Int J Clin Exp Pathol 2016;9(2):1533-1544
www.ijcep.com /ISSN:1936-2625/IJCEP0018827

Original Article

Characterization of *Mycobacterium tuberculosis* isolates from Shijiazhuang, China: genotypes and drug resistance

Yanan Li[1], Haican Liu[3], Huixia Gao[1], Yuling Wang[2], Zhi Zhang[1], Haibin Wang[2], Jinfeng Cao[2], Shumin Zhang[2], Xiaocui Wu[3], Qing Sun[3], Yuzhen Liu[1], Jianhua Lu[1], Erhei Dai[1], Kanglin Wan[3]

[1]Department of Laboratory Medicine, [2]Department of Tuberculosis, The Fifth Hospital of Shijiazhuang, Hebei Medical University, Shijiazhuang, Hebei, China; [3]State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

Received October 30, 2015; Accepted December 26, 2015; Epub February 1, 2016; Published February 15, 2016

Abstract: To understand the genetic diversity and phenotypic resistance of *Mycobacterium tuberculosis* (MTB) strains isolated in Shijiazhuang. Spoligotyping and MIRU-VNTR were performed to genotype MTB isolates and drug susceptibility testing (DST) was used to obtain phenotypic resistance. The genotyping result with drug resistance profiles and clinical characteristics was assessed. A total of 422 strains were collected from Shijiazhuang. The predominant lineage was the Beijing family (91.0%). The VNTR-$15_{standard}$ divided 69 strains into 27 clusters (2-7 strains per cluster) and 353 individual types. The clustering rate was 9.95% and HGDI value reached 0.999. In contrast, the VNTR-15_{China} with a clustering rate of 40.3% and a HGDI score of 0.989, grouped 229 strains into 59 clusters (2-32 strains per cluster) and 193 unique genotypes. MTB drug resistance rate in retreatment patients was statistically higher than that of initial treatment panels. However, no statistical difference of drug resistance rate was observed between the Beijing and the non-Beijing genotypes. The Beijing family is the predominant genotype in Shijiazhuang. Anti-TB drug treatment is associated with MTB drug resistance. VNTR-$15_{standard}$ has a higher discriminatory power than VNTR-15_{China}. Combination of spoligotyping and MIRU-VNTR is suitable for understanding the molecular epidemiology of MTB strains in Shijiazhuang.

Keywords: *Mycobacterium tuberculosis*, spoligotyping, vntr, drug resistance

Introduction

Tuberculosis (TB) remains one of the major threats to global public health, especially in developing countries. According to the report from the World Health Organization (WHO) [1], it was estimated that one third of the world population has ever been infected with the *Mycobacterium tuberculosis* (MTB). There was approximately 9.0 million new cases and 1.5 million died from this infectious disease each year. China, one of the highest TB burden countries in the world, held about 11% of global tuberculosis incidence with 1.3 million active TB patients and 0.98 million new cases in 2013. An estimated 5.7% of new infections and 26% of retreatment patients would develop into multi-drug resistant TB (MDR-TB) cases [1]. With the large infection group and the severe epidemic of MDR-TB, TB is still a major public health problem that China is facing and need to deal with.

Based on the advances in molecular biotechniques, some genotyping methods, such as spoligotyping and MIRU-VNTR, have developed as important tools to investigate the molecular diversity of TB strains. Spoligotyping is a rapid and easy-to-do genotyping method that has long been considered the gold standard for identification of the Beijing family MTB strains based on the absence of spacers 1-34 in the highly polymorphic direct repeat (DR) region of the MTB genome [2-4]. Due to the fairly low discriminatory power of spoligotyping, MIRU-VNTR was used as a complement to fully understand

the genetic diversity of MTB strains [5]. It can determine the different numbers of mycobacterial interspersed repetitive units with disparate VNTR loci [6]. Fast detection speed, high throughput and digitized records make it a popular molecular typing method all over the world [7-10].

Shijiazhuang, the provincial capital of Hebei, is located in the Taihang Mountains and the North China Plain, only 273 kilometers away from Beijing, which makes it a key transportation hub of China. It has a population of over 10 million people. Based on the 2000 National TB Epidemiology Survey in China, there were about 200,000 pulmonary TB patients in Hebei [11]. However, little local information about the molecular epidemiology of TB in Shijiazhuang is available. In this study, spoligotyping and MIRU-VNTR were performed to genotype MTB strains and DST was also conducted. In addition, molecular genotype with phenotypic drug resistance and clinical characteristics was assessed.

Materials and methods

Ethics statement

This study was approved by the Ethics Committee of the Fifth Hospital of Shijiazhuang and written informed consent was obtained from all participants.

Study population

A total of 422 isolates were collected for this study. The median age of the patients was 33 with 276 males. Patients aged from 16 to 35 years old accounted for 50.7% of the total patient population. Farmers were the largest infection population, followed by students and the unemployed. In addition, 131 (31.0%) had smoking history, 75 (17.8%) had drinking history, 103 (24.4%) were retreated cases, 47 (11.1%) were complicated with diabetes mellitus (DM), 61 (14.5%) had close contacts with patients suffering from TB and 34 (8.1%) had TB family history, which means at least one of his immediate family members were suffering from TB or had ever been infected with MTB strains (**Table 1**).

Bacterial strains

A total of 422 MTB strains isolated from 422 pulmonary TB patients from various regions of Shijiazhuang who had been treated at the Fifth Hospital of Shijiazhuang (The Infectious Disease Hospital of Shijiazhuang) during 2014, were enrolled in this study. The isolates were subcultured on Lowenstein-Jensen (L-J) medium for 4 weeks at 37°C. The clinical information was also collected in format of questionnaire. MTB H37Rv was used as the reference strain. Genomic DNA was extracted using Mericon™ DNA Bacteria Kit (Qiagen, USA). One loop of MTB colonies was resuspended in 200 μl fast lysis buffer by brief and vigorous vortex, and then incubated at 100°C for 10 minutes. After centrifugation at 12,000 rpm for 5 minutes, the supernatant containing DNA was collected and stored at -20°C for further use.

Drug susceptibility testing

For all the strains, DST was performed using the L-J proportion method, recommended by WHO, at Clinical Laboratory of the Fifth Hospital of Shijiazhuang, China [12]. The drug concentrations were 0.2 μg/ml for isoniazid (INH), 40 μg/ml for rifampicin (RIF), 4 μg/ml for streptomycin (SM), 2 μg/ml for ethambutol (EMB), 2 μg/ml for ofloxacin (OFX) and 30 μg/ml for kanamycin (KAN). H37Rv strain was used as quality control and was tested with each batch of DST. The results were read at the end of 28 days after inoculation. When the growth rate was more than 1% compared to the control, the strain was declared resistant to the specific drug. Strains resistant to at least RIF and INH were defined as MDR-TB [13].

Genotyping

Spoligotyping was performed as previously described by Kamerbeek et al [3]. To identify an optimal set of MIRU-VNTR loci for genotyping MTB isolates in Shijiazhuang, 19 loci were selected for analysis, including 11 common loci (ETRA, ETRC, ETRD, ETRE, MIRU10, MIRU16, MIRU26, MIRU40, Mtub21, Mtub30, Mtub39) shared by VNTR-15_{China} [14] and VNTR-$15_{standard}$ [9], 4 VNTR-15_{China} specific loci (ETRB, MIRU23, MIRU27, MIRU39) and 4 VNTR-$15_{standard}$ specific loci (Mtub04, QUB11b, QUB26, QUB4156). The PCR amplification and product electrophoresis were carried out as previously reported.

