

四川盆地结核分枝杆菌群体遗传学特征研究

李颖 付育红 原梅 董玲玲 黄海荣 李卫民 高基民

【摘要】 目的 探讨四川盆地(重庆、四川地区)流行的结核分枝杆菌(MTB)群体遗传学特征。方法 对413株源自四川盆地的MTB进行长序列多态性(LSP)谱系鉴定,采用 χ^2 检验分析不同谱系MTB在人群中的分布差异。对菌株进行15位点数目可变串联重复序列基因分型,评价各位点的分辨力。构建N-J树分析其系统进化;构建最小跨度树以模拟群体菌株的遗传结构;采用分子方差分析不同谱系MTB的群体间遗传分化;采用贝叶斯模型计算菌株最近共祖年代。结果 通过LSP可将413株MTB分为两大谱系,其中北京谱系占56.2%(232/413),欧美谱系占43.8%(181/413)。两谱系MTB在人群中的分布(性别、年龄、民族以及菌株耐药性)差异无统计学意义($P>0.05$)。北京谱系MTB的N-J树呈放射状分布,最小跨度树中有72.4%(168/232)的菌株被划分到同一“克隆复合群”;欧美谱系MTB的N-J树呈枝状分布,最小跨度树中呈现出多个“克隆复合群”。北京谱系MTB在该地区有显著的群体间遗传分化($F_{ST}=0.018\ 91, P<0.05$),而欧美谱系MTB群体间遗传分化不显著($F_{ST}=0.005\ 19, P>0.05$)。欧美谱系MTB中最大菌株复合群的最近共祖年代为723(95%CI:517~946)年。结论 北京谱系和欧美谱系MTB在四川盆地以“互相对抗”形式流行;两谱系MTB的群体间遗传分化存在差异;欧美谱系MTB的侵入可能与一次历史事件(约720年前的战争)有关。

【关键词】 结核分枝杆菌;长序列多态性;欧美谱系;群体遗传学

Study on the population-genetics of *Mycobacterium tuberculosis* from Sichuan Basin in China Li Ying¹, Fu Yuhong², Yuan Mei¹, Dong Lingling², Huang Hairong², Li Weimin², Gao Jimin¹. 1 Zhejiang Provincial Key Laboratory for Technology and Application of Model Organisms, School of Laboratory Medicine and Life Science, Wenzhou Medical University, Wenzhou 325035, China; 2 National Tuberculosis Clinical Laboratory of China, Beijing Tuberculosis and Thoracic Tumor Research Institute, Capital Medical University Affiliated Beijing Chest Hospital

Corresponding authors: Gao Jimin, Email: jimingao@yahoo.com; Li Weimin, Email: lwm_18@hotmail.com

This work was supported by grants from the National Natural Science Foundation of China (No. 81273144) and High Level Health Technical Personnel Training Plan of Beijing Health System (No. 20133041).

【Abstract】 **Objective** To explore the population-genetics characteristics of *Mycobacterium tuberculosis* (MTB) prevailing at the Sichuan basin of China. **Methods** A total of 413 MTB strains collected from Sichuan basin were genotyped by large sequence polymorphism (LSP) and 15 loci variable number tandem repeat (VNTR). Difference between the distribution of lineage population was analyzed by χ^2 -test and the discriminatory ability of each VNTR locus was evaluated under the Hunter-Gaston Discriminatory Index (HGI). Both phylogeny on population level and genetic structure were demonstrated through N-J tree and the Minimal Spanning Tree (MST). Genetic differentiation of different lineage strains was analyzed by Analysis of Molecular Variance (AMOVA). Time of the Most Recent Common Ancestor (TMRCA) was calculated based on the Bayesian model. **Results** Four hundred and thirteen MTBs were divided into two major lineages, in which the Beijing lineage accounted for 56.2% (232/413) and the Euro-American lineage for 43.8% (181/413). There was no

