

不同可变数目串联重复序列组合对中国流行结核分枝杆菌分辨力的评价研究

陈海霞 蔡超 刘静仪 张治国 原梅 贾俊楠 孙照刚 黄海荣 高基民
李卫民

325035 温州医科大学检验医学院生命科学学院 浙江省模式生物技术与应用重点实验室(陈海霞、刘静仪、高基民); 101149 北京市结核病胸部肿瘤研究所北京市结核病耐药重点实验室(蔡超); 101149 北京,首都医科大学附属北京胸科医院国家结核病临床实验室(陈海霞、刘静仪、贾俊楠、孙照刚、黄海荣、李卫民); 102200 北京市昌平区结核病防治所检验科(张治国); 044000 山西省运城市急救中心检验科(原梅); 100069 北京,首都医科大学公共卫生学院临床流行病学北京市重点实验室(贾俊楠、李卫民)

陈海霞、蔡超同为第一作者

通信作者: 高基民, Email: jimingao64@163.com; 李卫民, Email: lwm_18@aliyun.com

DOI: 10.3760/cma.j.issn.0254-6450.2017.06.021

【摘要】 目的 在规范化的结核分枝杆菌(*Mycobacterium tuberculosis*, MTB)可变数目串联重复序列(variable number of tandem repeats, VNTR)基因分型的基础上,构建我国31个省(自治区、直辖市)VNTR数据库,每个省优化一套VNTR位点组合,为我国结核病预防控制策略的制定提供科学依据。**方法** 对2007—2008年全国结核病耐药性基线调查的4 116株MTB15位点VNTR(15-VNTR)基因分型。汉高指数(Hunter-Gaston Index, HGI)分析每个位点的分辨率。依据谱系流行特征,以分辨率高和稳定性强为原则,为各省设计一套VNTR优化组合(12-VNTR、10-VNTR、8-VNTR和5-VNTR),采用HGI和成簇率进行评价。**结果** 完成了覆盖率为96.36%(3 966/4 116)MTB完整15-VNTR图谱。发现QUB11b、MIRU26等7个高分辨率位点; QUB26、MIRU16、Mtub21、QUB11b在部分地区遗传稳定性差。内蒙古自治区、重庆市、黑龙江省的最优组合为10-VNTR,其他各省的最佳组合为8-VNTR。**结论** VNTR数据库的建立将推动全国范围MTB传染源的追踪;各省优化VNTR组合的推出有助于当地结核病疫情的监测和群体遗传学的研究。

【关键词】 结核分枝杆菌; 可变数目串联重复序列; 基因分型; 分辨率; 成簇率

基金项目: 国家自然科学基金(81273144); 北京市自然科学基金B类重点项目(KZ201510025024); 北京市结核病临床数据和样本资源库项目(D131100005313012)

Discriminatory power of variable number on tandem repeats loci for genotyping *Mycobacterium tuberculosis* strains in China Chen Haixia, Cai Chao, Liu Jingyi, Zhang Zhiguo, Yuan Mei, Jia Junnan, Sun Zhaogang, Huang Hairong, Gao Jimin, Li Weimin

Zhejiang Provincial Key Laboratory for Technology, Application of Model Organisms, School of Laboratory Medicine and Life Science, Wenzhou Medical University, Wenzhou 325035, China (Chen HX, Liu JY, Gao JM); The Drug-resistant TB Key Laboratory of Beijing, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing 101149, China (Cai C); National Tuberculosis Clinical Laboratory of China, Beijing Chest Hospital, Capital Medical University, Beijing 101149, China (Chen HX, Liu JY, Jia JN, Sun ZG, Huang HR, Li WM); The Institute of Tuberculosis Prevention and Control of Changping District, Beijing 102200, China (Zhang ZG); Yuncheng City Emergency Center, Yuncheng 044000, China (Yuan M); Beijing Municipal Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing 100069, China (Jia JN, Li WM)

Chen Haixia and Cai Chao are the first authors who contributed equally to the article.

Corresponding authors: Gao Jimin, Email: jimingao64@163.com; Li Weimin, Email: lwm_18@aliyun.com

【Abstract】 Objective Using the standard genotype method, variable number of tandem repeats (VNTR), we constructed a VNTR database to cover all provinces and proposed a set of

optimized VNTR loci combinations for each province, in order to improve the preventive and control programs on tuberculosis, in China. **Methods** A total of 15 loci VNTR was used to analyze 4 116 *Mycobacterium tuberculosis* strains, isolated from national survey of Drug Resistant Tuberculosis, in 2007. Hunter-Gaston Index (HGI) was also used to analyze the discriminatory power of each VNTR site. A set combination of 12-VNTR, 10-VNTR, 8-VNTR and 5-VNTR was respectively constructed for each province, based on 1) epidemic characteristics of *M. tuberculosis* lineages in China, with high discriminatory power and genetic stability. **Results** Through the completed 15 loci VNTR patterns of 3 966 strains under 96.36% (3 966/4 116) coverage, we found seven high HGI loci (including QUB11b and MIRU26) as well as low stable loci (including QUB26, MIRU16, Mtub21 and QUB11b) in several areas. In all the 31 provinces, we found an optimization VNTR combination as 10-VNTR loci in Inner Mongolia, Chongqing and Heilongjiang, but with 8-VNTR combination shared in other provinces. **Conclusions** It is necessary to not only use the VNTR database for tracing the source of infection and cluster of *M. tuberculosis* in the nation but also using the set of optimized VNTR combinations in monitoring those local epidemics and *M. tuberculosis* (genetics in local) population.

