

中国土拉弗朗西斯菌 *holarctica* 亚种的遗传多样性

王艳华 彭遥 夏连续

【摘要】 目的 研究国内外土拉弗朗西斯菌的遗传进化关系。方法 选择17个单核苷酸多态性、4个插入/缺失和12个可变数目串联重复,采用单核苷酸多态性和插入/缺失、多位点可变数目串联重复分析方法单独和组合起来对39株土拉菌(10株中国土拉菌和29株已公布测序的土拉菌)进行系统进化分析。结果 组合分析显示,3株中国土拉菌和日本的FSC022被分配到B5;剩余3株中国土拉菌和瑞典的FSC200被分配到B1;3株和美国的OSU18被分配到B2;1株和法国的FTNF002-00、德国的F92与美国的OR96246一起被分配到B4。10株中国土拉菌分为4种亚型,研究表明中国土拉菌具有广泛的遗传多样性。结论 本研究针对土拉菌B型建立了一套简易高效的分型方法,并以此为基础得出土拉菌B型的起源可能是亚洲地区。

【关键词】 土拉弗朗西斯菌;土拉热;系统进化;单核苷酸多态性;插入/缺失;多位点可变串联重复分析

Genetic diversity of *Francisella tularensis* subsp. *holarctica* in China Wang Yanhua, Peng Yao, Xia Lianxu. National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

Corresponding author: Wang Yanhua, Email: wangyanhua@icdc.cn

This work was supported by grants from the National Natural Science Foundation of China (No. 81473032) and National Science and Technology Major Project of China (No. 2012ZX10004219-007).

【Abstract】 Objective To explore the genetic relationship between the Chinese and the foreign species of *Francisella tularensis*. **Methods** Based on our own findings and from the literature, 17 SNP, 4 INDEL, and 12 VNTR were selected for phylogenetic analysis on 39 strains of *F. tularensis*, including 10 strains of Chinese *F. tularensis* and 29 strains of foreign *F. tularensis* that had been sequenced and published. SNP-INDEL and MLVA were used for the separation and combination. **Results** Data from the combined analysis indicated that 3 strains of Chinese *F. tularensis* with Japanese FSC022 were assigned to B5; 3 strains, with Swedish FSC200 to B1; 3 strains with American OSU18 to B2 and 1 strain with French FTNF002-00, German F92, and American OR96246 to B4, respectively. 10 strains of Chinese *F. tularensis* were assigned to 4 clades and the result demonstrated a wide diversity of *F. tularensis* subsp. *holarctica* in China. **Conclusion** A set of simple and robust typing tools for *F. tularensis* subsp. *holarctica* were established in this study. Based on the results, *F. tularensis* subsp. *holarctica* might have had its origins in Asia.

【Key words】 *Francisella tularensis*; Tularemia; Phylogeography; Single-nucleotide polymorphism; Insertion-deletion; Multiple-locus variable number tandem repeat analysis

土拉弗朗西斯菌(土拉菌)是一种具有较强传染性的、能够引起土拉热的胞内寄生菌。土拉热主要由2个亚种引起: subsp. *tularensis* (A型)和 subsp. *holarctica* (B型)。A型分布仅限于北美地区。B型亚种的研究主要集中在欧洲地区,根据地理分布特征,欧洲地区的土拉菌分为两个亚群^[1]。在西班牙、

法国和瑞士等国几乎分离的土拉菌都属于B.Br. FTNF002-00亚群^[2-5],而从捷克、俄罗斯等国分离的大多数土拉菌属于B.Br.013亚群内的多个谱系^[2,4,6]。目前,B型各亚群在亚洲地区的分布不是很清楚,关于中国土拉菌B型亚种和欧美地区B型亚种之间关系的研究更是少见。为了解中国土拉菌和国外土拉菌在亚型上的异同点,从而能够在土拉热暴发和流行时进行菌株溯源。前期以单核苷酸多态性(SNP)为基础,初步进行了中国土拉菌的遗传进化关系研究^[7]。本研究拟通过针对土拉菌B型亚种建立一套简易高效的分型方法,进行更加深入地探讨。

DOI: 10.3760/cma.j.issn.0254-6450.2015.12.021

基金项目: 国家自然科学基金(81473032); 国家科技重大专项(2012ZX10004219-007)