DOI: 10.3760/cma.j.issn.0254-6450.2015.04.017

基金项目:国家自然科学基金(81273144);北京市卫生系统高层次卫生技术人才培养计划项目(20133041)

作者单位:325035 温州医科大学检验医学院生命科学院 浙江省模式生物技术与应用重点实验室(李颖、原梅、高基民);首都医科大学附属北京胸科医院北京市结核病胸部肿瘤研究所 国家结核病临床实验室(付育红、董玲玲、黄海荣、李卫民)

通信作者:高基民, Email: jimingao@yahoo.com; 李卫民, Email: lwm_18@hotmail.com

significant difference of population distribution between the two lineages ($P>0.05$). The N-J tree of Beijing lineage MTB presented distinctly "star-like" and 72.4% strains were grouped to one clonal complex in MST. The Euro-American lineages MTB presented "branch-like" in N-J tree and were grouped into multiple clonal complexes in MST. There was significant genetic differentiation in Beijing lineage MTBs between Chongqing and Sichuan ($F_{ST}=0.018\ 91, P<0.05$), but not in the Euro-American lineage MTB ($F_{ST}=0.005\ 19, P>0.05$). TMRCA of the largest clonal complex in Euro-American lineage MTB appeared to be 723 (95% CI: 517-946) years. **Conclusion** Both Beijing lineage and Euro-American lineage MTBs were competitively prevalent in the Sichuan basin. There was some difference noticed between the two lineages referring to genetic differentiation. The invasion of Euro-American lineages MTB to the basin area might be associated with a war occurred in this area about 720 years ago.

【Key words】 *Mycobacterium tuberculosis*; Large Sequence Polymorphism; Euro-American lineage; Population genetics

结核分枝杆菌(MTB)是结核病的病原菌。据WHO统计^[1],我国结核病患病人数及耐药菌感染人数均居全球第二,而我国西南地区结核病疫情较严重,其中四川盆地(重庆、四川地区)由于其独特的人文地理位置,结核病疫情对整个地区的影响不容忽视。前期Li等^[2]采用Spoligotyping分型技术将四川地区292株MTB分为北京谱系(54.8%)和非北京谱系(45.2%)两个大群。但有关该地区MTB群体遗传结构以及群体间遗传分化的研究未见报道。为此本研究采用长序列多态性(large sequence polymorphism, LSP)和15位点数目可变串联重复序列(variable number tandem repeat, VNTR)两种基因分型技术,对收集自四川盆地的413株MTB进行群体遗传学特征分析,探讨该地区MTB的流行特征。

材料与方法

1. 实验菌株:432株MTB取自2007—2008年全国结核病耐药基线调查中重庆市和四川省分离株^[3],由国家结核病临床实验室提供。其中重庆市包括丰都县(采样点收集60株/本研究纳入60株,60/60)、奉节县(67/67)2个采样点;四川省包括峨眉山市(60/54)、通江县(62/62)、武胜县(65/61)、江油市(58/53)、涪城区(60/56)5个采样点。因19株MTB复苏未成功而舍弃,最终413株(95.6%)MTB纳入研究。

2. 试剂和仪器:2×Taq Master Mix、100 bp Ladder以及RD105鉴定北京谱系MTB试剂盒均购自北京康为世纪生物科技有限公司;Bio Numerics 5.0分析软件购自比利时Applied Maths公司。

3. LSP鉴定谱系:参照文献[4-5]选定RD105、RD239、RD702、RD711、RD750以及pks15/1Δ7 bp共6个DNA片段作为MTB谱系定义靶标,针对每一RD片段分别进行PCR扩增,以检测相应片段有无

缺失鉴定MTB谱系。其中北京谱系MTB的鉴定采用RD105商品化试剂盒;对经鉴定的北京谱系MTB进一步行RD181为靶标的“古老”或“现代”亚谱系分型^[6]。pks15/1Δ7 bp引物设计参照文献[5],扩增该目的基因可分别获得398 bp(缺失7个碱基)和405 bp(完整)两种片段长度的扩增产物,欧美谱系特异性缺失5'-GCCGCGG-3' 7个碱基。通过基因测序比对完成欧美谱系MTB的鉴定。一株菌对应一名病例,采用 χ^2 检验分析该地区不同谱系MTB在人群性别、年龄、民族以及耐药情况上的差异是否具有统计学意义($P<0.05$)。

