (*M. tuberculosis*) was prevailing not only in china, but also in the global. It was reported Beijing strains associated with drug resistance and greater virulence. Our study shown that Beijing strains accounted for 64.9% of TB strains isolated in China[2]. Recently, the Beijing genotype was renamed as Beijing lineage *M. tuberculosis* and was determinate from other lineages by RD105. Based on phylogenic theory, Beijing lineage *M. tuberculosis* could be divided into modern and ancient two sub-lineages by RD181. Luo et al.[3] and van Soolingen et al.[4] hoped to respectively use eight SNP loci and 51 SNP loci to classify more modern and ancient sub-lineages of Beijing lineage *M. tuberculosis*. Yang et al.[5] first reported that Beijing lineage *M. tuberculosis* was not associated with some clinical symptom, such as treatment history and drug resistance pattern, so on, but was stronger transmission capacity than that of no Beijing-lineage based on the theory of clustering. In addition, he reported again transmission capacity of modern Beijing lineage *M. tuberculosis* was stronger than that of ancient by the same theory. We speculated that this observation was related to the pathogenicity and epidemiological characteristics of Beijing lineage *M. tuberculosis*. Microevolution refers to changes in allele frequencies over time within a TB strain population, which can be used to investigate high-differentiation gene loci in the entire *Mycobacterium* genome and the molecular mechanism of adaptation to an extreme environment. Jiang et al.[6] constructed homology modeling software to predict a protein's structure and function. In this study, we sought to evaluate

*doi: 10.3967/bes2015.077*

*This study was supported by the National Natural Science Foundation of China (No. 81273144), Beijing Natural Science Foundation Program and Scientific Research Key Program of Beijing Municipal Commission of Education (KZ201510025024) and Beijing Municipal Administration of Hospitals Clinical Medicine Development of Special Funding Support (ZYLX201304).

1. Tianjin Children's Hospital, Tianjin 300074, China; 2. Division of Infectious Disease and Gastroenterology, Children's Hospital, Chongqing Medical University, Chongqing 400014, China; 3. Department of Epidemiology, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing Chest Hospital, Capital Medical University, Beijing 101149, China; 4. National Tuberculosis Clinical Lab of China, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing Chest Hospital, Capital Medical University, Beijing 101149, China; 5. Key Laboratory of Protein and Peptide Pharmaceuticals, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China; 6. Beijing Key Laboratory for Drug-Resistant Tuberculosis Research, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing Chest Hospital, Capital Medical University, Beijing 101149, China; 7. Zhejiang Provincial Key Laboratory for Technology & Application of Model Organisms, School of Laboratory Medicine and Life Science, Wenzhou Medical University, Wenzhou 325035, Zhejiang, China; 8. Clinical Epidemiology Unit, Peking Union Medical College, Beijing 101149, China

the different SNP loci between modern and ancient sub-lineages in order to search for genes related to the recent transmission and host adaptation of modern Beijing lineage *M. tuberculosis*, and analyze the prevalence of Beijing lineage *M. tuberculosis* and the clinical correlation between its modern and ancient sub-lineages. In addition, we screened several strong differentiation SNP loci, whose corresponding genes may be associated with recent transmission and host adaptation of modern Beijing strains.

All patients diagnosed with culture-positive TB from January 1, 2013 to March 31, 2013 were included from Beijing Chest Hospital. The study was approved by the Ethics Committee of Beijing Chest Hospital affiliated to Capital Medical University. Informed consent was obtained from all participants. Sputum smear using Ziehl-Neelsen staining and cultures on Lowenstein-Jensen medium were performed on samples collected from all participants[8]. *Mycobacterium* genomic DNA was extracted using a standard method. Beijing lineage *M. tuberculosis* was identified using the deletion-targeted multiplex PCR (DTM-PCR) method. The 51 SNP loci recommended by van Soolingen et al.[4] and the 8 SNP loci recommended by Luo et al.[3] were used to analyze Beijing sub-lineage *M. tuberculosis*. The SNP loci were detected using the MALDI-TOF method (Sequenom Inc. USA). The standard VNTR-15 and another three hypervariable loci (VNTR3820, VNTR4120, and VNTR3232) were used to genotyped the Beijing strains[3]. H37Rv was used as the standard strain. In this study, there were 421 *Mycobacterium* strains, including eight *No-tuberculosis Mycobacterium.* However, 413 *M. tuberculosis* were implemented the next analysis. In addition, there were 389 strains with the RD105 deletion; thus, Beijing lineage *M. tuberculosis* accounted for 94.2% (389/413) of the strains in this study.