【Key words】 *Mycobacterium tuberculosis*; Variable number of tandem repeats; Enotype; Discriminatory; Clustered rate

Fund programs: National Natural Science Foundation of China (81273144); Beijing Natural Science Foundation Program and Scientific Research Key Program of Beijing Municipal Commission of Education (KZ201510025024); Beijing Bio-Bank of Clinical Resources-Tuberculosis, BBCR-Tuberculosis (D131100005313012)

据统计,2014年我国结核病(tuberculosis, TB)分别有930 000例新发病例和38 000例死亡病例, TB例数居世界第三,是我国所面临的一项严峻公共卫生问题^[1]。结核基因分型是分子流行病学的核心技术。目前国际上使用较为广泛的结核分枝杆菌(*M. tuberculosis*, MTB)基因分型方法主要包括:可变量串联重复序列(variable number of tandem repeats, VNTR)以及目前逐渐被广泛使用的长序列多态性(large sequence polymorphism, LSP)、单核苷酸多态性(single nucleotide polymorphism, SNP)和全基因组测序。其中,全基因组测序作为一项新兴技术,因其快速、准确、高效的优点,在结核病基础研究中发挥了重要作用^[2-4];但是其昂贵的价格以及复杂的分析方法使之在我国的结核病预防控制领域广泛推广尚需时日。VNTR分型技术因其操作简便,分析快速以及便于跨实验室比较等优点,对暴发流行、近期传播、疫情监测等结核流行病学实际工作具有指导意义。目前国际上使用较广泛的有15-VNTR及24-VNTR^[5-6]。但VNTR分型技术具有同源异质性的缺陷。因此采用LSP和SNP先行谱系定义,再行VNTR分型是目前MTB基因分型的一种高效策略。Luo等^[7]曾在采用SNP鉴定北京谱系,及其“现代”和“古老”亚谱系MTB的基础上,评价了几个对其分辨率较高的VNTR位点,并以此为工具在我国开展了卓有成效的分子流行病学研究工作^[8-9]。但是除北京谱系外,我国还有约40%的其他谱系MTB的流行,有必要找出一套适合我国结核病预防控制的最优VNTR分型方案。

本研究对2007年全国结核病耐药性基线调查

收集的31省4 116株MTB 15-VNTR基因分型,汉高指数(Hunter-Gaston Index, HGI)评估各地区VNTR位点的分辨力,并根据分辨力及遗传稳定性排列分析,为各省推出一套高效的VNTR位点组合,筛选出适合我国国情的MTB基因分型策略,并将其应用于结核病预防控制的实际工作。

材料与方法

1. 材料:

(1)实验菌株:2007年全国结核病耐药性基线调查涵盖中国31个省(自治区、直辖市),共收集4 116株MTB。经鉴定均为MTB^[10-11]。标准菌株H37Rv由北京胸科医院国家结核病临床实验室提供。

(2)主要试剂和仪器:2×Taq Master Mix、100 bp Ladder均购自北京康为世纪生物科技有限公司;15-VNTR位点引物序列参照文献[6],由北京擎科生物公司合成;Bio Rad凝胶成像仪以及Quantity One软件购自美国Bio Rad生命科学公司。

2. 实验室方法:

(1)DNA提取:①煮沸法:吸取200 μl冻存菌液于无菌EP管,100℃水浴20~30 min,12 000 r/min离心4 min,上清液即含DNA。②酚-氯仿法:对于扩增失败的菌株,吸取50 μl冻存菌液接种于中性罗氏培养基斜面,37℃恒温培养4周;待长出菌落,用接种环轻轻刮取细菌,100℃,30 min灭活。标准酚-氯仿法提取。

(2)VNTR基因分型:采用标准的15-VNTR方法。PCR扩增条件:94℃预变性5 min;94℃变性30 s,65℃退火30 s,72℃延伸45 s,30~35个循环;延伸72℃10 min;PCR扩增产物由1.2%琼脂糖凝

胶分离, Bio-Rad 凝胶成像仪系统上处理分析, 100 bp DNA Ladder 作为参照。

3. 分析方法:

(1) 样本收集质控: 复苏或 PCR 扩增失败(经过重复 2 次 PCR 扩增有 ≥ 5 个 VNTR 位点失败)的菌株不纳入本研究。

(2) 实验室质控: 实验设立对照菌株 H37Rv 在同一 VNTR 位点的扩增片段大小是否相同。从检测标本中随机挑 10 个菌株, 每株不同 VNTR 位点重复检测 3 次, 比较每次扩增片段是否一致。