Data and statistical analysis

Spoligotyping data in binary format and MIRU-VNTR data in decimal format were analyzed

Table 1. Clinical characteristics of MTB patients: comparison of MDR patients with pan-sensitive patients

Characteristic	Total (N=422), n (%)	Pan-sensitive (N=311), n (%)	MDR[a] (N=58), n (%)	OR[b] (95% CI)	*p*-Value
Gender					
Male	276 (65.4)	203 (65.3)	35 (60.3)	0.81 (0.46-1.44)	0.472
Female	146 (34.6)	108 (34.7)	23 (39.7)	1.0 (Ref.)	
Age, years					
≤15	5 (1.2)	4 (1.3)	1 (1.7)		0.145
16-25	136 (32.2)	101 (32.5)	16 (27.6)		
26-35	78 (18.5)	61 (19.6)	11 (19.0)		
36-45	49 (11.6)	34 (10.9)	8 (13.8)		
46-55	54 (12.8)	35 (11.3)	10 (17.2)		
56-65	60 (14.2)	40 (12.9)	12 (20.7)		
66-75	25 (5.9)	21 (6.8)	0		
≥76	15 (3.6)	14 (4.5)	0		
Smoking					
Yes	131 (31.0)	98 (31.5)	16 (27.6)	0.83 (0.44-1.55)	0.553
No	291 (69.0)	213 (68.5)	42 (72.4)	1.0 (Ref.)	
Drinking					
Yes	75 (17.8)	54 (17.4)	13 (22.4)	1.38 (0.69-2.72)	0.360
No	347 (82.2)	257 (82.6)	45 (77.6)	1.0 (Ref.)	
TB Contact history					
Yes	61 (14.5)	44 (14.1)	12 (20.7)	1.58 (0.78-3.22)	0.202
No	361 (85.5)	267 (85.9)	46 (79.3)	1.0	
TB Family history					
Yes	34 (8.1)	26 (8.4)	7 (12.1)	1.51 (0.62-3.65)	0.327
No	388 (91.9)	285 (91.6)	51 (87.9)	1.0 (Ref.)	
Occupation					
Farmer	240 (56.9)	176 (56.6)	29 (50.0)		0.764
Student	63 (14.9)	47 (15.1)	9 (15.5)		
Worker	41 (9.7)	33 (10.6)	5 (8.6)		
Office clerk	20 (4.7)	15 (4.8)	4 (6.9)		
Unemployed	44 (10.4)	30 (9.6)	9 (15.5)		
Others	14 (3.3)	10 (3.2)	2 (3.4)		
Landform					
Mountainous region	107 (25.4)	72 (23.2)	17 (29.3)	1.38 (0.74-2.57)	0.314
Plain	315 (74.6)	239 (76.8)	41 (70.7)	1.0 (Ref.)	
Socieconomic status[b]					
Upper/middle	96 (22.7)	71 (22.8)	15 (25.9)	1.18 (0.62-2.25)	0.616
Lower	326 (77.3)	240 (77.2)	43 (74.1)	1.0 (Ref.)	
Medical history					
Initial treatment	319 (75.6)	260 (83.6)	17 (29.3)	1.0 (Ref.)	<0.01
Retreatment	103 (24.4)	51 (16.4)	41 (70.7)	12.30 (6.48-23.3)	
Diabetes melllitus					
Yes	47 (11.1)	37 (11.9)	7 (12.1)	1.02 (0.43-2.41)	0.970
No	375 (88.9)	274 (88.1)	51 (87.9)	1.0 (Ref.)	

[a]MDR, multi-drug resistant.[b]OR, odds ratio; CI, confidence interval.[c]Yin P, Zhang M. *et al*. Prevalence of COPD and its association with socioeconomic status in China: findings from China Chronic Disease Risk Factor Surveillance 2007. [J] BMC Public Health, 2011. 11: p. 586.

Table 2. Distribution of drug susceptibility patterns of the MTB strains with different medical history

DST[a]	n (%)	Initial treatment, n (%)	Retreatment, n (%)	OR (95% CI)	*p*-value
Total	422	319 (75.6)	103 (24.4)		
Pan sensitive	311 (73.7)	260 (81.5)	51 (49.5)	1.0 (Ref.)	
Any resistant	111 (26.3)	59 (18.5)	52 (50.5)	4.49 (2.78-7.25)	<0.01
Any SM resistant	86 (20.4)	47 (14.7)	39 (37.9)	4.23 (2.52-7.12)	<0.01
Any INH resistant	78 (18.5)	32 (24.8)	46 (44.7)	7.33 (4.26-12.60)	<0.01
Any RIF resistant	67 (15.9)	23 (7.2)	44 (42.7)	9.75 (5.42-17.54)	<0.01
Any EMB resistant	24 (5.7)	10 (3.1)	14 (13.6)	7.14 (3.00-16.96)	<0.01
Any OFX resistant	17 (4.0)	5 (1.6)	12 (11.7)	12.24 (4.13-36.24)	<0.01
Any KAN resistant	14 (3.3)	3 (0.9)	11 (10.7)	18.69 (5.04-69.40)	<0.01
MDR	58 (13.7)	17 (5.3)	41 (39.8)	12.30 (6.48-23.32)	<0.01

[a]DST, drug resistance testing; OR, odds ratio; CI, confidence interval; SM, streptomycin; INH, isoniazid; RIF, rifampicin; EMB, ethambutol; OFX, ofloxacin; KAN, kanamycin; MDR, multi-drug resistance.

using BioNumerics 5.0 software (Applied-Maths, Sint-Martens-Latem, Belgium). Dendrogram and minimum spanning tree were generated for clustering analysis. The discriminatory power and allelic diversity of each locus were determined by Hunter-Gaston discriminatory index (HGDI) value:

$$HDGI = 1 - \left[\frac{1}{N(N-1)}\sum_{j=1}^{s} n_j(n_j-1)\right]$$

Where N is the total number of strains in the typing method, s is the total number of different patterns discriminated by MIRU-VNTR, and n_j is the number of strains belonging to the j^{th} pattern. The clustering rate was defined as $(n_c-c)/N$, where n_c is the total number of clustered strains, c is the number of clusters and N is the total number of strains.

The statistical analysis was done by SPSS 16.0 software. *Chi-square* test or *Fisher's exact probability* test was used to compare the proportions of different groups and One-way ANOVA was used to explore the data of different age groups. Two-sided *p* value of less than 0.05 was considered to be statistically significant.

Results

Drug resistance

DST showed that 311 (73.7%) were sensitive strains, and 111 (26.3%) were resistant to one or more of the drugs tested, including 58 (13.7%) MDR isolates. The resistance rate of SM, INH, RIF, EMB, OFX and KAN was 20.4%, 18.5%, 15.9%, 5.7%, 4.0% and 3.3%, respectively. Statistical analysis showed that both the total resistance rate and MDR rate of the retreatment group were significantly higher than that of the initial treatment group, and the same results were seen in SM, INH, RIF, EMB, OFX and KAN mono-resistance rate between-the two groups (**Table 2**). Compared to initial treatment panels, retreatment patients had a significantly higher frequency of developing MDR [*P*<0.01; OR (95% CI): 12.30 (6.481-23.32)]. However, the patients with diabetes had no statistical difference in any drug resistance rate compared to the patients without diabetes (Table S1). In addition, there was no statistical difference between the pan-sensitive group and MDR group in any clinical information collected (**Table 1**).

Spoligotyping

According to SpolDB4.0, 384 (91.0%) strains belonged to the Beijing family and 38 (9.0%) strains were the non-Beijing family, including 8 strains of T1, 7 of T2, 3 of T3, 1 of T4, 3 of MANU2, 1 of LAM1, 3 of LAM9, 2 of H3, 1 of U, 2 of AMBIGOUS: T3 T2 and 7 presenting spoligotypes that were not found in the database (**Table 3**). Clustering analysis showed that 392 (92.9%) strains were grouped into 10 clusters (2-361 strains per cluster), and the remaining 30 (7.1%) strains were unique. The largest cluster belonged to the typical Beijing family, containing 361 (85.5%) strains characterized by the absence of the first 34 spacers and the presence of spacers 35 to 43. The remaining 23 (5.5%) strains were defined as the atypical Beijing family for missing one or more dots in spacers 35-43 (Figure S1). The HGDI of spoligotyping for all strains was 0.271.