4. VNTR基因分型:采用标准15位点VNTR基因分型(VNTR-15)方法^[7]。计算VNTR-15对不同谱系MTB的总体分辨力以及各VNTR位点的分辨力(汉高分辨指数 Hunter-Gaston Discriminatory Index, HGI)^[8]。

5. 群体遗传学分析:将413株MTB的LSP和VNTR数据一并上传至MIRU-VNTR数据库(<http://www.miru-vntrplus.org>)与代表菌株进行比对分析。结合VNTR-15数据,以*M. prototuberculosis*(包含*M. canettii*)为外群构建N-J进化树;采用Bio Numerics 5.0软件构建最小跨度树(minimum spanning tree, MST)。本研究将该地区不同谱系MTB分别设定为一个群体,采用Arlequin 3.5软件通过分子方差分析(analysis of molecular variance, AMOVA)计算成对群体间遗传分化系数(F_{ST})^[9],推测对应群体之间是否存在显著的遗传分化。

6. 欧美谱系MTB最近共祖年代的计算:前期Comas等^[10]对全球范围内主要谱系MTB的起源进化以及共祖年代进行较细致的研究,推测北京谱系MTB起源于东亚地区,欧美谱系MTB起源于欧洲大陆。Pepperell等^[11]推导了DS6_{QUE}欧美亚谱系MTB通过法国人皮毛贸易由加拿大东海岸登陆并跟随贸易活动向西海岸传播的过程。参照

Pepperell 方法,采用贝叶斯模型,突变速率为 10^{-3} 每位点每代,增长速率为 10^3 每代,利用 Ytime 软件计算该群欧美谱系 MTB 的最近共祖年代(time of the most recent common ancestor, TMRCA)。

结 果

1. LSP 鉴定谱系:北京谱系占收集菌株的 56.2% (232/413), 欧美谱系占 43.8% (181/413), 未发现其他谱系 MTB。两谱系 MTB 与 MIRU-VNTR 数据库中的北京谱系和欧美谱系 MTB 代表菌株汇聚良好。北京谱系 MTB 中“现代”亚谱系所占比例为 90.9% (211/232), “古老”亚谱系仅占 9.1% (21/232)。结合国内现有研究数据,设计适合我国 MTB 的 LSP 谱系鉴定流程图(图 1)。经 χ^2 检验验证,该地区两谱系 MTB 在人群性别、年龄、民族分布以及耐药上的差异无统计学意义($P > 0.05$)(表 1)。

表 1 两谱系 MTB 在不同特征人群中的分布

特征	北京谱系株数	欧美谱系株数	χ^2 值	P 值
性别			0.518	>0.05
男	167	136		
女	65	45		
年龄组(岁)			1.398	>0.05
1~	20	16		
21~	83	62		
41~	82	64		
≥60	41	39		
民族			1.408	>0.05
汉	231	177		
其他	1	4		
药敏状态			1.239	>0.05
敏感	136	112		
耐药	76	59		
耐多药	20	10		

2. VNTR-15 基因分型的分辨力:VNTR-15 将 232 株北京谱系 MTB 和 181 株欧美谱系 MTB 分为 210 和 175 个基因型, HGI 分别为 0.998 6 和 0.999 6。QUB11b、QUB26、Mtub21、MIRU26 位点对该地区北京谱系和欧美谱系 MTB 均具有较高的分辨力(HGI > 0.5); MIRU04、MIRU40、MIRU16、ETR-A、Mtub39 位点对该地区欧美谱系的分辨力均高于北京谱系 MTB(表 2)。