(3) 结果分析: Quantity one 软件分析每个 VNTR 位点的 PCR 结果, H37Rv 为参照标准。重复单元数目公式: 各 VNTR 位点重复单元数目 = (PCR 产物片段长度 - VNTR 位点侧翼片段长度) / 重复单元片段长度。

(4) HGI 计算: 计算各 VNTR 位点以及不同 VNTR 位点组合的分辨率^[12]。公式:

$$HGI = 1 - \frac{1}{N(N-1)} \sum_{j=1}^S n_j(n_j-1)$$

式中 N 代表样本菌株总数; S 代表基因型总数; n_j 代表拥有第 j 种基因型的菌株总数。

由于不同地区同一 VNTR 位点偶尔存在极端值的情况, 因此将 VNTR 位点 HGI 的中位数值 (M) 用来估算其对全国范围内 MTB 的总分辨率。

(5) VNTR 优化组合: 针对每个省份 MTB 谱系流行特征, 依据参考文献[13]提示谱系不同, VNTR 分辨率不同。15 个 VNTR 位点根据分辨率不同划分为高、中、低 3 个水平, 高、中分辨率且稳定性好的 VNTR 位点作为一线候选位点, 分辨率低 ($HGI > 0.1$) 但稳定性好的位点作为二线候选位点, 分辨率高但稳定性差(该地区扩增失败的菌株数超过该地区总菌株数的 5%)位点则舍弃。并以 15-VNTR 的分辨率及成簇率为参考, 为每个省份均分别排列推出 12-VNTR、10-VNTR、8-VNTR、5-VNTR^[14]。并采用 HGI 和成簇率进行评估, 成簇率 = (成簇的菌株数 - 成簇基因型数) / 总菌株数。

结 果

1. 纳入分析菌株及重复性:

(1) 菌株收集: 2007 年全国结核病耐药性基线调查所收集的 4 116 株 MTB, 150 株因年久缺失、复苏或 PCR 扩增失败, 最终 3 996 株 MTB 纳入本研究。

(2) 检测结果的重复性: 对照菌株 H37Rv 在同一 VNTR 位点的扩增片段大小相同。每株 MTB 在不同 VNTR 位点的 3 次重复检测结果均完全一致。

2. 15 位点 VNTR 分辨率和稳定性: 所有菌株的分型结果提示: QUB26、QUB11b、Mtub21、MIRU16 位点扩增失败的菌株数量较多, 比例分别为 11.85% (470/3 996)、2.67% (106/3 996)、3.08% (122/3 996)、4.06% (122/3 996), 提示这些位点相对其他 VNTR 位点遗传稳定性较差。并发现不同地区同一 VNTR 位点稳定性不同, 同一地区不同 VNTR 位点遗传稳定性也不同(表 1)。其中 QUB26 在广西、广东、湖南、上海、宁夏、山东等 13 省份扩增失败的菌株数超过 5%。QUB11b 在广西、广东、江西、河北 4 省份扩增失败的菌株数超过 5%。Mtub21 在山西、河北、江西 3 省扩增失败的菌株数超过 5%。MIRU16 在湖南、浙江、山西、青海 4 省扩增失败的菌株数超过 5%。比较各个 VNTR 位点的 HGI 中值发现: QUB11b、Mtub21、QUB26、MIRU26、Mtub04、MIRU10、MIRU31、MIRU16 位点 HGI 中值分别高达 0.834、0.762、0.760、0.817、0.757、0.704、0.731、0.662(表 2), 属于高分辨率位点 ($HGI > 0.6$); MIRU40、ETR-A、Mtub30、QUB4156c、MIRU04、Mtub39 属于中分辨率位点 ($0.3 \leq HGI \leq 0.6$), ETR-C 属于低分辨率位点 ($HGI < 0.3$)^[15]。

3. 各省最佳 VNTR 组合位点分析: 31 个省份的 15-VNTR 分辨率均 $> 0.994 3$, 成簇率均低于 10%, 其中除去重庆、贵州、江西、宁夏、甘肃、新疆、吉林、辽宁、黑龙江 9 个省份外, 22 个省份分辨率和成簇率均分别 1 和 0。VNTR 位点在不同省份的分辨率存在不同程度的差异, MIRU10、MIRU40、ETR-A、Mtub30、QUB4156c、MIRU04、Mtub39 位点在个别省份分辨率过低, 而 ETR-C 分辨率则是在全国普遍偏低。高分辨率位点 QUB11b 在某些省份稳定性较差。因此本研究为每省分别筛选 12-VNTR、10-VNTR、8-VNTR、5-VNTR 的 4 套组合(表 4)。发现 QUB26 位点被纳入西部地区 VNTR 组合明显高于东、中部地区, 二线候选位点 ETR-C 被纳入南方地区 VNTR 组合高于北方地区, 相邻地区间各 VNTR 组合通常是相同或仅有 1 个 VNTR 位点的差别。内蒙古、重庆、黑龙江地区选用 10-VNTR 作为最优组合, 其他省份可用 8-VNTR 作为最优组合。