作者单位: 102206 北京, 中国疾病预防控制中心传染病预防控制所

通信作者: 王艳华, Email: wangyanhua@icdc.cn

材料与方法

1. 土拉菌来源:本研究的10株中国土拉菌中,410105、410108、410109、410111、410112、410113和920607来源于西藏,410116和410117来源于新疆,410107来源于黑龙江。另外,还有29个参考菌株,包括SCHU S4、FSC147、FTNF002-00、U112、WY96-3418、OSU18、LVS、FSC198、FSC033、MA00-2987、NE061598、257、FSC022、FSC200、GA99-3548、GA99-3549、RC503、MI001730、OR96246、F92、URFT1、ATCC6223、80700103、AS_713、831、70102010、TIGB03、TI0902和80700075,

基因组信息获取网址:<http://www.ncbi.nlm.nih.gov/genome/genomes/511> 或 http://www.broadinstitute.org/annotation/genome/francisella_tularensis_group/Regions.html。

2. 选择适合的SNP和插入/缺失(INDEL):根据前期研究^[7],从*fopA*和*tul4*基因中选出6个SNP位点(F1、F2、F3、F4、F5和F6)。从以往用于B型亚种分型的遗传标记中^[2,4],选择更有效力的11个SNP位点(B2、B3、B4、B5、B6、B12、B16、B17、B19、B20和B23)和4个INDEL位点(Ftind33、Ftind38、Ftind48和Ftind49),共21个遗传标记作为本研究的靶位点(表1)。为了获得准确的SNP信息,对于中国土拉

表1 21个SNP和INDEL标记和引物

SNP/INDEL	SNP在SCHU S4中的位置	SNP状况	引物	序列	退火温度(°C)
F1	909934	T	F	GCTGTATCATCATTTAATAAACTGCTG	55
		G	R	TTGGGAAGCTTGTATCATGGCACT	
F2	909969	G	F	GCTGTATCATCATTTAATAAACTGCTG	55
		A	R	TTGGGAAGCTTGTATCATGGCACT	
F3	910179	A	F	GCTGTATCATCATTTAATAAACTGCTG	55
		G	R	TTGGGAAGCTTGTATCATGGCACT	
F4	910232	A	F	GCTGTATCATCATTTAATAAACTGCTG	55
		G	R	TTGGGAAGCTTGTATCATGGCACT	
F5	599580	C	F	CTTGAGTCTTATGTTTCGGCATGTGAATAG	50
		T	R	CCAACTAATTGGTTGTACTGTACAGCGAAG	
F6	599411	T	F	CTTGAGTCTTATGTTTCGGCATGTGAATAG	50
		C	R	CCAACTAATTGGTTGTACTGTACAGCGAAG	
B2	5162	A	F	TTAGTCTATGAGCAGCCAG	50
		C	R	TAATATACCAAGGTAGCC	
B3	470842	A	F	ACGCTAGGTGTCTTGGT	50
		G	R	CTATATCCGCTCAACAT	
B4	823673	T	F	TAGACGCACTGGATTTAGGT	53.5
		A	R	AACCATCAGCCACCATAAG	
B5	1853699	T	F	TGGATCAAACAACCGT	50
		C	R	TCTCAAGAGCTGGTGC	
B6	713648	A	F	AGTAGTGGTAGCGAGGC	53.5
		G	R	TACCGTTAGCCCAACAG	
B12	109781	T	F	TACTGCCCAACATAGAG	55
		A	R	ATCGTGATAAGGCTGGA	
B16	608246	T	F	ATGCTAGCAAATTACCATCAAAAAG	57
		G	R	AACTCTTCTCGCCATCAACTTCTAT	
B17	1743251	T	F	CCAAGAGCTAAATTAGCTTCAA	53.5
		G	R	TGACCAAGAAGGTAGAGGTATTGGTT	
B19	1374034	A	F	TTGCTACTGATGGTTAACT	57
		C	R	CAATACGTCACCTTATGCAGTGAT	
B20	1396117,789461	G	F	ATGGGTCGGACTATCACATC	56
		A	R	ATTATTGTAAAACGGCATCG	
B23	253121	T	F	GGCAACAGCAGATTCGTGAG	56
		G	R	TGAAAGCAGGTTTAGAAGGACAG	
Ftind33	512045,512063	IN		TCTAAATTAAGCAATGTTTCTAACT	53
		OUT		ATCATCGTATAAGAAATCAACTT	
		CP		TCAACCTTACAGAATAAGAATGT	
Ftind38	95661,95674	IN		ACCCAATAAGCTCACCATCA	50
		OUT		ATCTTTCTCAGGTACAGACTTTA	
		CP		AGTACTATTTGCTTATCCAAGTGAA	
Ftind48	960987,961011	IN		ATCCTACTAATATCAATTCCAGT	53
		OUT		CCTTCAGCTTGAGTATTTTGACGT	
		CP		ACTGTTATATTCAGTTAATTTGCT	
Ftind49	834341,834349	IN		AAGATTAAGTGGCAATTTAC	52
		OUT		TTCAACCTGGACAACCACTA	
		CP		AGGATCCCAGTTAGGTTTAGTA	