Table 3. Spoligotypes of *Mycobacterium tuberculosis* strains in this study (n=422)

No.	Spoligotype	SIT[a]	Family[b]	N (%)[c]
1	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	1	Beijing	361 (85.54)
2	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■□■■■	190	Beijing	8 (1.90)
3	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■□□■■■■■	260	Beijing	1 (0.24)
4	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■□■■■■■■	265	Beijing	2 (0.47)
5	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■	269	Beijing	1 (0.24)
6	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■□□■■	541	Beijing	1 (0.24)
7	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■□■■■■■■	621	Beijing	4 (0.95)
8	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■□■■	941	Beijing	1 (0.24)
9	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■□□□□	1311	Beijing	1 (0.24)
10	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■□□■■■■	1364	Beijing	1 (0.24)
11	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■□□□□□□□	2101	Beijing	1 (0.24)
12	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■□	2610	Beijing	1 (0.24)
13	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■□□□■■■	ORPHAN	Beijing	1 (0.24)
14	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	53	T1	4 (0.95)
15	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□■■■■■■	240	T1	1 (0.24)
16	■□■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	334	T1	2 (0.47)
17	■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□□□□□□	1793	T1	1 (0.24)
18	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□■■■	52	T2	4 (0.95)
19	■■■□■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□■■■	848	T2	1 (0.24)
20	□□■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■□■■■	1613	T2	1 (0.24)
21	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□■■■	ORPHAN	T2	1 (0.24)
22	■■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■■■■■	37	T3	1 (0.24)
23	■□■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■■■■■	ORPHAN	T3	2 (0.47)
24	■■■■■■■■■■■■■■■■□■■■■■■■■■■■■□□□□■■■■■■■	40	T4	1 (0.24)
25	■□□□□□■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	961	LAM1	1 (0.24)
26	■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	42	LAM9	1 (0.24)
27	■■■■■■■■■■■■■□□□□□□□□□□■■■■■■■■□□□□■■■■■■■	803	LAM9	1 (0.24)
28	■■■■■■■■■■■■■□□□□■□□□□□□■■■■■■■■□□□□■■■■■■■	2191	LAM9	1 (0.24)
29	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□■■■■■■■■■	54	MANU2	3 (0.71)
30	■■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■	50	H3	1 (0.24)
31	■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□■□□□□■■■■■■■	1908	H3	1 (0.24)
32	■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□□□□□□■■■■■■■	602	U	1 (0.24)
33	■□■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■□■■■	ORPHAN	AMBIGOUS:T3T2	2 (0.47)
34	■■■■■■■□■■■■■■■■□■■■□■■■■■■■■■□□■■■■■■■■■	-	NEW	1 (0.24)
35	■■■■■■■■■■■■■■■■■■■■■■■■□■■■■■■□□□□□□■■■■■	-	NEW	1 (0.24)
36	■□□■■■□□■■■■■■■■■■■■■□□■■■□□□□□□■■□□□■□■■	-	NEW	1 (0.24)
37	■■■■□■■■■■■■■■■■■■■■■■■■■□□□□□□□□□□□□□□□	-	NEW	1 (0.24)
38	■■■■■■■■■■■■□□■■■■■■■■■■■■□■□■■□□■■■■■■■■■	-	NEW	1 (0.24)
39	□□□□■■■■■■■■■■■■■■■■■■■■■■■■□□□□□■■■□■■■	-	NEW	1 (0.24)
40	■■■■■■■■■■■■□■■■■■■■■■■■■■■□■■□□□□■■□■■■■	-	NEW	1 (0.24)

[a]SIT, spoligotype international type. [b]Spoligotype families as assigned in SpolDB4.0. [c]The number of the strains with a common SIT.

The Beijing family strains had a higher risk for developing MDR than the non-Beijing family, but no significant difference was observed. Meanwhile, the drug resistance rate of the Beijing family and the non-Beijing family did not show any significance in SM, INH, RIF, EMB, OFX and KAN (Table S2). Of the Beijing genotype patients, 29.4% had a smoking history, sig-

Table 4. Clinical characteristics of Beijing family and non-Beijing family strains

Characteristic	Total (N=422), n (%)	Beijing (N=384), n (%)	Non-Beijing (N=38), n (%)	OR (95% CI)	*p*-Value
Gender					
Male	276 (65.4)	247 (64.3)	29 (76.3)	0.56 (0.26-1.22)	0.138
Female	146 (34.6)	137 (35.7)	9 (23.7)	1.0 (Ref.)	
Age, years					
≤15	5 (1.2)	4 (1.0)	1 (2.6)		0.497
16-25	136 (32.2)	128 (33.3)	8 (21.1)		
26-35	78 (18.5)	71 (18.5)	7 (18.4)		
36-45	49 (11.6)	42 (10.9)	7 (18.4)		
46-55	54 (12.8)	48 (12.5)	6 (15.8)		
56-65	60 (14.2)	52 (13.5)	8 (21.1)		
66-75	25 (5.9)	24 (6.3)	1 (2.6)		
≥76	15 (3.6)	14 (3.6)	1 (2.6)		
Smoking history					
Yes	131 (31.0)	114 (29.7)	18 (47.4)	2.13 (1.09-4.18)	0.025
No	291 (69.0)	270 (70.3)	20 (52.6)	1.0 (Ref.)	
Drinking history					
Yes	75 (17.8)	64 (16.7)	11 (28.9)	0.49 (0.23-1.04)	0.059
No	347 (82.2)	320 (83.3)	27 (71.1)	1.0 (Ref.)	
TB Contact history					
Yes	61 (14.5)	60 (15.6)	1 (2.6)	6.85 (0.92-50.92)	0.028
No	361 (85.5)	324 (84.4)	37 (97.4)	1.0 (Ref.)	
TB Family history					
Yes	34 (8.1)	34 (8.9)	0	7.58 (0.45-126.2)	0.059
No	388 (91.9)	350 (91.1)	38 (100)	1.0 (Ref.)	
Occupation					
Farmer	240 (56.9)	215 (56.0)	25 (65.8)		0.242
Student	63 (14.9)	60 (15.6)	3 (7.9)		
Worker	41 (9.7)	38 (9.9)	3 (7.9)		
Office clerk	20 (4.7)	20 (5.2)	0		
Unemployed	44 (10.4)	40 (10.4)	4 (10.5)		
Others	14 (3.3)	11 (2.9)	3 (7.9)		
Landform					
Mountainous region	107 (25.4)	99 (25.8)	8 (21.1)	1.30 (0.58-2.94)	0.523
Plain	315 (74.6)	285 (74.2)	30 (78.9)	1.0 (Ref.)	
Socieconomic status[a]					
Upper/middle	96 (22.7)	90 (23.4)	6 (15.8)	1.63 (0.66-4.03)	0.283
Lower	326 (77.3)	294 (76.6)	32 (84.2)	1.0 (Ref.)	
Medical history					
Initial treatment	319 (75.6)	292 (76.0)	27 (71.1)	1.29 (0.62-2.71)	0.495
Retreatment	103 (24.4)	92 (24.0)	11 (28.9)	1.0 (Ref.)	
Diabetes mellitus					
Yes	47 (11.1)	44 (11.5)	3 (7.9)	1.51 (0.45-5.12)	0.786
No	375 (88.9)	340 (88.5)	35 (92.1)	1.0 (Ref.)	

[a]Yin P *et al*. Prevalence of COPD and its association with socioeconomic status in China: findings from China Chronic Disease Risk Factor Surveillance 2007. [J] BMC Public Health, 2011. 11: p. 586.