讨 论

基因分型技术能够用于研究结核的暴发、近期传播、克隆扩张以及区分内源性复燃和外源性再感染, 因此是结核病防控的重要手段。发达国家早在 20 世纪 90 年代就建有大型的 MTB 基因型数据库, 在原有的暴发流行和接触调查的基础上开展成簇调

表1 各地区结核分枝杆菌分别在15个VNTR位点扩增失败的菌株数

VNTR位点	未能成功扩增VNTR位点的菌株数(株)
QUB11b	广西(9)广东(16)福建(5)湖南(14)湖北(6)江西(17)江苏(1)浙江(4)上海(3)安徽(4)陕西(2)宁夏(1)山西(3)山东(3)河南(3)河北(10)吉林(1)黑龙江(1)青海(3)
Mtub21	广西(6)福建(1)湖南(6)江西(30)江苏(2)浙江(1)上海(2)安徽(3)山西(8)山东(1)河南(3)河北(52)北京(1)天津(2)吉林(1)黑龙江(1)青海(2)
QUB26	广西(17)广东(13)福建(19)湖南(63)湖北(26)江西(71)江苏(23)浙江(11)上海(12)安徽(47)陕西(3)宁夏(11)山西(25)内蒙古(1)山东(26)河南(47)河北(44)天津(1)辽宁(1)黑龙江(1)青海(2)
MIRU26	湖南(1)安徽(1)宁夏(6)山东(2)河北(2)辽宁(1)
Mtub04	广西(1)海南(2)湖北(1)江西(1)江苏(3)浙江(2)安徽(3)陕西(3)宁夏(11)山西(25)内蒙古(1)山东(2)河南(2)河北(2)吉林(1)辽宁(1)
MIRU10	湖南(1)湖北(2)江西(1)江苏(4)浙江(6)安徽(3)山西(1)山东(7)河南(5)河北(6)青海(1)
MIRU31	福建(1)江苏(1)浙江(1)安徽(2)山东(1)河南(4)河北(14)
MIRU40	浙江(1)山东(1)河北(12)
QUB4156c	广西(1)福建(1)江苏(1)浙江(1)宁夏(1)山东(1)河北(3)黑龙江(1)
ETR-A	广西(3)广东(12)海南(4)福建(2)湖南(6)湖北(6)江苏(1)浙江(1)上海(1)安徽(1)陕西(3)宁夏(4)山西(1)山东(1)河南(6)河北(4)北京(4)天津(1)青海(2)
Mtub30	江苏(1)浙江(1)山东(2)河北(1)吉林(1)
Mtub39	广西(1)宁夏(4)内蒙古(3)山东(1)河南(1)辽宁(2)黑龙江(2)
MIRU16	广西(3)广东(1)福建(3)湖南(26)湖北(7)江西(7)江苏(7)浙江(28)上海(2)安徽(8)宁夏(1)山西(21)山东(10)河南(17)河北(11)北京(1)黑龙江(3)青海(5)
MIRU04	江西(1)安徽(1)山西(1)河北(1)吉林(1)
ETR-C	安徽(1)河北(1)

表2 15个VNTR位点的汉高指数中值

VNTR位点	QUB11b	Mtub21	QUB26	MIRU26	Mtub04	MIRU10	MIRU31	MIRU40	QUB4156c	ETR-A	Mtub30	Mtub39	MIRU16	MIRU04	ETR-C
HGI	0.834	0.762	0.760	0.817	0.757	0.704	0.731	0.563	0.426	0.574	0.504	0.515	0.662	0.344	0.266