Compared to cases managed in TB dispensaries, those in TB-specialized hospitals were generally more severe. Few studies have evaluated the epidemiological characteristics of cases from TB-specialized hospitals of China. Our study found that the proportion of Beijing lineage *M. tuberculosis* in our hospital was 94.2% (389/413), which was also higher than the 64.9% (265/408) ($P$<0.05) (data not shown) as national averages[2]. These results suggest that it was interested to study *M. tuberculosis* from TB specialized hospitals.

The sets of 8 and 51 SNP loci were analyzed using Network v.4.6.1.1 (Fluxus Technology Ltd, Germany). Data from the standard VNTR-15 loci were used to construct a minimum spanning tree (MST) using BioNumerics v.7.1 (Applied Maths, Belgium). The VNTR-18 was composed of the standard VNTR-15 and three hyper-variable VNTR loci. Every *M. tuberculosis* isolate corresponded to a single TB case. The TB patients were classified based on radiology results (cavity *vs.* non-cavity), treatment history (treatment naïve *vs.* re-treat), drug resistance pattern and clustering rate of the infecting strain. The number of modern Beijing lineage *M. tuberculosis* cases was determined on the basis of specific RD and SNP patterns. The association between different clinical characteristics and modern Beijing sub-lineage *M. tuberculosis* were evaluated using $\chi^2$ test. Data were analyzed using SPSS version17.0 (SPSS Inc., Chicago, IL, USA). All hypothesis testing was 2-sided, with α level of 0.05.

379 (97.4%; 379/389) Beijing strains had enough DNA for VNTR and SNP genotyping, which were grouped. The eight sub-lineages determined by eight SNP loci and the diversity were shown that 'Modern' Beijing lineages including Bmyc10, Bmyc16, and Bmyc210 was comprised 85.5% (324/379) of the samples, while 'ancient' Beijing lineages including Bmyc2, Bmyc4, Bmyc6, Bmyc25, and Bmyc26 was comprised 14.5% (55/379) of the samples (Table 1). Additionally, 51 SNP loci were used to classify the 379 strains into three groups. Except for the standard strain, H37Rv, 55 Beijing strains were in group 1, which was congruent with the group of ancient sub-lineages defined by the eight SNP loci. Group 2 and group 3 completely overlap with modern Beijing lineage *M. tuberculosis* defined by the eight SNPs.

To further study the evolution and diversity of each sub-lineage, we constructed a composite tree, in which the minimal spanning tree (MST) of each sub-lineage was combined with a SNP phylogeny drawing. Bmyc4 and Bmyc6 sub-lineages were excluded due to small sample size (Figure 1). The composite tree suggested that the modern Beijing strains were more recent expansion than those ancient ones. There were no significant associations between the ancient and modern sub-lineages and the clinical characteristics of patients, including treatment history, radiology, and drug resistance patterns.

Except for six non-coding SNP loci, the non-synonymous SNPs of the remaining 45 SNP loci were evaluated by PROVEAN[8]. The cut-off value for

a deleterious gene was set at <-2.5. Based on the screening scoring, the protein functions of the two interesting genes, Rv0245 and Rv2494, were predicted. Flavin reductase (PDB ID 1RZ1, (see http://www.rcsb.org/pdb/home/home.do) was used as the structure template for the target sequence of the *M. tuberculosis* Rv0245 gene (residues 1-162; NCBI accession: NP_214759; (see http://www.ncbi.nih.gov/). The toxin-antitoxin complex (PDB ID 3H87, (see http://www.rcsb.org/pdb/home/home.do) was used as the structure template for the target sequence of the *M. tuberculosis* Rv2494 gene (residues 1-141; NCBI accession: NP_217010; see http://www.ncbi.nih.gov/).