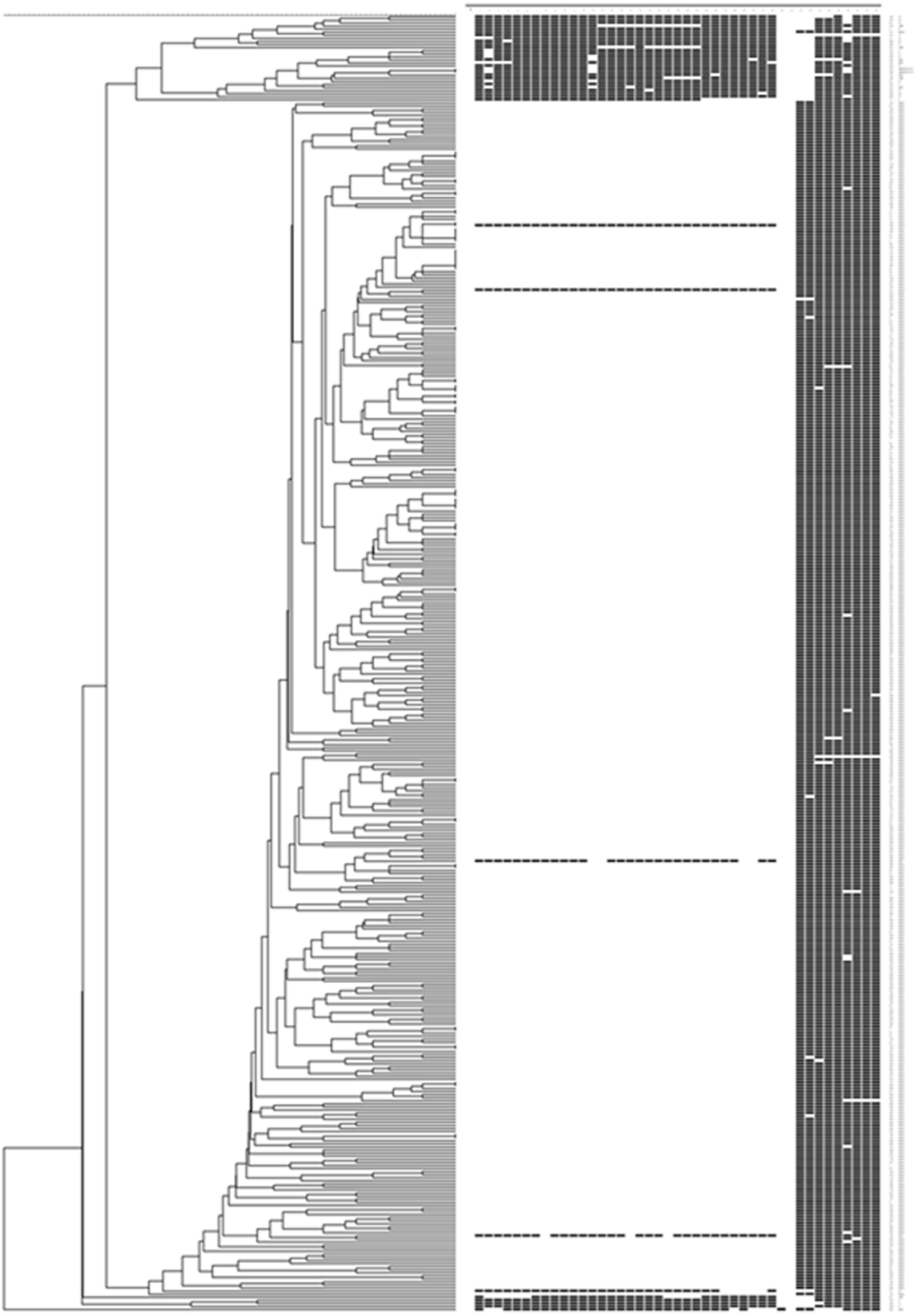

Figure 1. Molecular typing of 422 *M. tuberculosis* strains with VNTR-$15_{Standard}$ and spoligotyping. From left to right: UPGMA dendrogram generated by VNTR-$15_{Standard}$ method, spoligotyping pattern, Spoligo-International-Type (SIT) number, genetic lineage based on SpolDB4.0, strain numbers.

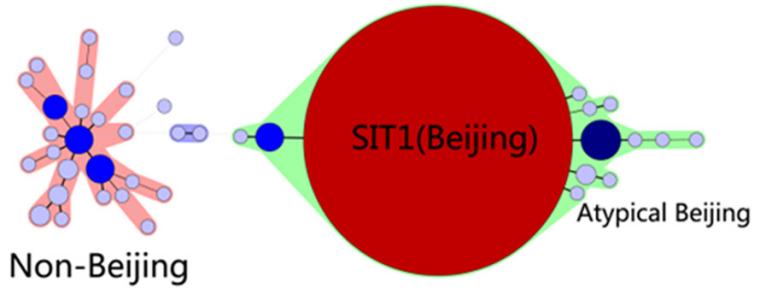

Figure 2. Minimum spanning tree showing the clustering by spoligotyping of 422 *M. tuberculosis* strains from Shijiazhuang. Each nodal point represents a cluster of a particular spoligotypes, the colour and the size of nodal points is relative to the number of strains within that cluster.

nificantly lower than that of the non-Beijing ones [*P*=0.025; OR (95% CI): 2.13 (1.09-4.18)]. Moreover, 15.6% of the Beijing lineage group had a TB contact history, significantly higher than that of the non-Beijing group [*P*=0.028; OR (95% CI): 6.85 (0.92-50.92)]. There was no statistical difference within other clinical factors between the Beijing and non-Beijing family (**Table 4**).

MIRU-VNTR

The VNTR-15$_{standard}$ cluster analysis divided 69 strains into 27 cluster types (2-7 strains per cluster) and the remaining 353 were individual types (**Figure 1**). The clustering rate was 9.95% and its HGDI value reached 0.999. In contrast, the VNTR-15$_{China}$ method with a clustering rate of 40.3% and HGDI of 0.989, grouped 229 strains into 59 clusters (2-32 strains per cluster) and the remaining 193 were unique genotypes (Figure S2). The HGDI scores of the 19 loci varied significantly from 0.729 for QUB11b to 0.065 for ETRC. The MIRU-VNTR loci were classified into highly (>0.6), moderately (0.3 to 0.6) and poorly (<0.3) discriminatory power based on HGDI values [15]. The VNTR-15$_{standard}$ consisted of 5 highly discriminatory loci (QUB11b, QUB26, QUB4156, Mtub04, Mtub21), 5 moderately discriminatory loci (MIRU26, ETRE, ETRA, MIRU10, MIRU40) and 5 poorly discriminatory loci (MIRU16, MIRU39, ETRD, Mtub30, ETRC). But the VNTR-15$_{China}$ had only 1 highly discriminatory locus (Mtub21), 5 moderate loci (MIRU26, ETRE, ETRA, MIRU10, MIRU40) and 9 poor loci (MIRU16, MIRU39, Mtub39, ETRD, MIRU27, Mtub30, ETRB, MIRU23, ETRC) (Table S3).

Minimum spanning tree

To understand the major clusters and describe their genetic links, minimum spanning tree was mapped (**Figure 2**). Each nodal point represented a particular genotype, and the size of each nodal point was determined by the number of strains within that genotype [7]. The Beijing and the non-Beijing family were classified into two large groups. The biggest cluster containing 361 strains (SIT1) represented the typical Beijing family and surrounded by small clusters of the atypical Beijing family. The three larger clusters on the left were MANU2, T1 and T2 family. In particular, 7 newly-defined strains, similar to MANU2 or T family, were clustered into the non-Beijing family.