菌,所有的 SNP 位点都通过测序获得, F1 ~ F6 位点 SNP 状况从以前的测序结果中得到^[7]。每个 INDEL 位点均通过 2 对引物(CP-IN 和 CP-OUT)扩增以确定是否有插入片段的存在。CP-OUT 引物作为阳性对照。对于 29 个测序菌株,在基因组序列中通过比对上下游引物,可以找到相应位置的 SNP 位点以及确定检测片段的插入和缺失。

3. 选择适合的 VNTR:不考虑 A 型亚种,仅考虑对于 B 型亚种分型能力较强,同时考虑时间和费用的最佳组合。从 25 个 VNTR 位点中,选出对于 B 型亚种分型能力较强的 12 个位点(Ft-M3、Ft-M4、Ft-M6、Ft-M8、Ft-M10、Ft-M13、Ft-M16、Ft-M18、Ft-M20、Ft-M21、Ft-M22 和 Ft-M24),引物序列和退火温度参照文献^[8]。通过 PCR 方法在 10 株中国

能够很好地将 4 个亚种区分出来,而且,不仅能够将 A 型的两个主要分支 A1 和 A2 区分开,对于 B 型也能够将日本土拉菌和北半球其他国家的土拉菌区分开(图 2A)。从主干到分支的进化分析为:F1 可以将 subsp. *novicida* 区分开; F2 和 F6 可以将 subsp. *tularensis* 区分开; F3 可以将 subsp. *holarctica* 区分开。在 subsp. *tularensis* 内, F4 可以将 A2 区分出来, F5 可以进一步将毒力弱的 ATCC6223 和其他 A2 菌株区分出来。引导分析的节点值大于 50 的,在节点左上方标出(图 2),在节点的左下方标出了分型标记(图 2A)。