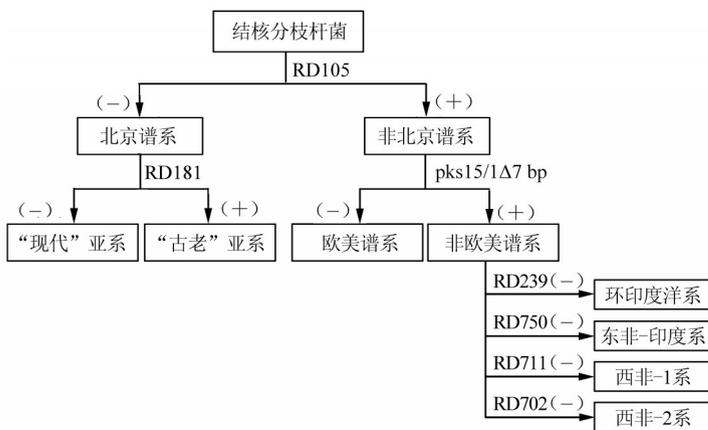
表 2 四川盆地不同谱系 MTB 中 VNTR-15 各位点的 HGI 比较

VNTR 位点	北京谱系 (n=232)	欧美谱系 (n=181)	合计 (n=413)
Mtub 04	0.495	0.589	0.528
ETR-C	0.059	0.107	0.082
MIRU 04	0.132	0.420	0.128
MIRU 40	0.294	0.668	0.349
MIRU 10	0.396	0.501	0.437
MIRU 16	0.223	0.478	0.244
Mtub 21	0.663	0.545	0.615
QUB 11b	0.768	0.821	0.795
ETR-A	0.274	0.464	0.313
Mtub 30	0.239	0.313	0.274
MIRU 26	0.627	0.706	0.674
MIRU 31	0.393	0.401	0.369
Mtub 39	0.266	0.645	0.374
QUB 26	0.633	0.803	0.755
QUB 4156	0.285	0.416	0.342

注:黑体数字为 VNTR 位点对该谱系 MTB 具有较高的分辨能力

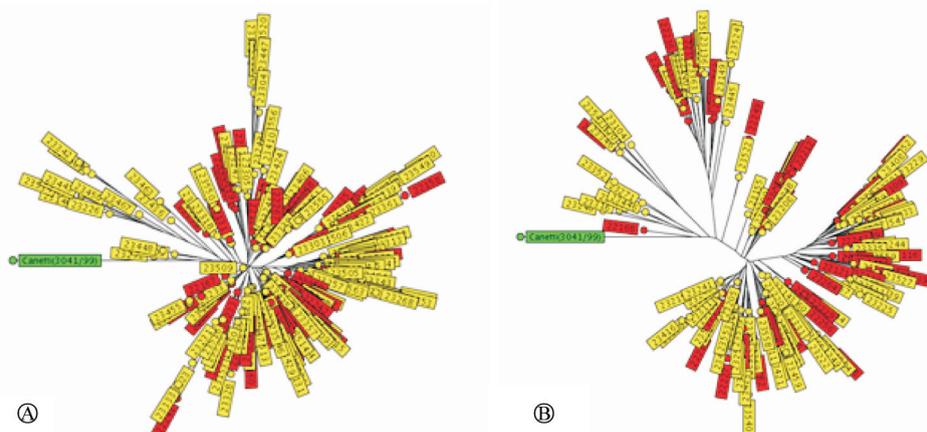
3. 系统发育分析:232 株北京谱系 MTB 的 N-J 树呈“放射状”分布(图 2A), 提示这些菌株可能由某一“克隆”进化而来; MST 中有 72.4% (168/232) 的菌株被划分到同一“克隆复合群”(图 3A), 表明该群北京谱系 MTB 的克隆扩展能力较强。欧美谱系 MTB 的 N-J 树呈“枝状”分布(图 2B), 提示这些菌株可能是“多克隆”起源; MST 显示有多个较明显的“克隆复合群”, 其中最大的克隆复合群包含 39.2% (71/181) 的菌株(图 3B)。