The minimum spanning tree presented the clustering results of VNTR-15$_{standard}$ (a) and VNTR-15$_{China}$ (b) genotyping methods (**Figure 3**). The figure clearly showed that the polymorphism of strains enrolled was intensively high and VNTR-15$_{standard}$ grouped the strains into more clusters than VNTR-15$_{China}$, suggesting that VNTR-15$_{standard}$ typing method have relatively higher discriminatory power than VNTR-15$_{China}$. The largest cluster belonged to the Beijing family (light red and yellow) and separated from the non-Beijing family (green and blue). In particular, 7 newly-defined strains were clustered into the non-Beijing group, consistent with clustering results from spoligotyping.

Discussion

The present study aimed to investigate the genetic diversity and drug resistance of MTB strains isolated from Shijiazhuang, an area of little information about MTB molecular characteristics. A total of 422 MTB strains were collected during 2014 from the Fifth Hospital of Shijiazhuang, which has the largest TB outpatient capacity in Shijiazhuang. Based on the Nationwide Census, 50.06% of Shijiazhuang population was male. Taking this into account, we found that men had a higher TB morbidity rate than women, and farmers were the largest TB infection group, followed by students and

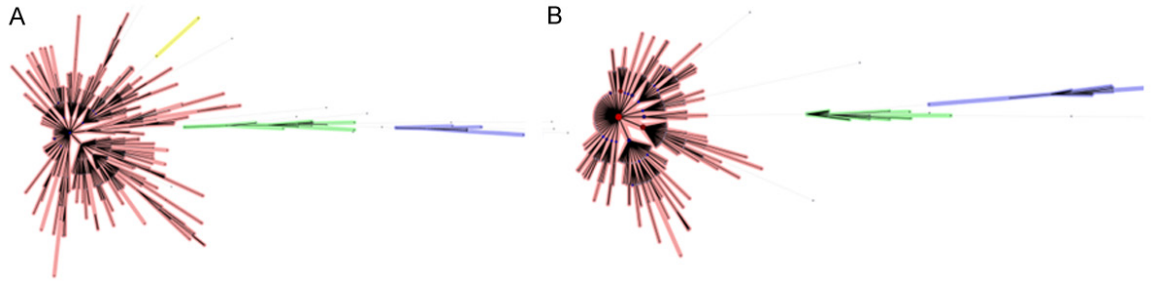

Figure 3. Minimum spanning tree showing the clustering by the VNTR-15$_{Standard}$ (A) and VNTR-15$_{China}$ (B) of 422 *M. tuberculosis* strains from Shijiazhuang. Each nodal point represents a cluster of a particular spoligotypes, the colour and the size of nodal points is relative to the number of strains within that cluster.

the unemployed. Although people of all ages can be infected with MTB, 16 to 35 years old young adults suffered the most, leading to the loss of social labor. DST showed that 13.7% of all strains were MDR, higher than the data reported on the National Baseline Survey in Drug-resistant Tuberculosis during 2007-2008 [16]. Retreatment patients had a significantly higher MDR rate than new cases, consistent with early reported data [17, 18], and the similar results were seen in SM, INH, RIF, EMB, OFX and KAN resistance rate between the initial treatment group and the retreatment group in our study, indicating that unfavorable antibiotic treatment may contribute to the development of drug resistant *M. tuberculosis* [19]. However, our data suggested that other clinical factors, such as gender, age, smoking, drinking, contact history, family history, occupation, socioeconomic status, DM and landform, exhibited no statistical association with drug resistance partly due to the large proportion of the Beijing family strains. A DM screening survey among TB patients and non-TB controls, revealed that no association between DM and TB was found, in agreement with our findings [20]. Instead, another study found that TB patients with DM had more cavitation, higher smear grade and more MDR-TB than patients without DM [21]. More studies may be needed to further explain the relationship between the two diseases, DM and TB.

The Beijing family is the predominant lineage of TB strains in Northern China with a prevalence rate of 67.12%-96.3%, as shown in Figure S3 [5, 10, 22-29]. Our data demonstrated the Beijing family was also the predominantly epidemic genotype in Shijiazhuang, accounting for 91.0%. The genetic diversity of *M. tuberculosis* strains in Shijiazhuang appeared to be highly polymorphic, as many spoligotypes, such as Beijing, T1, T2, T3, T4, MANU2, LAM1, LAM9, H3 and U, were found in this area. Prior studies considered Beijing family is related to drug resistance [30-32], while others believe that the Beijing lineage are no more likely to acquire drug resistance than the non-Beijing genotypes [33]. In our study, no evidence supported that the Beijing family was related to any drug resistance. However, patients with a TB contact history were more susceptible to the Beijing genotype strains, indicating that the Beijing genotype strains possess a stronger ability to cause transmission than the non-Beijing strains [34]. Interestingly, smoking patients were more likely to be infected with the non-Beijing strains than the Beijing strains, but reasons for this phenomenon remain poorly understood. The estimation for the difference between the Beijing family and the non-Beijing family in other clinical factors was probably biased by the small size of the non-Beijing family strains.

Due to the low discriminatory power of spoligotyping, MIRU-VNTR was used for further analysis. Considering the geographical difference, strain polymorphism and high prevalence of the Beijing family strains in China, a suitable set of VNTR loci for molecular epidemiology should be evaluated [35]. In order to get VNTR loci suitable for Shijiazhuang strains, we selected 19 loci that were included in the classical VNTR-15$_{standard}$ method or VNTR-15$_{China}$ method, and the allelic diversity varied significantly at each VNTR locus. Among the VNTR-15$_{standard}$ loci set, QUB11b had the highest HGDI score of 0.729, followed by QUB26 (HGDI=0.676), QUB4156 (HGDI=0.624) and ETRC had the lowest HGDI of 0.065, similar to previous studies

from China [19, 32]. The highest discriminative locus of VNTR-15_{China} was Mtub21 (HGDI=0.602), and similar with the results from Beijing (0.688) [6] and Inner Mongolia (0.655) [36]. Four loci in VNTR-15_{China}, not belonging to VNTR-$15_{standard}$ subset, presented rather poor allelic diversity, and the same results were seen in Gansu [37]. VNTR-15_{China} method contained much more low discriminative loci than VNTR-$15_{standard}$ group, which means unrelated strains will be clustered together. The clustering analysis also indicated that VNTR-15_{China} can group more clusters and put more strains into a particular cluster than VNTR-$15_{standard}$ method. Although VNTR-$15_{standard}$ subset gives proper HGDI scores, much more individual types were aslo generated, which means related strains did not be clustered [38]. Therefore, more studies are in needed to achieve a more reasonable set of MIRU-VNTR loci for strains in Shijiazhuang.

Conclusions

In summary, the *Mycobacterium tuberculosis* isolates in Shijiazhuang is of highly genetic diversity. The Beijing family is the predominant genotype in this area. Anti-TB drug treatment is associated with MTB drug resistance. VNTR-$15_{standard}$ has a higher discriminatory power than VNTR-15_{China}. Combination of spoligotyping and MIRU-VNTR is suitable for understanding the molecular epidemiology of MTB strains in Shijiazhuang.

Acknowledgements

This study was supported by the project 2013ZX10003002-004 of the National Key Program of Mega Infectious Disease. We thank the stuff of the Fifth Hospital of Shijiazhuang and Chinese Center for Disease Control and Prevention for their excellent contributions to this research.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Kanglin Wan, State Key Laboratory for Infectious Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, 155 Changbai Road, Changping, Beijing 102206, China. E-mail: cdcwankanglin@163.com; Dr. Erhei Dai, Department of Laboratory Medicine, The Fifth Hospital of Shijiazhuang, Hebei Medical University, 42 Tanan Road, Shijiazhuang 050021, Hebei, China. E-mail: daieh2008@126.com

References

[1] World Health Organization. Global tuberculosis report 2014. Geneva, Switzerland: World Health Organization; 2014.