2. MLVA 分析:MLVA 不能将 subsp. *novicida* 和 subsp. *mediasiatica* 区分开,能够将 A 型中的 A1 和 A2 以及 B 型中的日本来源菌株和其他来源菌株区

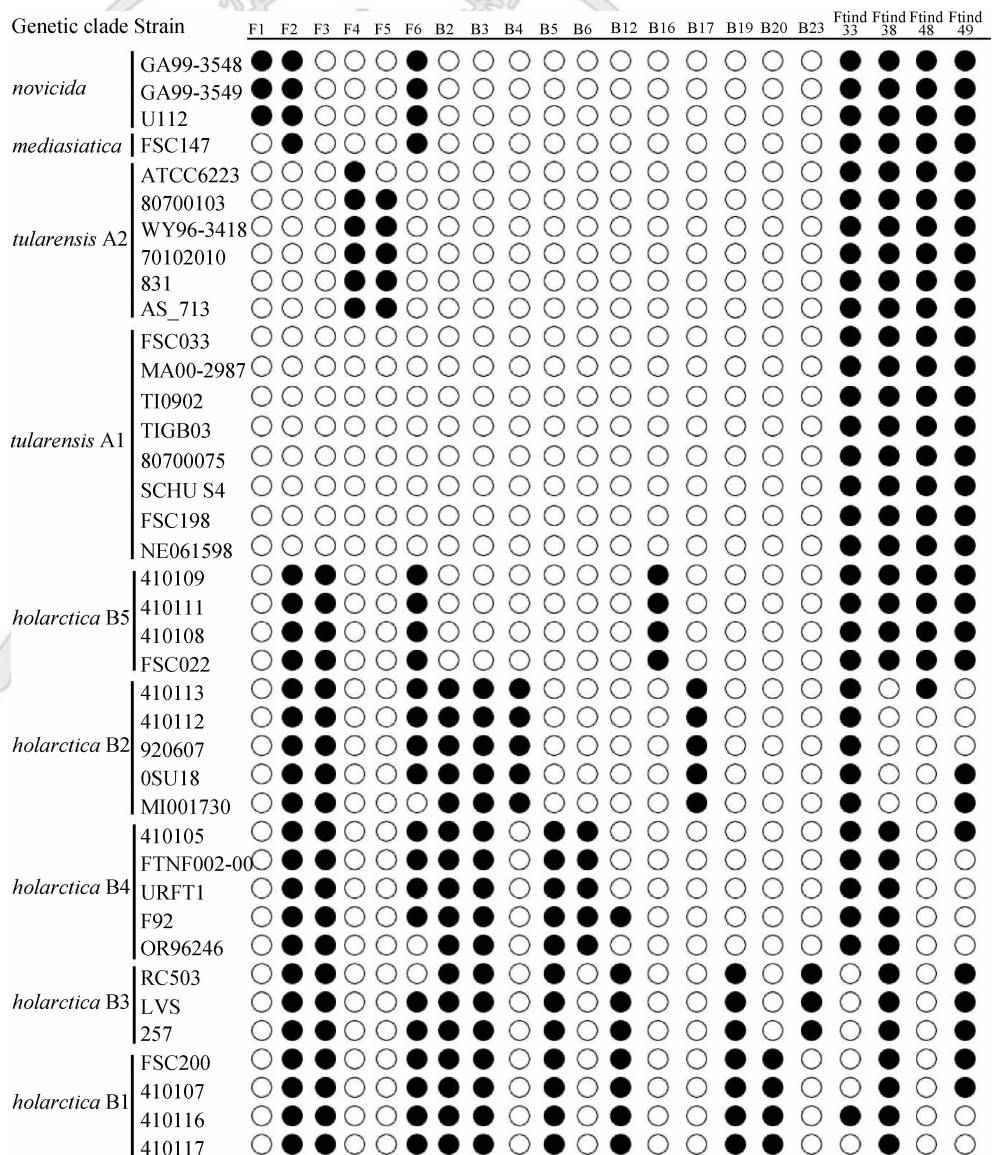
土拉菌中 12 个 VNTR 位点重复扩增 3 次。在每个 VNTR 位点,保证每种片段长度的 PCR 产物都要通过测序来获得碱基序列,再查找重复基元的拷贝数。对于 29 个参考菌株,在基因组序列中通过比对上下游引物找到 12 个 VNTR 区域,再查找重复基元的拷贝数。

4. 系统进化分析:采用 PAUP 4.0b10 软件建立系统进化树,采用最大似然法(MP)构建,MP 分析采用 50 次替换。引导分析用 1 000 次替换的启发式研究,节点值小于 50 被剔除。

结 果

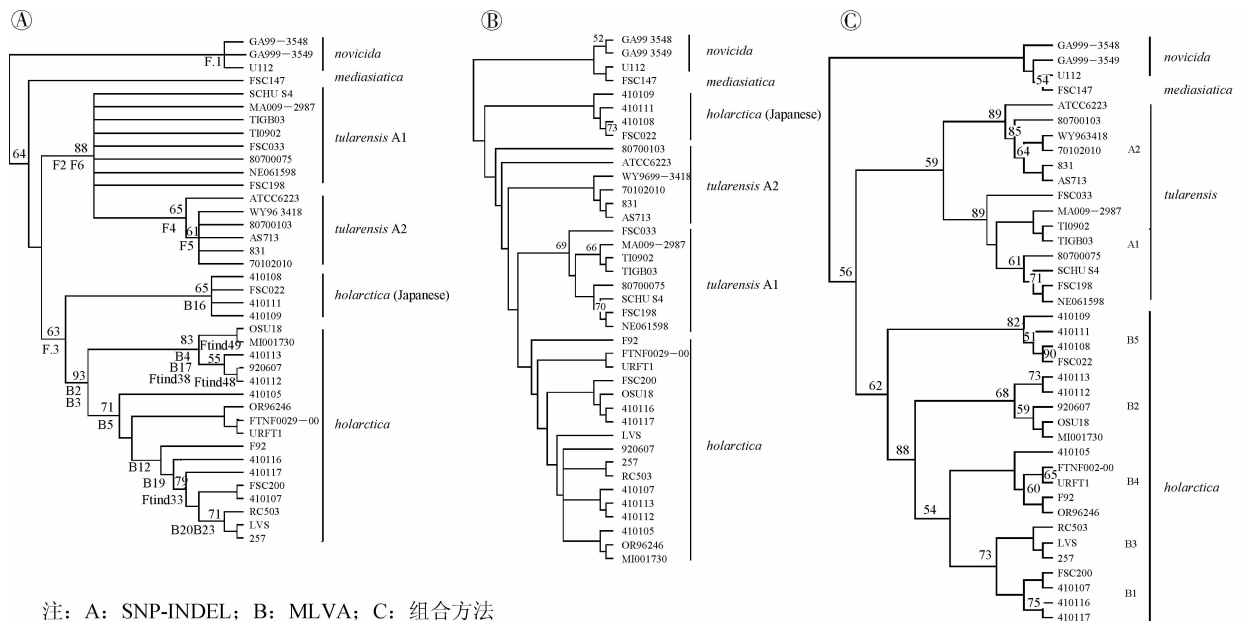
1. SNP-INDEL 分析:21 个 SNP 和 INDEL 在 39 株土拉菌中的分布情况见图 1。