4. 群体菌株间的分子方差分析:就北京谱系 MTB 而言,重庆、四川两群菌株群体间遗传分化较显著($F_{ST} = 0.018\ 91, P < 0.05$); 就欧美谱系 MTB 而言,重庆、四川两群菌株群体间遗传分化不显著($F_{ST} = 0.005\ 19, P > 0.05$)。以全部菌株为总体,计算北京谱系与欧美谱系 MTB 之间的 F_{ST} , 得到显著的 P 值($F_{ST} = 0.304\ 09, P = 0.000$), 两谱系 MTB 群体间遗传分化显著。



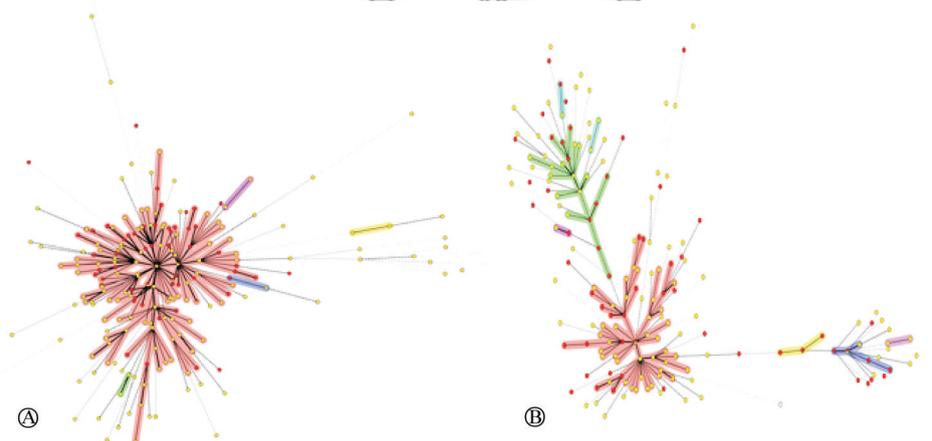
注: -, 缺失, +, 存在

图 1 LSP 谱系鉴定流程



注:绿色为外群参考菌株 *M. prototuberculosis* (含 *M. canettii*); 红色代表重庆地区菌株; 黄色代表四川地区菌株

图2 232株北京谱系MTB(A)和181株欧美谱系MTB(B)基于VNTR-15的系统发育N-J树



注:圆点示VNTR-15基因型,其大小对应拥有该基因型菌株的数量,其中黄色代表四川地区菌株,红色代表重庆地区菌株;不同的背景颜色代表不同的克隆复合群

图3 232株北京谱系MTB(A)和181株欧美谱系MTB(B)基于VNTR-15的MST

5. 针对欧美谱系MTB的TMRCA计算:选定欧美谱系MTB中最大的菌株复合群(共71株,处于N-J树同一进化分支,在MST中属于同一克隆复合群)进行共祖年代计算,得到TMRCA为723(95%CI:517~946)年。

讨 论

LSP是基于MTB基因组中分化区(region of difference, RD)的缺失而建立的谱系鉴定技术,通过该技术可将全球的MTB分为6大谱系,即环印度洋系、北京谱系、欧美谱系、东非-印度系、西非-1系、西非-2系,且每个谱系MTB均与对应的地理人群存在着一定联系^[5,12]。本研究采用LSP对四川盆地的413株MTB进行谱系鉴定,得到北京(56.2%)和欧美(43.8%)两个谱系。结合我国MTB的总体流行情况,北京谱系MTB是我国主流流行菌(流行比例约为62.2%,南北地区略有不同^[13]),欧美谱系MTB在我国亦有一定程度的流行,而西非谱系MTB几乎未