Of 17 coding SNP loci, 14 non-synonymous SNP loci that showed polymorphisms between the ancient and modern types of Beijing lineage *M. tuberculosis* (excluding three non-coding loci) were evaluated by the PROVEAN scoring. Five high differentiation non-synonymous SNP loci were identified, including Rv0245 (S103F), Rv1163 (P179R), Rv2248 (V199G), Rv2494 (V48A), and Rv3283 (E276K). The protein functions of the wild and mutant types of Rv0245 (S103F) and Rv2494 (V48A)

**Table 1.** The Sublineage Distribution of the 379 Strains of Beijing Lineage *M. tuberculosis*

| Subtype of Beijing Strain | SNPs | | | | | | | | Modern/ Ancient | Number of Strains (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| | Ogt (37) | LigD (580) | adhE2 (124) | UvrD1 (462) | mutT2 (58) | recX (59) | mutT4 (48) | recR (44) | | |
| Bmyc2 | W | M | W | W | W | W | W | W | ancient | 21 (5.5%) |
| Bmyc4 | W | W | W | W | W | W | W | M | ancient | 8 (2.1%) |
| Bmyc6 | W | W | W | W | W | W | M | M | ancient | 4 (1.1%) |
| Bmyc25 | M | W | W | W | W | W | M | M | ancient | 20 (5.3%) |
| Bmyc26 | W | W | W | W | W | M | M | M | ancient | 2 (0.5%) |
| Bmyc10 | W | W | W | W | M | M | M | M | modern | 210 (55.4) |
| Bmyc16 | M | M | M | M | M | W | W | W | modern | 13 (3.4%) |
| Bmyc21 | W | W | M | W | M | M | M | M | modern | 101 (26.6%) |

***Note.*** W: wild-type; M: mutant.

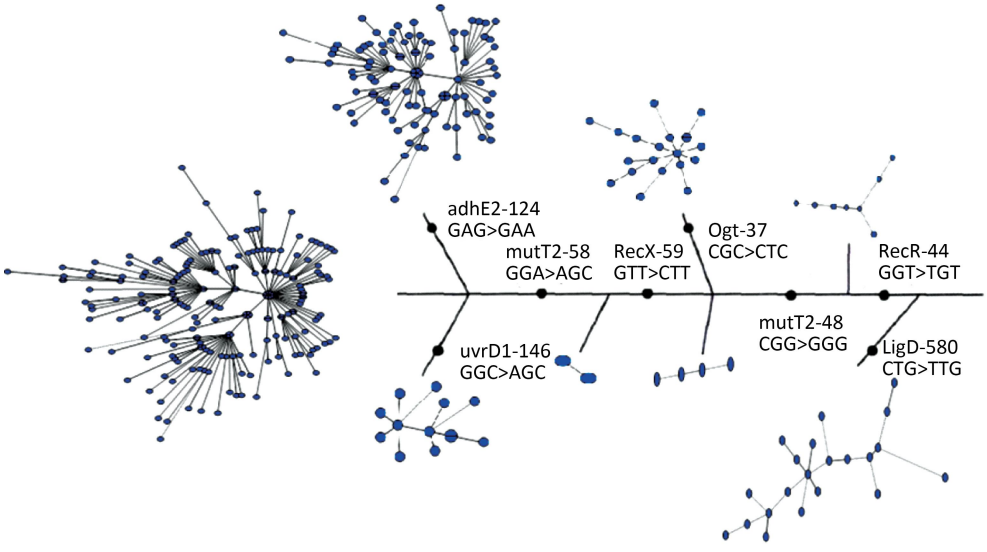


**Figure 1.** Composite phylogenetic tree based on eight SNP loci and 15 VNTR loci. The terminus of each branch is the Minimal Spanning Tree (MST) of each sub-lineage strains based on VNTR-15. Excluding the three sub-lineages, of which the strains were little, each MST of the left six sub-lineages was respectively represented as a star-like and branch-like network. In fact, all MSTs of modern Beijing strains were shown star-like, which indicated the homogenous population and more recent clonal expansion.