[2] Jia ZW, Jia XW, Liu YX, Dye C, Chen F, Chen CS, Zhang WY, Li XW, Cao WC and Liu HL. Spatial analysis of tuberculosis cases in migrants and permanent residents, Beijing, 2000-2006. Emerg Infect Dis 2008; 14: 1413-1419.

[3] Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M and van Embden J. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. J Clin Microbiol 1997; 35: 907-914.

[4] Kremer K, Glynn JR, Lillebaek T, Niemann S, Kurepina NE, Kreiswirth BN, Bifani PJ and van Soolingen D. Definition of the Beijing/W lineage of Mycobacterium tuberculosis on the basis of genetic markers. J Clin Microbiol 2004; 42: 4040-4049.

[5] Liu Y, Tian M, Wang X, Wei R, Xing Q, Ma T, Jiang X, Li W, Zhang Z, Xue Y, Zhang X, Wang W, Wang T, Hong F, Zhang J, Wang S and Li C. Genotypic diversity analysis of Mycobacterium tuberculosis strains collected from Beijing in 2009, using spoligotyping and VNTR typing. PLoS One 2014; 9: e106787.

[6] Weniger T, Krawczyk J, Supply P, Niemann S and Harmsen D. MIRU-VNTRplus: a web tool for polyphasic genotyping of Mycobacterium tuberculosis complex bacteria. Nucleic Acids Res 2010; 38: W326-331.

[7] Bauer J, Kok-Jensen A, Faurschou P, Thuesen J, Taudorf E and Andersen AB. A prospective evaluation of the clinical value of nation-wide DNA fingerprinting of tuberculosis isolates in Denmark. Int J Tuberc Lung Dis 2000; 4: 295-299.

[8] Allix-Beguec C, Fauville-Dufaux M and Supply P. Three-year population-based evaluation of standardized mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing of Mycobacterium tuberculosis. J Clin Microbiol 2008; 46: 1398-1406.

[9] Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C and van Soolingen D. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Myco-

bacterium tuberculosis. J Clin Microbiol 2006; 44: 4498-4510.
[10] Mears J, Abubakar I, Cohen T, McHugh TD and Sonnenberg P. Effect of study design and setting on tuberculosis clustering estimates using Mycobacterial Interspersed Repetitive Units-Variable Number Tandem Repeats (MIRU-VNTR): a systematic review. BMJ Open 2015; 5: e005636.
[11] Ministry of Health of the People's Republic of China. The 2000 national tuberculosis epidemiological survey data compilation. People's Medical Publishing House; 2003. pp. 15-16.
[12] WHO. WHO Guidelines for surveillance of drug resistance in tuberculosis.Available: http://www.who.int/tb/publications/2009/surveillance_guidelines/en/index.html. Accessed 20 Aug 2013. 2013.
[13] Shao Y, Yang D, Xu W, Lu W, Song H, Dai Y, Shen H and Wang J. Epidemiology of anti-tuberculosis drug resistance in a Chinese population: current situation and challenges ahead. BMC Public Health; 2011.pp. 110.
[14] Wan K, Liu J, Hauck Y, Zhang Y, Zhao X, Liu Z, Lu B, Dong H, Jiang Y, Kremer K, Vergnaud G, van Soolingen D and Pourcel C. Investigation on Mycobacterium tuberculosis diversity in China and the origin of the Beijing clade. PLoS One 2011; 6: e29190.
[15] Mokrousov I, Jiao WW, Wan K and Shen A. Stranger in a strange land: Ibero-American strain of Mycobacterium tuberculosis in Tibet, China. Infect Genet Evol 2014; 26: 323-326.
[16] Zhao Y, Xu S, Wang L, Chin DP, Wang S, Jiang G, Xia H, Zhou Y, Li Q, Ou X, Pang Y, Song Y, Zhao B, Zhang H, He G, Guo J and Wang Y. National survey of drug-resistant tuberculosis in China. N Engl J Med 2012; 366: 2161-2170.
[17] Liu Q, Yang D, Xu W, Wang J, Lv B, Shao Y, Song H, Li G, Dong H, Wan K and Wang H. Molecular typing of Mycobacterium tuberculosis isolates circulating in Jiangsu province, China. BMC Infect Dis 2011; 11: 288.
[18] Ayaz A, Hasan Z, Jafri S, Inayat R, Mangi R, Channa AA, Malik FR, Ali A, Rafiq Y and Hasan R. Characterizing Mycobacterium tuberculosis isolates from Karachi, Pakistan: drug resistance and genotypes. Int J Infect Dis 2012; 16: e303-309.
[19] Hang NT, Maeda S, Keicho N, Thuong PH and Endo H. Sublineages of Mycobacterium tuberculosis Beijing genotype strains and unfavorable outcomes of anti-tuberculosis treatment. Tuberculosis (Edinb) 2015; 95: 336-342.
[20] Haraldsdottir TL, Rudolf F, Bjerregaard-Andersen M, Carlos Joaquim L, Stochholm K, Gomes VF, Beck-Nielsen H, Ostergaard L, Aaby P and Wejse C. Diabetes mellitus prevalence in tuberculosis patients and the background population in Guinea-Bissau: a disease burden study from the capital Bissau. Trans R Soc Trop Med Hyg 2015; 109: 400-407.
[21] Magee MJ, Kempker RR, Kipiani M, Gandhi NR, Darchia L, Tukvadze N, Howards PP, Narayan KM and Blumberg HM. Diabetes mellitus is associated with cavities, smear grade, and multidrug-resistant tuberculosis in Georgia. Int J Tuberc Lung Dis 2015; 19: 685-692.
[22] Dong H, Liu Z, Lv B, Zhang Y, Liu J, Zhao X and Wan K. Spoligotypes of Mycobacterium tuberculosis from different Provinces of China. J Clin Microbiol 2010; 48: 4102-4106.
[23] Li WM, Wang SM, Pei XY, Liu ZQ, Zhong Q, Qian M, Zhao B and Duanmu HJ. [DNA fingerprinting of Mycobacterium tuberculosis strains from Beijing, Guangdong and Ningxia]. Zhonghua Liu Xing Bing Xue Za Zhi 2003; 24: 381-384.
[24] Han H, Wang F, Xiao Y, Ren Y, Chao Y, Guo A and Ye L. Utility of mycobacterial interspersed repetitive unit typing for differentiating Mycobacterium tuberculosis isolates in Wuhan, China. J Med Microbiol 2007; 56: 1219-1223.
[25] Kam KM, Yip CW, Tse LW, Wong KL, Lam TK, Kremer K, Au BK and van Soolingen D. Utility of mycobacterial interspersed repetitive unit typing for differentiating multidrug-resistant Mycobacterium tuberculosis isolates of the Beijing family. J Clin Microbiol 2005; 43: 306-313.
[26] Dou HY, Tseng FC, Lin CW, Chang JR, Sun JR, Tsai WS, Lee SY, Su IJ and Lu JJ. Molecular epidemiology and evolutionary genetics of Mycobacterium tuberculosis in Taipei. BMC Infect Dis 2008; 8: 170.
[27] Zhou A, Nawaz M, Duan Y, Moore JE, Millar BC, Xu J and Yao Y. Molecular characterization of isoniazid-resistant Mycobacterium tuberculosis isolates from Xi'an, China. Microb Drug Resist 2011; 17: 275-281.
[28] Chen L, Li N, Liu Z, Liu M, Lv B, Wang J, Zhang J, Wan K and Zhang H. Genetic diversity and drug susceptibility of Mycobacterium tuberculosis isolates from Zunyi, one of the highest-incidence-rate areas in China. J Clin Microbiol 2012; 50: 1043-1047.
[29] Qiao K WH, Yang CG, et al. The application of variable number of tandem repeats in the microevolution study of Mycobacterium tuberculosis strain of Beijing genotype in Chongming Island, Shanghai. Journal of Microbes and Infections 2010; 5: 208-213.
[30] Duong DA, Nguyen TH, Nguyen TN, Dai VH, Dang TM, Vo SK, Do DA, Nguyen VV, Nguyen HD, Dinh NS, Farrar J and Caws M. Beijing genotype of Mycobacterium tuberculosis is significantly associated with high-level fluoroquinolone resistance in Vietnam. Antimicrob Agents Chemother 2009; 53: 4835-4839.