在我国发现^[12],因此设计了适合我国MTB谱系鉴定的LSP流程图。该地区北京谱系与欧美谱系MTB在人群年龄、性别、民族分布以及耐药性上的差异无统计学意义,提示欧美谱系MTB与当地宿主人群的联系和北京谱系MTB基本一致。本研究表明,在四川盆地欧美谱系与北京谱系MTB呈“相互对抗”式的流行,与东、中部地区北京谱系MTB高水平的流行特征不同^[14-15];欧美谱系MTB和该地区宿主人群的联系与Gagneux等“其与欧美人群存在稳定联系^[12]”的结论不同。因此有必要以群体遗传学为切入点,探讨导致该地区较为特殊的MTB流行现状的原因。

本研究中VNTR-15对该地区两谱系MTB的总体分辨力均较高。北京谱系MTB在该地区呈现出较强的克隆扩展状态;而欧美谱系MTB则趋向于多克隆、多中心起源状态。前期研究表明北京谱系MTB无论是在东亚还是在美洲大陆等地区其菌株群均呈克隆状态扩展^[14,16];欧美谱系MTB则由于谱系内亚型较多以及欧美国家人口密度较小等因素导

致其多呈枝状扩展^[17];可见该地区流行的北京谱系和欧美谱系MTB的群体遗传结构并未发生明显改变。

本研究将四川盆地分成重庆、四川两个地区进行分析,以期能够发现该地区北京谱系和欧美谱系MTB的群体间遗传分化是否存在差异。分子方差分析显示,北京谱系MTB在重庆、四川两地区间存在较显著的群体间遗传分化,而欧美谱系MTB的群体间遗传分化不显著。前期研究表明,北京谱系MTB起源于东亚,早在新石器时代(约8 000年前)就已在我国出现,且随着现代人的不断繁衍迁移在东亚大范围扩展^[10];北京谱系MTB在接下来的历史长河中可能先后遇到了不同的自然选择压力和不同的群体内遗传漂变以及有限范围内的基因漂流等多重因素,从而导致不同地区菌株间出现了较显著的遗传分化。欧美谱系MTB起源于欧洲大陆,且遗传多样性较大^[10],这一特点使其作为外来物种能够更好地适应新的环境,可能是其成功入侵定植的原因之一;欧美谱系MTB作为入侵物种可能在随后较短的时间内遇到了较为一致的自然选择压力以及来自当地北京谱系MTB的“挤压”,从而导致不同地区菌株间遗传分化不显著。分析该地区欧美谱系MTB主要菌株复合群的最近共祖年代为720年,结合该地区特殊的地理历史(重庆市合川区钓鱼城)^[18],本研究据此推测:700多年前的合川钓鱼城战争,即公元1260年(宋末元初)蒙古铁骑从征战中的欧洲大陆收兵,全力攻打巴蜀咽喉合川钓鱼城,战争持续36年之久,而蒙古骑兵很可能也将欧洲大陆较为流行的欧美谱系MTB带到该地区,并此后随人类活动传播扩展。

我国由于北京谱系MTB流行程度高,因此忽视了流行程度较低但在欧美国家呈高水平流行的欧美谱系MTB的研究。本研究以四川盆地北京谱系和欧美谱系MTB呈“对抗”式流行现象为切入点,设计了适合我国MTB的LSP谱系鉴定流程图,并发现该地区北京谱系与欧美谱系MTB在群体遗传结构以及群体间遗传分化上存在差异;计算最近共祖年代推测出欧美谱系MTB可能的入侵年代。此结果有助于探讨北京谱系、欧美谱系乃至整个MTB复合群的生物进化机制,为当地乃至全国结核病防治提供科学依据。