表3 不同地区15个位点VNTR的汉高指数

地区	QUB11b	Mtub21	QUB26	MIRU26	Mtub04	MIRU10	MIRU31	MIRU40	QUB4156c	ETR-A	Mtub30	Mtub39	MIRU16	MIRU04	ETR-C
北京	0.824	0.736	0.750	0.780	0.767	0.704	0.65	0.296	0.433	0.655	0.355	0.497	0.649	0.237	0.410
天津	0.763	0.558	0.740	0.769	0.642	0.718	0.662	0.681	0.188	0.374	0.189	0.496	0.544	0.159	0.493
河北	0.832	0.762	0.825	0.781	0.984	0.701	0.715	0.455	0.188	0.389	0.331	0.256	0.805	0.147	0.125
山西	0.829	0.707	0.721	0.829	0.756	0.716	0.728	0.697	0.127	0.601	0.219	0.582	0.819	0.210	0.074
内蒙古	0.655	0.512	0.635	0.682	0.645	0.256	0.259	0.197	0.352	0.340	0.133	0.515	0.192	0.549	0.197
辽宁	0.668	0.607	0.709	0.769	0.724	0.240	0.550	0.405	0.198	0.315	0.405	0.610	0.153	0.390	0.153
吉林	0.720	0.471	0.760	0.708	0.700	0.193	0.326	0.413	0.198	0.235	0.324	0.565	0.309	0.529	0.053
黑龙江	0.756	0.547	0.685	0.671	0.687	0.294	0.483	0.281	0.159	0.343	0.327	0.467	0.354	0.311	0.070
上海	0.890	0.861	0.790	0.835	0.808	0.698	0.748	0.603	0.426	0.659	0.786	0.581	0.767	0.187	0.442
江苏	0.851	0.839	0.865	0.879	0.835	0.811	0.840	0.751	0.426	0.738	0.772	0.770	0.868	0.482	0.526
浙江	0.791	0.787	0.746	0.841	0.860	0.689	0.702	0.531	0.680	0.658	0.518	0.471	0.761	0.524	0.572
安徽	0.882	0.758	0.849	0.818	0.794	0.770	0.824	0.728	0.327	0.681	0.737	0.731	0.820	0.396	0.575
福建	0.862	0.810	0.865	0.863	0.846	0.782	0.823	0.711	0.642	0.691	0.602	0.697	0.735	0.657	0.361
江西	0.840	0.875	0.826	0.838	0.836	0.793	0.816	0.609	0.278	0.692	0.697	0.633	0.814	0.553	0.246
山东	0.833	0.806	0.800	0.863	0.836	0.791	0.815	0.696	0.634	0.659	0.777	0.711	0.820	0.339	0.430
河南	0.850	0.805	0.873	0.831	0.826	0.757	0.843	0.728	0.701	0.574	0.674	0.628	0.848	0.315	0.407
湖北	0.835	0.803	0.856	0.829	0.833	0.768	0.788	0.744	0.576	0.689	0.701	0.753	0.845	0.537	0.478
湖南	0.885	0.815	0.879	0.872	0.829	0.799	0.848	0.748	0.517	0.751	0.752	0.684	0.815	0.504	0.596
广东	0.877	0.822	0.884	0.864	0.786	0.787	0.858	0.749	0.487	0.708	0.790	0.718	0.711	0.682	0.272
广西	0.901	0.850	0.891	0.836	0.812	0.718	0.754	0.572	0.639	0.710	0.510	0.489	0.703	0.384	0.121
海南	0.756	0.821	0.679	0.808	0.910	0.154	0.731	0.385	0.385	0.692	0.385	0.295	0.564	0.282	0.000
重庆	0.829	0.740	0.690	0.757	0.553	0.592	0.638	0.519	0.291	0.521	0.504	0.503	0.346	0.527	0.032
四川	0.849	0.767	0.740	0.805	0.638	0.624	0.659	0.563	0.399	0.544	0.520	0.490	0.346	0.305	0.104
贵州	0.847	0.737	0.721	0.744	0.726	0.521	0.659	0.561	0.388	0.545	0.513	0.444	0.251	0.218	0.068
云南	0.858	0.774	0.698	0.747	0.667	0.559	0.642	0.548	0.383	0.487	0.444	0.484	0.422	0.102	0.053
西藏	0.833	0.592	0.758	0.783	0.575	0.433	0.742	0.492	0.313	0.242	0.233	0.608	0.242	0.125	0.000
陕西	0.879	0.743	0.878	0.866	0.757	0.762	0.807	0.659	0.515	0.740	0.444	0.290	0.633	0.152	0.317
甘肃	0.673	0.699	0.674	0.824	0.443	0.434	0.376	0.443	0.598	0.389	0.342	0.278	0.206	0.344	0.167
青海	0.834	0.779	0.860	0.618	0.768	0.749	0.800	0.606	0.550	0.565	0.256	0.143	0.722	0.096	0.266
宁夏	0.726	0.593	0.720	0.691	0.545	0.399	0.305	0.270	0.567	0.240	0.162	0.442	0.405	0.567	0.043
新疆	0.860	0.654	0.851	0.817	0.692	0.794	0.748	0.373	0.739	0.473	0.526	0.378	0.662	0.514	0.376