[31] Sun YJ, Lee AS, Wong SY, Heersma H, Kremer K, van Soolingen D and Paton NI. Genotype and phenotype relationships and transmission analysis of drug-resistant tuberculosis in Singapore. Int J Tuberc Lung Dis 2007; 11: 436-442.

[32] Zhang Z, Lu J, Liu M, Wang Y, Qu G, Li H, Wang J, Pang Y, Liu C and Zhao Y. Genotyping and molecular characteristics of multidrug-resistant Mycobacterium tuberculosis isolates from China. J Infect 2015; 70: 335-345.

[33] Wang XH, Ma AG, Han XX, Gu XM, Fu LP, Li PG, Li FY, Wang QZ, Liang H, Katar A and Wang LJ. Correlations between drug resistance of Beijing/W lineage clinical isolates of Mycobacterium tuberculosis and sublineages: a 2009-2013 prospective study in Xinjiang province, China. Med Sci Monit 2015; 21: 1313-1318.

[34] Yang C, Luo T, Sun G, Qiao K, DeRiemer K, Mei J and Gao Q. Mycobacterium tuberculosis Beijing strains favor transmission but not drug resistance in China. Clin Infect Dis 2012; 55: 1179-1187.

[35] Jiang Y, Liu HC, Zheng H, Dou X, Tang B, Zhao XQ, Zhu Y, Lu B, Wang S, Dong HY, Zhang YY, Zhao G and Wan K. 19-VNTR loci used in genotyping Chinese clinical Mycobacterium tuberculosis complex strains and in association with spoligotyping. J Basic Microbiol 2013; 53: 562-580.

[36] Yu Q, Su Y, Lu B, Ma Y, Zhao X, Yang X, Dong H, Liu Y, Lian L, Wan L, Wu Y and Wan K. Genetic diversity of Mycobacterium tuberculosis isolates from Inner Mongolia, China. PLoS One 2013; 8: e57660.

[37] Liu J, Tong C, Jiang Y, Zhao X, Zhang Y, Liu H, Lu B and Wan K. First insight into the genotypic diversity of clinical Mycobacterium tuberculosis isolates from Gansu Province, China. PLoS One 2014; 9: e99357.

[38] Deng YF, Zhang YA, Zheng JL, Jing H, Wang Y, Wang HY, Ma X and Liu ZM. [Study on molecular characteristics regarding DNA genotype of Mycobacterium tuberculosis clinical strains in Shandong]. Zhonghua Liu Xing Bing Xue Za Zhi 2010; 31: 321-323.

Table S1. Distribution of drug susceptibility patterns of the MTB strains between diabetes group and non-diabetes group

DST	*n* (%)	Diabetes, *n* (%)	Non-diabetes, *n* (%)	OR (95% CI)	*p*-value
Total	422	47 (11.1)	375 (88.9)		
Pan sensitive	311 (73.7)	37 (78.7)	274 (73.1)	1.0 (Ref.)	
Any resistant	111 (26.3)	10 (21.3)	101 (26.9)	0.73 (0.35-1.53)	0.406
Any SM resistant	86 (20.4)	9 (19.1)	77 (20.5)	0.87 (0.40-1.87)	0.714
Any INH resistant	78 (18.5)	7 (14.9)	71 (18.9)	0.73 (0.31-1.71)	0.466
Any RIF resistant	67 (15.9)	7 (14.9)	60 (34.9)	0.86 (0.37-2.03)	0.737
Any EMB resistant	24 (5.7)	2 (4.3)	22 (5.9)	0.67 (0.15-2.98)	1.000
Any OFX resistant	17 (4.0)	1 (2.1)	16 (4.3)	0.46 (0.06-3.59)	0.704
Any KAN resistant	14 (3.3)	2 (4.3)	12 (3.2)	1.23 (0.27-5.74)	0.679
MDR	58 (13.7)	7 (14.9)	51 (13.6)	1.02 (0.43-2.41)	0.970

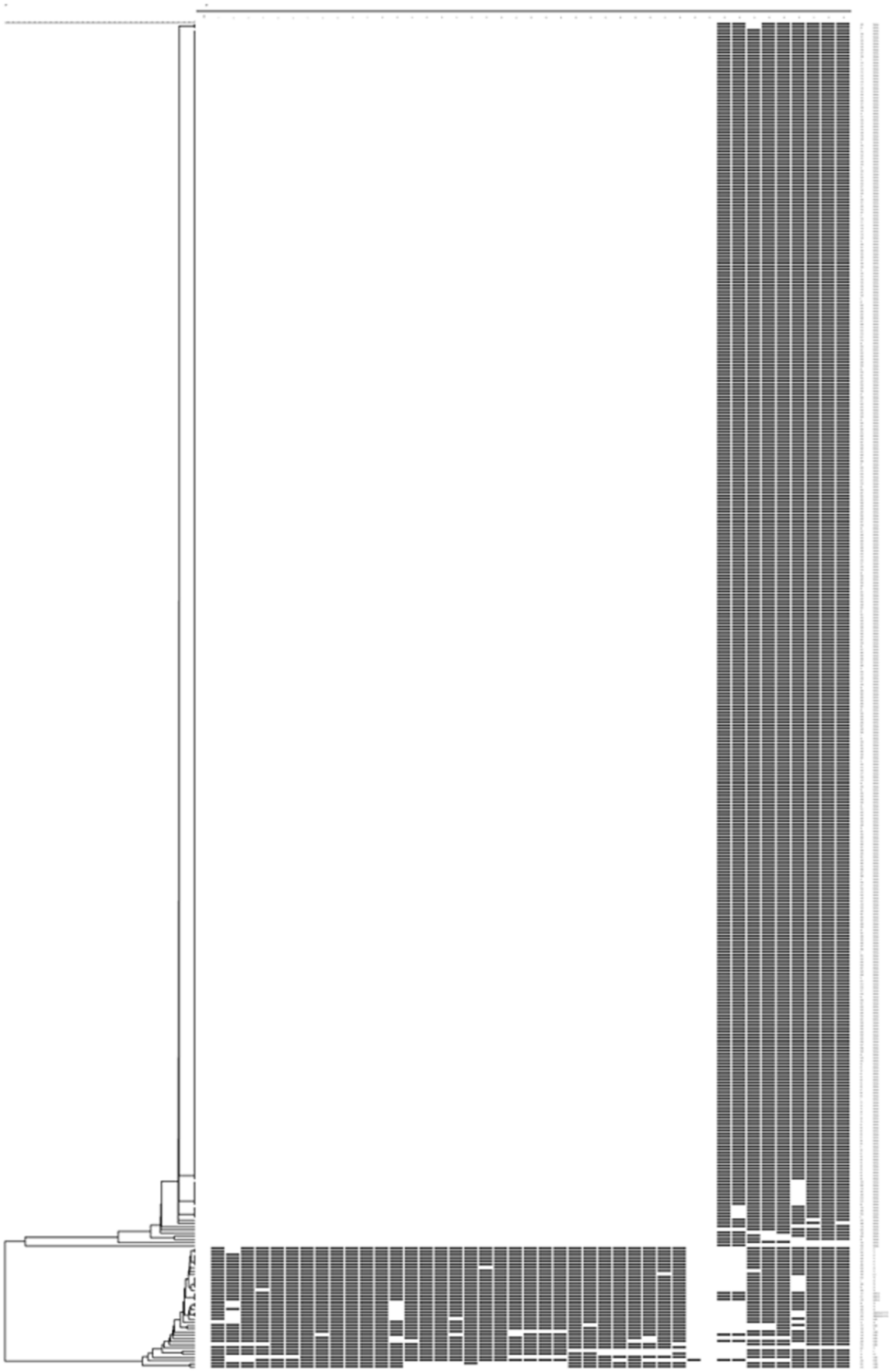

Figure S1. Molecular typing of 422 *M. tuberculosis* strains with spoligotyping. The clustering analysis was conducted using BoiNumerics 5.0 software. From left to right: UPGMA dendrogram generated by spoligotyping, spoligotyping pattern, Spoligo-International-Type (SIT) number, genetic lineage based on SpolDB4.0, strain numbers.