SNP-INDEL 分析



注:对于 SNP,菌株状况和 SCHU S4 一致的,用 ○ 表示,反之则用 ● 表示;对于 INDEL,缺失用 ○ 表示,插入则用 ● 表示

图 1 21 个 SNP 和 INDEL 在 39 株土拉菌中的分布



注: A: SNP-INDEL; B: MLVA; C: 组合方法

图2 采用SNP-INDEL、MLVA单独和组合方法分析土拉菌之间的遗传进化关系

分开(图2B)。在树状图主干进化支上,SNP-INDEL的节点值较高(图2A),而MLVA的节点值均低于50(图2B)。

3. 组合分析:SNP-INDEL和MLVA组合分析显示,与SNP-INDEL、MLVA单独分析相比,在远端分支上的节点值增加(图2C),这说明对于菌株间的区分能力增强。组合分析方法将subsp. *holarctica*分为5个亚型,分别命名为B1~B5。值得注意的是,中国的10株土拉菌分到了4个亚型中,3株和瑞典的FSC200分到B1;3株和美国的OSU18分到B2;1株和法国的FTNF002-00、德国的F92、美国的OR96246分到B4;最后3株和日本的FSC022分到B5。

讨论

细菌的分子分型和系统进化分析的方法很多,目前一致认为,SNP是缓慢突变的标记,在分支的主干上能够保证系统进化分类的正确性;而MLVA是非常快速突变的标记,在远端的分支上对各亚种内的菌株个体能够进行细致的区分。如果将SNP和MLVA组合,能够互补不足、发挥各自的优势,达到更加理想的分析效果。另外,INDEL能够增加SNP在主要进化节点上的支持度和结果的可靠性,并增强节点上的分辨力^[4,9]。因此,将SNP-INDEL和MLVA组合起来,在主干进化支上SNP-INDEL保证分类的正确性和可靠性,在远端进化支上MLVA对菌株个体进行细致的区分,二者组合就会对土拉菌的进化分析提供更高的精确性和分辨力^[10]。为了最小化遗传标记的数量,根据前期的工作^[7]和以往的

研究结果^[2,4],选出17个SNP和4个INDEL建立了一个分型系统。SNP-INDEL和MLVA单独分析构建的树状图,虽然在分类上略有不同,但是对于所有B型菌株均分为日本土拉菌和非日本土拉菌,在这两个亚组的主要分界点上是一致的。本研究显示:SNP-INDEL和MLVA组合分析,既在主干系统进化关系上获得了更高的精确性,同时对于亚种内的菌株个体之间又表现出更强的分辨力。因此,本研究针对土拉菌B型亚种建立了一套简易高效的分型方法。

在本研究的SNP-INDEL分析中,测序菌株SCHU S4、FSC033和WY96-3418、ATCC6223分别被分到A1和A2亚群中,分类结果与Svensson等^[4]一致。与SNP-INDEL相比,MLVA明显地表现出较高的多态性,对描述菌株个体的特征很有帮助。组合分析方法将subsp. *holarctica*分为5个亚型(B1~B5),和Svensson等^[4]的分析结果一致。中国10株土拉菌被分配到4个亚型中,表现出多样性的遗传特征。虽然设计时选择的遗传标记未考虑subsp. *tularensis*,但SNP-INDEL和MLVA组合分析同样也能将subsp. *tularensis*区分为两个主要亚群,只是未能将subsp. *novicida*和subsp. *mediasiatica*区分开。