参 考 文 献

- [1] WHO. Global tuberculosis control: a short update to the 2009 report [R]. Global Tuberculosis Control, Geneva: World Health Organization, 2009.
- [2] Li X, Xu P, Shen X, et al. Non-Beijing strains of *Mycobacterium tuberculosis* in China [J]. J Clin Microbiol, 2011, 49(1):392-395.
- [3] The Ministry of Health of the People's Republic of China. The national TB drug resistance baseline survey report (2007-2008) [M]. Beijing: People's Medical Publishing House, 2010: 1-230. (in Chinese)
中华人民共和国卫生部. 全国结核病耐药性基线调查报告(2007-2008年)[M]. 北京:人民卫生出版社, 2010: 1-230.
- [4] Gagneux S, DeRiemer K, Van T, et al. Variable host pathogen compatibility in *Mycobacterium tuberculosis* [J]. Proc Natl Acad Sci USA, 2006, 103(8):2869-2873.
- [5] Martinez-Gamboa A, Ponce-de-Leon A, Galindo-Fraja A, et al. Molecular analysis of *Mycobacterium tuberculosis* strains with an intact *pkb15/1* gene in a rural community of Mexico [J]. Arch Med Res, 2008, 39(8):809-814.
- [6] Tsolaki AG, Gagneux S, Pym AS, et al. Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis* [J]. J Clin Microbiol, 2005, 43(7):3185-3191.
- [7] Supply P, Allix C, Lesjean S, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit VNTR typing of *Mycobacterium tuberculosis* [J]. J Clin Microbiol, 2006, 44(12):4498-4510.
- [8] Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity [J]. J Clin Microbiol, 1988, 26(11):2465-2466.
- [9] Weir BS, Hill WG. Estimating F-statistics [J]. Ann Rev Genet, 2002, 36:721-750.
- [10] Comas I, Coscolla M, Luo T, et al. Out-of-Africa migration and Neolithic co-expansion of *Mycobacterium tuberculosis* with modern humans [J]. Nat Genet, 2013, 45(10):1176-1182.
- [11] Pepperell CS, Granka JM, Alexander DC, et al. Dispersal of *Mycobacterium tuberculosis* via the Canadian fur trade [J]. Proc Natl Acad Sci USA, 2011, 108(16):6526-6531.
- [12] Hirsh AE, Tsolaki AG, DeRiemer K, et al. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations [J]. Proc Natl Acad Sci USA, 2004, 101(14):4871-4876.
- [13] Pang Y, Zhou Y, Zhao B, et al. Spoligotyping and drug resistance analysis of *Mycobacterium tuberculosis* strains from national survey in China [J]. PLoS One, 2012, 7(3):e32976.
- [14] Ding PJ, Hui M, Liang Q, et al. Origin, phylogeny, and spread of *Mycobacterium tuberculosis* Beijing lineage in the five provinces of northern China [J]. Chin J Epidemiol, 2013, 34(4):374-378. (in Chinese)
丁朋举, 惠明, 梁倩, 等. 中国北方部分地区北京谱系结核分枝杆菌群体遗传关系的分析 [J]. 中华流行病学杂志, 2013, 34(4):374-378.
- [15] Dai GM, Zhang ZG, Ding PJ, et al. Differences in the population of genetics of *Mycobacterium tuberculosis* between urban migrants and local residents in Beijing, China [J]. Chin Med J, 2013, 126(21):4066-4071.
- [16] Glynn JR, Whiteley J, Bifani PJ, et al. Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review [J]. Emerg Infect Dis, 2002, 8(8):843-849.
- [17] Rindi L, Medici C, Bimbi N, et al. Genomic variability of *Mycobacterium tuberculosis* strains of the Euro-American lineage based on large sequence deletions and 15-locus MIRU-VNTR polymorphism [J]. PLoS One, 2014, 9(11):e114676.
- [18] Xiao W. The miracle in cold war period: the war of Hechuan Fishing Town [J]. Hist Refer, 2011(1):60-63. (in Chinese)
小兀. 冷兵器战争的奇迹:合川钓鱼城之战 [J]. 文史参考, 2011(1):60-63.

(收稿日期:2014-10-04)

(本文编辑:张林东)