were predicted (Figure 2A, B). Rv2494 is a toxin protein component of vapBC38 in the toxin-antitoxin system. The protein function prediction suggested that the V48A mutation may significantly reduce the binding capacity of the pocket region and the efficiency of Rv2494 under physiological conditions. A study referred to toxin-antitoxin[9] suggested the activity change of the protein would affect the *M. tuberculosis* status, such as growth or dormancy. Rv0245 is an FAD- or FMN-dependent oxidoreductase.

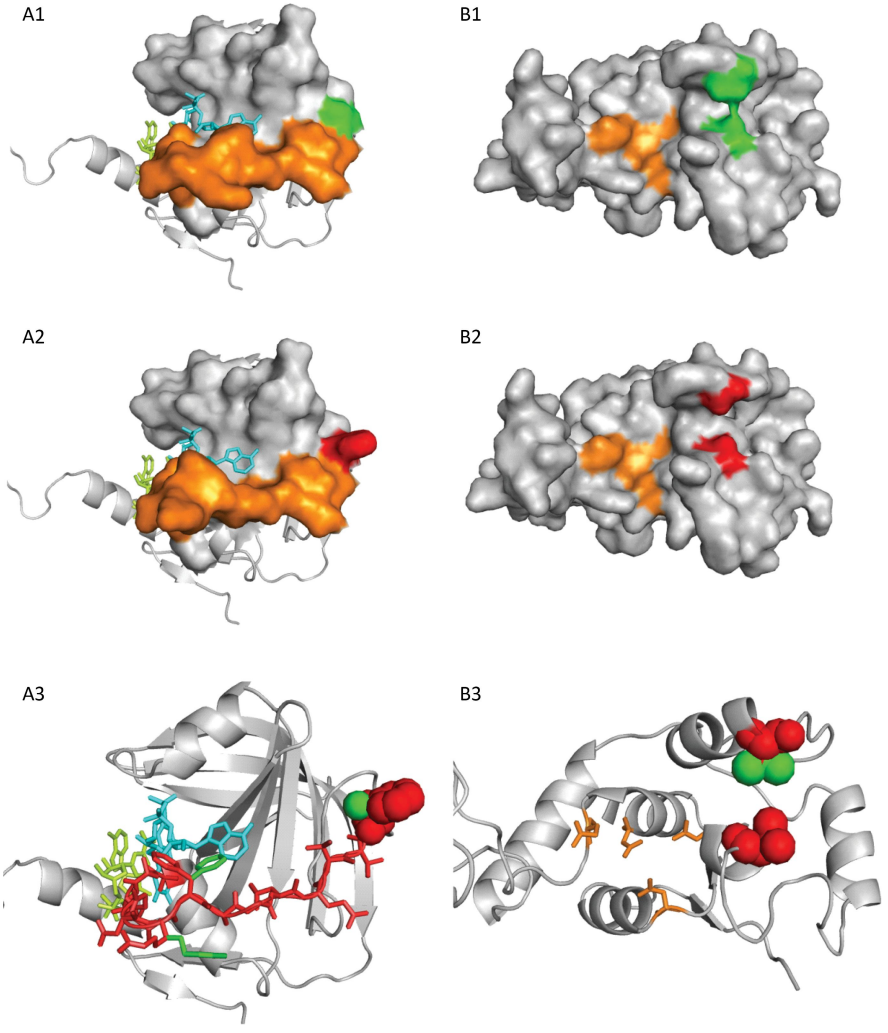


**Figure 2.** Three-dimensional structure models of the Rv0245 and Rv2494 genes of *M.tuberculosis*. The structure models A1 and B1 are the wild-type structures of Rv0245 and Rv2494, respectively. The structure models A2 and B2 are the homologous structures of Rv0245 (with S103F substitution) and Rv2494 (with S103F substitution), respectively. The structure model A3 is the overlapping structure between A1 and A2, and the structure model B3 shows the overlap between B1 and B2. The green color in all of the images indicates the wild-type regions, and the red color indicates mutant regions. In Figure A, the small cyan color molecules are FAD, and the small yellow colored molecules are NAD. The A1 and A2 surfaces show the pocket region, and the orange region is a loop region from A89 to V102 that constitutes the pocket. In Figure B, the orange region is a pocket with negative charges (D5, Q99, D102, and D120). Green represents the wild-type V48, and red represents the mutant type A48. The surfaces of B1 and B2 are shown, while B3 uses ribbons to represent the structure model, sticks to represent the pocket region, and spheres to represent mutation sites. Pymol software (www.pymol.org/) was used to draw the structure models.

The structural model showed that S103F was close to the $Ala^{89}$-$Val^{102}$ loop region (Figure 2A3) and this mutation, which changed the polar Ser to a non-polar Phe, could impact the conformation of this loop, thus affect the recognition and binding of FAD or FMN and increase Rv0245 catalytic activity. Another study[10] reported that FAD- or FMN-dependent oxidoreductases play important roles in lipid (such as cholesterol) metabolism in TB. Based on the functional prediction of the above two genes, we speculate that Rv0245 (S103F) and Rv2494 (V48A) polymorphisms in the modern sub-lineage bear a close relationship with the dormant status of *M. tuberculosis*. In a word, the reason that we identified high differentiation gene loci and predicted their function was to analysis underlying mechanisms of transmission capacity of modern Beijing strains.