Table S2. Distribution of drug susceptibility patterns of the MTB strains between Beijing and non-Beijing family

DST	*n* (%)	Beijing, *n* (%)	Non-Beijing, *n* (%)	OR (95% CI)	*p*-value
Total	422	384 (91.0)	38 (9.0)		
Pan sensitive	311 (73.7)	282 (73.4)	29 (76.3)	1.0 (Ref.)	
Any resistant	111 (26.3)	102 (26.6)	9 (23.7)	1.17 (0.53-2.55)	0.701
Any SM resistant	86 (20.4)	78 (20.3)	8 (21.1)	1.00 (0.44-2.28)	0.995
Any INH resistant	78 (18.5)	72 (18.8)	6 (15.8)	1.23 (0.49-3.09)	0.652
Any RIF resistant	67 (15.9)	63 (16.4)	4 (10.5)	1.62 (0.55-4.77)	0.479
Any EMB resistant	24 (5.7)	23 (6.0)	1 (2.6)	2.37 (0.31-18.17)	0.709
Any OFX resistant	17 (4.0)	15 (3.9)	2 (5.3)	0.77 (0.17-3.54)	0.668
Any KAN resistant	14 (3.3)	14 (3.6)	0	3.03 (0.18-52.11)	0.623
MDR	58 (13.7)	55(14.3)	3 (7.9)	1.89 (0.56-6.41)	0.446

Genotypes and drug resistance of TB isolates

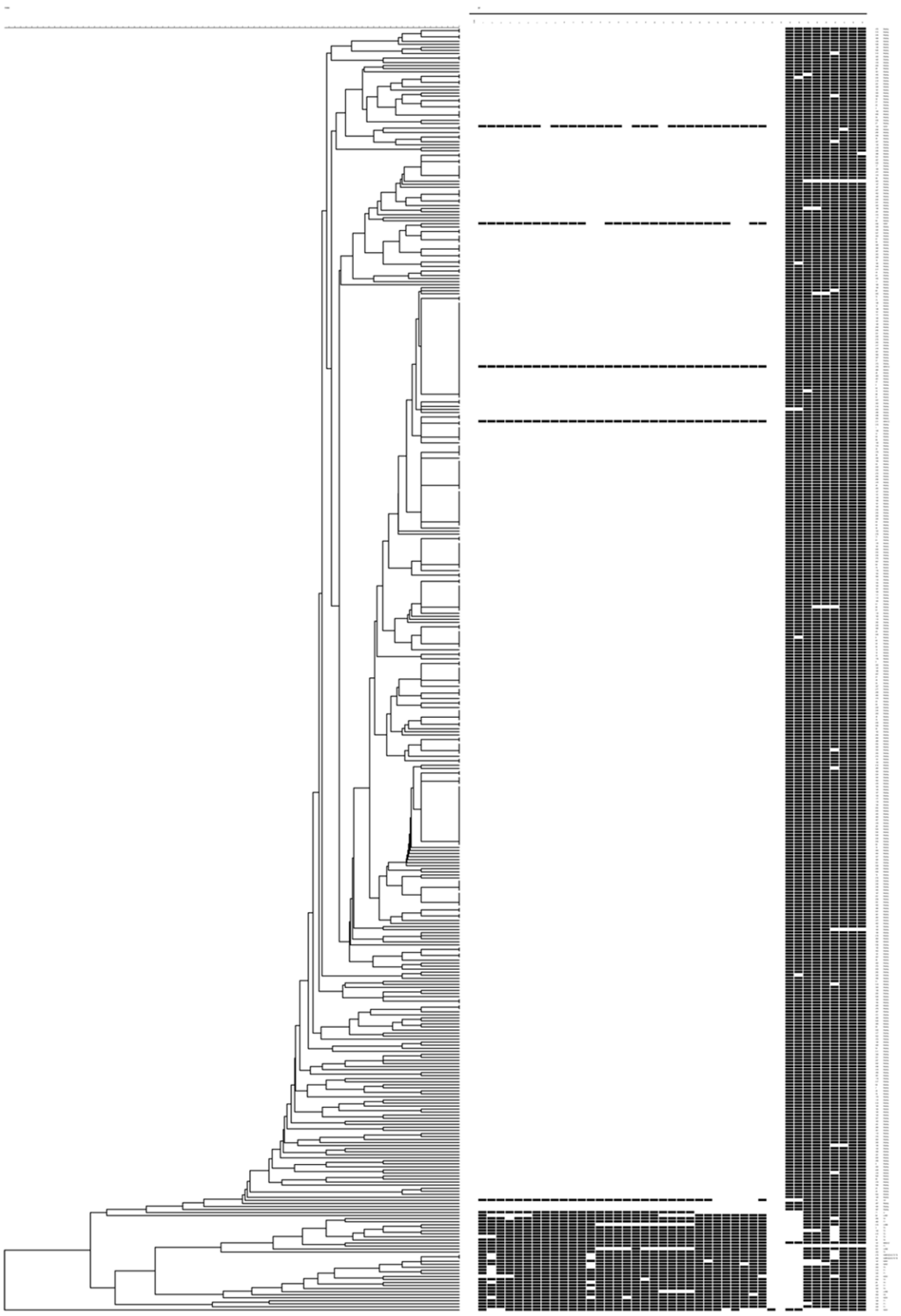

Figure S2. Molecular typing of 422 *M. tuberculosis* strains with VNTR-15$_{China}$ and spoligotyping. From left to right: UPGMA dendrogram generated by VNTR-15$_{China}$ method, spoligotyping pattern, Spoligo-International-Type (SIT) number, genetic lineage based on SpolDB4.0, strain numbers.

Table S3. HGDI scores of the VNTR-15$_{Standard}$ loci and VNTR-15$_{China}$ loci for all strains

Locus	HGDI	95% CI	No. of alleles	VNTR-15$_{standard}$	VNTR-15$_{China}$
QUB11b	0.729	0.718-0.741	9	√	
QUB26	0.676	0.655-0.696	15	√	
QUB4156	0.624	0.606-0.643	6	√	
Mtub04	0.604	0.585-0.623	6	√	
Mtub21	0.602	0.581-0.623	8	√	√
MIRU26	0.579	0.555-0.603	10	√	√
ETRE	0.560	0.538-0.582	6	√	√
ETRA	0.364	0.341-0.388	5	√	√
MIRU10	0.361	0.335-0.387	6	√	√
MIRU40	0.301	0.274-0.327	6	√	√
MIRU16	0.248	0.223-0.273	5	√	√
MIRU39	0.214	0.190-0.237	5	√	√
Mtub39	0.211	0.187-0.236	8		√
ETRD	0.211	0.186-0.235	7	√	√
MIRU27	0.202	0.179-0.225	3		√
Mtub30	0.179	0.156-0.201	5	√	√
ETRB	0.121	0.101-0.140	3		√
MIRU23	0.083	0.066-0.100	5		√
ERTC	0.065	0.050-0.080	4	√	√

Figure S3. Map of China showing the proportion of Beijing family strains in previous studies. The red star showed the location of the Capital of China, Beijing, and the blue point showed the location of Shijiazhuang.