表 4 不同地区各 VNTR 组合位点及其成簇率与 HGI

地区	12-VNTR 组合	HGI	成簇率(%)	10-VNTR 组合	HGI	成簇率(%)	8-VNTR 组合	HGI	成簇率(%)	5-VNTR 组合	HGI	成簇率(%)
北京	1,2,3,4,5,6,7,9,10,12,13,15	0.999 3	1.82	12-VNTR-9,15	0.999 3	1.82	10-VNTR-10,12	0.999 3	1.82	8-VNTR-6,7,13	0.998 7	3.64
天津	1,2,3,4,5,6,7,8,10,12,13,15	1.000 0	0.00	12-VNTR-10,15	1.000 0	0.00	10-VNTR-12,13	1.000 0	0.00	8-VNTR-2,5,7	0.998 3	6.67
河北	4,5,6,7,8,9,10,11,12,13,14,15	0.999 8	1.78	12-VNTR-14,15	0.999 6	1.78	10-VNTR-10,11	0.999 3	3.55	8-VNTR-7,8,12	0.998 4	11.24
山西	1,2,4,5,6,7,8,9,10,11,12,14	0.999 4	4.72	12-VNTR-9,14	0.999 4	5.66	10-VNTR-11,12	0.998 7	7.55	8-VNTR-2,8,10	0.977 1	8.49
内蒙古	1,2,3,4,5,6,7,8,9,10,12,14	1.000 0	0.00	12-VNTR-6,8	0.997 5	3.45	10-VNTR-7,10	0.992 6	10.34	8-VNTR-2,9,12	0.987 7	13.79
辽宁	1,2,3,4,5,6,7,8,10,11,12,14	0.998 5	2.70	12-VNTR-6,10	0.997 0	5.41	10-VNTR-11,14	0.997 0	5.41	8-VNTR-2,7,8	0.988 0	18.92
吉林	1,2,3,4,5,6,7,8,10,11,12,13,14	0.994 3	5.26	12-VNTR-10,13	0.994 3	5.26	10-VNTR-7,11	0.994 3	5.26	8-VNTR-2,8,14	0.988 6	18.42
黑龙江	1,2,3,4,5,6,7,10,11,12,13,14	0.997 4	7.14	12-VNTR-6,11,14	0.995 4	10.71	10-VNTR-10,11	0.993 8	11.90	8-VNTR-7,12,13	0.987 7	22.62
上海	1,2,4,5,6,7,8,10,11,12,13,15	1.000 0	0.00	12-VNTR-12,15	1.000 0	0.00	10-VNTR-6,8,10	1.000 0	0.00	8-VNTR-7,13	0.998 4	5.63
江苏	1,2,4,5,6,7,8,10,11,12,13,15	1.000 0	0.00	12-VNTR-10,15	0.999 9	1.94	10-VNTR-8,11	0.999 9	2.33	8-VNTR-5,12	0.999 3	5.81
浙江	1,2,4,5,6,7,8,9,10,11,14,15	1.000 0	0.00	12-VNTR-11,14	0.999 9	0.83	10-VNTR-8,15	0.999 9	0.83	8-VNTR-6,9,10	0.998 6	10.00
安徽	1,2,4,5,6,7,8,10,11,12,13,15	1.000 0	0.00	12-VNTR-13,15	0.999 9	1.84	10-VNTR-8,10	0.999 9	1.84	8-VNTR-2,11,12	0.998 7	8.76
福建	1,2,4,5,6,7,8,9,10,12,13,14	0.999 9	0.77	12-VNTR-9,14	0.999 9	0.77	10-VNTR-10,12	0.999 8	2.31	8-VNTR-6,8,13	0.998 9	6.15
江西	2,4,5,6,7,8,9,10,11,12,13,14	0.999 9	1.03	12-VNTR-8,14	0.999 9	1.03	10-VNTR-10,12	0.999 7	2.58	8-VNTR-6,7,11	0.997 1	13.40
山东	1,2,4,5,6,7,8,9,10,11,12,13	1.000 0	0.99	12-VNTR-9,10	0.999 9	0.85	10-VNTR-8,12	0.999 9	1.27	8-VNTR-2,6,11	0.999 2	10.17
河南	1,2,4,5,6,7,8,9,10,11,12,13	1.000 0	0.00	12-VNTR-10,12	1.000 0	0.00	10-VNTR-9,11	0.999 9	0.61	8-VNTR-2,6,8	0.998 7	9.15
湖北	1,2,4,5,6,7,8,9,10,11,12,13	0.999 9	0.38	12-VNTR-9,10	0.999 9	0.38	10-VNTR-8,11	0.999 9	1.15	8-VNTR-6,7,12	0.999 3	7.69
湖南	1,2,4,5,6,7,8,9,10,11,12,15	0.999 9	0.30	12-VNTR-9,15	0.999 9	0.30	10-VNTR-8,12	0.999 9	0.90	8-VNTR-6,10,11	0.999 6	6.31
广东	2,4,5,6,7,8,9,10,11,12,13,14	0.999 9	0.46	12-VNTR-9,14	0.999 9	0.46	10-VNTR-10,13	0.999 9	1.39	8-VNTR-5,8,12	0.999 3	6.48
广西	2,4,5,6,7,8,9,10,11,12,13,14	1.000 0	0.00	12-VNTR-12,14	1.000 0	0.00	10-VNTR-7,11	0.999 8	1.03	8-VNTR-6,10,13	0.998 2	6.32
海南	1,2,3,4,5,7,8,9,10,11,12,13	1.000 0	0.00	12-VNTR-9,12	1.000 0	0.00	10-VNTR-8,11	1.000 0	0.00	8-VNTR-3,10,13	1.000 0	0.00
重庆	1,2,3,4,5,6,7,8,9,10,11,12,14	0.998 5	4.76	12-VNTR-11,12	0.998 0	9.52	10-VNTR-8,10	0.997 2	11.90	8-VNTR-5,6,14	0.986 2	26.19
四川	1,2,3,4,5,6,7,8,9,10,11,12	0.999 6	4.33	12-VNTR-9,12	0.999 3	4.69	10-VNTR-10,11	0.999 1	3.61	8-VNTR-5,6,8	0.996 3	24.91
贵州	1,2,3,4,5,6,7,8,9,10,11,12	0.999 6	2.07	12-VNTR-9,12	0.999 2	3.45	10-VNTR-6,9	0.998 5	6.21	8-VNTR-7,8,10	0.995 7	14.48
云南	1,2,3,4,5,6,7,8,10,11,12,13	1.000 0	0.00	12-VNTR-11,13	0.998 6	2.63	10-VNTR-10,12	0.998 6	2.63	8-VNTR-6,7,8	0.998 6	2.63
西藏	1,2,3,4,5,6,7,8,9,10,12,13	1.000 0	0.00	12-VNTR-10,13	1.000 0	0.00	10-VNTR-6,9	1.000 0	0.00	8-VNTR-5,8	1.000 0	0.00
陕西	1,2,3,4,5,6,7,8,9,10,11,13	1.000 0	0.00	12-VNTR-9,11	1.000 0	0.00	10-VNTR-8,13	1.000 0	0.00	8-VNTR-5,6,10	0.998 7	2.04
甘肃	1,2,3,4,5,6,7,8,9,10,11,14	0.997 1	6.52	12-VNTR-11,14	0.997 1	6.52	10-VNTR-7,10	0.996 1	8.70	8-VNTR-5,6,8	0.996 1	8.70
青海	1,2,3,4,5,6,7,8,9,10,11,15	1.000 0	0.00	12-VNTR-11,15	1.000 0	0.00	10-VNTR-9,10	0.997 6	4.88	8-VNTR-4,6,8	0.993 9	12.20
宁夏	1,2,3,4,5,6,7,8,9,12,13,14	0.991 3	15.22	12-VNTR-7,8	0.991 3	15.22	10-VNTR-6,13	0.988 4	19.57	8-VNTR-5,12,14	0.963 3	41.30
新疆	1,2,3,4,5,6,7,9,10,11,13,14	0.999 1	4.12	12-VNTR-10,14	0.999 1	4.12	10-VNTR-2,11	0.999 1	4.12	8-VNTR-5,9,13	0.998 5	6.19