In summary, we found the character of the genetic population construction of *M. tuberculosis* from Chinese largest TB specialized-hospital. Through analysis the high differentiation SNP loci of modern Beijing strains, we found that the recent transmission of Beijing strains was associated with its favor dormant.

†These authors contributed equally to this work.

#Correspondence should be addressed to LI Wei Min, Tel: 86-10-89509359, E-mail: lwm_18@aliyun.com; HUANG Yan Feng, Tel: 86-10- 89509159, E-mail: ymhykq@126.com; HUANG Hai Rong, Tel: 86-10-89509159, E-mail: hairong. huangcn@gmail.com

Biographical notes of the first authors: LI Xiao Ying, female, born in 1987, MS degree, a clinical Doctor; LI Ying, female, born in 1988, MS degree, a PhD student; YAO Zhang, male, born in 1977, MS and PhD, a clinical Doctor.

Received: January 8, 2015;
Accepted: June 15, 2015

## REFERENCES

1. World Health Organization. Global tuberculosis control: WHO report 2011. Geneva: World Health Organization, 2011.
2. Li WM, Wang SM, Li CY, et al. Molecular epidemiology of *Mycobacterium tuberculosis* in China: a nationwide random survey in 2000. Int J Tuberc Lung Dis, 2005; 9, 1314-9.
3. Luo T, Yang C, Gagneux S, et al. 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, 2012; 50, 633-9.
4. Schürch AC, Kremer K, Warren RM, et al. Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the *Mycobacterium tuberculosis* Beijing genotype. Infect Genet Evol, 2011; 11, 587-97.
5. Yang C, Luo T, Sun G, et al. *Mycobacterium tuberculosis* Beijing strains favor transmission but not drug resistance in China. Clin Infect Dis, 2011; 55, 1179-87.
6. Miao ZC, Cao Y, Jiang TJ. RASP: rapid modeling of protein side chain conformations. Bioinformatics, 2011; 27, 3117-22.
7. Chinese Medicine Association. The rule of clinical technique operation-Tuberculosis Branch. Beijing, PR China: People's Military Medical Press, 2003; 36-44.
8. Choi Y, Sims GE, Murphy S, et al.Predicting the functional effect of Amino acid substitutions and Indels. PLoS One, 2012; 7, e46688.
9. Ramage HR, Connolly LE, Cox JS. Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. PLoS Genet, 2009; 5, e1000767.
10. Dresen C, Lin LY, DAngelo I, et al. A flavin-dependent monooxygenase from *Mycobacterium tuberculosis* involved in cholesterol catabolism. J Biol Chem, 2010; 285, 22264-75.