注:1~15 分别代表不同的 VNTR 位点: QUB11b、Mtb21、QUB26、MIRU26、Mtb04、MIRU10、MIRU31、MIRU40、QUB4156c、ETR-A、Mtb30、Mtb39、MIRU16、MIRU04、ETR-C; HGI 代表各 VNTR 汉高指数

查,以发现基因型相匹配的患者(成簇)之间的流行病学联系。所以结核病近期传播必须要有两个先决条件:①菌株基因形成簇;②患者有流行病学联系。因此本研究构建涵盖31个省3 996株MTB的15-VNTR数据库,有助于MTB传染源全国范围的追踪和簇调查。

目前我国虽已有多个不同VNTR组合评估的研究,但均仅针对某一谱系或某一地区,尚无多个谱系MTB流行的国家层面综合评估。本研究分析15个VNTR位点的分辨力发现,QUB11b、Mtub21、QUB26、MIRU26、Mtub04、MIRU10和MIRU31位点分辨率较高,与Comas等^[13]和Chen等^[16]研究结果一致。同时发现QUB26、MIRU16、Mtub21、QUB11b位点在部分省份遗传稳定性差。我国是多个MTB谱系流行且分布不均衡的结核病高负担国家。为提高结核病预防控制的效率,有必要向各省推荐一套优化的VNTR位点组合。优先推荐QUB11b、Mtub21、QUB26、MIRU26、Mtub04、MIRU10、MIRU31这7个较高分辨率位点,并以此为基础的12-VNTR、10-VNTR、8-VNTR 3套VNTR组合。

目前全基因组测序技术因居高的价格尚未应用于结核病的预防控制,纵然其价格下降,其以系统进化和群体遗传学的分析也是深奥的。鉴于VNTR基因分型同源异质性指数(homoplasy index, HI)较Spoligotyping更接近SNP^[13],在谱系定义(LSP和SNP)基础上结合VNTR分型便是这一时期最高效的基因分型工具。因此对一个地区MTB进行全方位的VNTR基因分型并建立数据库,可利用成簇率和簇即代表近期的传播的概念,从时间和空间监测结核病疫情的变化。一个地区成簇率高,则疫情发展需密切关注,如成簇率低,则疫情平稳。根据Wirth等^[17]以及之前研究结果提示^[18-19],利用LSP、SNP以及VNTR数据研究MTB的群体遗传结构特征及系统进化,例如研究MTB的最近共祖年代、系统进化等,从时间和空间探讨该地区MTB的流行特征。总之,各省的结核病工作者可利用本研究推荐的VNTR组合建立当地结核菌基因分型数据库(含LSP或SNP鉴定和VNTR分型)进行疫情监测、初步系统进化和群体遗传学的分析,为当地结核病疫情控制策略的制定提供科学依据。

志谢 感谢国家结核病参比实验室赵雁林教授等开展的2007—2008年度全国结核病耐药性基线调查;感谢国家结核病临床实验室霍凤敏实习研究员、姜广路副主任技师等为该工作提供的菌株资源和数据资源。

利益冲突 无

参考文献

- [1] WHO. Global tuberculosis report 2016[R/OL]. Geneva: WHO, 2016. http://www.who.int/tb/publications/global_report/en/.
- [2] Luo T, Comas I, Luo D, et al. Southern East Asian origin and coexpansion of *Mycobacterium tuberculosis* Beijing family with Han Chinese[J]. Proc Natl Acad Sci USA, 2015, 112(26): 8136–8141. DOI: 10.1073/pnas.1424063112.
- [3] Zhang HT, Li DF, Zhao LL, et al. Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance[J]. Nat Genet, 2013, 45(10): 1255–1260. DOI: 10.1038/ng.2735.
- [4] Gardy JL, Johnston JC, Sui SJH, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak[J]. N Engl J Med, 2011, 364(8): 730–739. DOI: 10.1056/NEJMoa1003176.
- [5] Supply P, Lesjean S, Savine E, et al. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units[J]. J Clin Microbiol, 2001, 39(10): 3563–3571. DOI: 10.1128/JCM.39.10.3563–3571.2001.
- [6] Supply P, Allix C, Lesjean S, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*[J]. J Clin Microbiol, 2006, 44(12): 4498–4510. DOI: 10.1128/JCM.01392–06.
- [7] Luo T, Yang CG, 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]. J Clin Microbiol, 2012, 50(3): 633–639. DOI: 10.1128/JCM.05539–11.
- [8] Luo T, Yang CG, Pang Y, et al. Development of a hierarchical variable-number tandem repeat typing scheme for *Mycobacterium tuberculosis* in China[J]. PLoS One, 2014, 9(2): e89726. DOI: 10.1371/journal.pone.0089726.
- [9] Luo T, Yang CG, Gao Q. Mycobacterial interspersed repetitive-unit locus PCR amplification and Beijing strains of *Mycobacterium tuberculosis*[J]. J Clin Microbiol, 2011, 49(11): 4026–4027. DOI: 10.1128/JCM.05389–11.
- [10] Zhao YL, Xu SF, Wang LX, et al. National survey of drug-resistant tuberculosis in China[J]. N Engl J Med, 2012, 366(23): 2161–2170. DOI: 10.1056/NEJMoa1108789.
- [11] 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. DOI: 10.1371/journal.pone.0032976.
- [12] 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.
- [13] Comas I, Homolka S, Niemann S, et al. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies[J]. PLoS One, 2009, 4(11): e7815. DOI: 10.1371/journal.pone.0007815.
- [14] Kremer K, Arnold C, Cataldi A, et al. Discriminatory power and reproducibility of novel DNA typing methods for *Mycobacterium tuberculosis* complex strains[J]. J Clin Microbiol, 2005, 43(11): 5628–5638. DOI: 10.1128/JCM.43.11.5628–5638.2005.
- [15] Sola C, Filliol I, Legrand E, et al. Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics[J]. Infect Genet Evol, 2003, 3(2): 125–133. DOI: 10.1016/S1567-1348(03)00011-X.
- [16] Chen YY, Chang JR, Huang WF, et al. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on SNP and VNTR typing profiles in Asian countries[J]. PLoS One, 2012, 7(7): e39792. DOI: 10.1371/journal.pone.0039792.
- [17] Wirth T, Hildebrand F, Allix-Béguec C, et al. Origin, spread and demography of the *Mycobacterium tuberculosis* complex[J]. PLoS Pathog, 2008, 4(9): e1000160. DOI: 10.1371/journal.ppat.1000160.
- [18] 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. DOI: 10.3760/cma.j.issn.0254-6450.2013.04.015.
- [19] 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. DOI: 10.3760/cma.j.issn.0254-6450.2013.04.015.
- [19] 李颖, 付育红, 原梅, 等. 四川盆地结核分枝杆菌群体遗传学特征研究[J]. 中华流行病学杂志, 2015, 36(4): 374–378. DOI: 10.3760/cma.j.issn.0254-6450.2015.04.017.
- [19] Li Y, Fu YH, Yuan M, et al. Study on the population-genetics of *Mycobacterium tuberculosis* from Sichuan Basin in China[J]. Chin J Epidemiol, 2015, 36(4): 374–378. DOI: 10.3760/cma.j.issn.0254-6450.2015.04.017.

(收稿日期: 2016-10-27)
(本文编辑: 斗智)