一项日本研究显示:日本33株菌中有31株被分配到在Cluster I,仅有2株和其他国家来源的菌株被分配到Cluster II,表明日本菌株具有显著的本国特有的遗传特征^[11]。值得注意的是,在SNP-INDEL和MLVA的单独、以及二者的组合分析中,中国土拉菌410108、410109、410111和日本的FSC022均被分配到一个分支,这说明中国土拉菌和日本土拉菌具

有某些相同的遗传特征。由图1可看到,3株中国土拉菌在B2、B3的SNP状况与FSC022完全相同,而和其他的subsp. *holarctica*的SNP状况都不相同。另外,在B16中国这3株土拉菌的SNP状况仍然与FSC022完全相同,而其他包括4个亚种的35株土拉菌的SNP状况都和这4株菌的不同。按照Svensson等^[4]的推论,中国这3株土拉菌可以和日本土拉菌一同归入B5亚型。除本研究发现的中国菌株,到目前为止,其他采用SNP/MLVA分析方法,日本来源的菌株都是单独被分到一个亚组^[2,8-9,12],没有任何其他国家的菌株能和日本来源的菌株分到一起的。在SNP-INDEL和MLVA的组合分析中,除了B3是俄罗斯菌株独有的亚型外,其余4个亚型(B1、B2、B4和B5)都包含有中国土拉菌,再次证实中国土拉菌的遗传多样性^[13]。

目前研究认为,土拉菌subsp. *holarctica*之间的遗传差异性很小,说明这个亚种仅仅是最近克服了一个遗传学障碍才得以出现,然后迅速地播散到整个北半球^[8-9,14-17]。然而,对于subsp. *holarctica*的起源问题一直有争议。Vogler等^[2]持两种观点,一种是根据FSC022在进化支中的基础位置,认为这个亚种可能起源于亚洲地区;另外,根据MIP SNP分析中加利福尼亚谱系也在较基础的位置,认为这个亚种也可能起源于北美,然后通过B辐射播散到现在分布的地区。遗憾的是,系统进化分析中的土拉菌数据仅包括了日本菌株,而没有亚洲地区其他来源的菌株。这些假设是建立在一个中心多样性的理论基础之上的,它的含义是如果一个地区能够观察到最多的遗传多样性,那么这个地区可能就是实验组的起源地,实验组在这里分化,然后再进一步从这里向外播散^[2]。本研究显示,中国3株土拉菌和日本FSC022一同位于B型系统进化结构中的基础位置,同时也表现出高水平的遗传多样性。那么是否可以认为,亚洲地区可能是subsp. *holarctica*的起源地?从1株全基因组测序的格鲁吉亚土拉菌中筛选出几个SNP,然后将这几个SNP放在全球系统进化分析的标记中,结果发现格鲁吉亚谱系被分配到B.Br.013内B.Br.026分支的基础位置,于是认为格鲁吉亚可能是B.Br.013更老的起源地^[6]。同理,将中国本土的土拉菌测序后补充新的SNP,放在subsp. *holarctica*的全球遗传进化背景下分析,可能会得出一些更接近于事实真相的结论。

参 考 文 献

[1] Gyuranecz M, Birdsell DN, Spletstoeser W, et al.

- Phylogeography of *Francisella tularensis* subsp. *holarctica*, Europe [J]. *Emerg Infect Dis*, 2012, 18(2): 290-293.
- [2] Vogler AJ, Birdsell D, Price LB, et al. Phylogeography of *Francisella tularensis*: global expansion of a highly fit clone [J]. *J Bacteriol*, 2009, 191(8): 2474-2484.
- [3] Vogler AJ, Birdsell DN, Lee J, et al. Phylogeography of *Francisella tularensis* ssp. *holarctica* in France [J]. *Lett Appl Microbiol*, 2011, 52(2): 177-180.
- [4] Svensson K, Granberg M, Karlsson L, et al. A real-time PCR array for hierarchical identification of *Francisella* isolates [J]. *PLoS One*, 2009, 4(12): e8360.
- [5] Pilo P, Johansson A, Frey J. Identification of *Francisella tularensis* cluster in central and western Europe [J]. *Emerg Infect Dis*, 2009, 15(12): 2049-2051.
- [6] Chanturia G, Birdsell DN, Kekelidze M, et al. Phylogeography of *Francisella tularensis* subspecies *holarctica* from the country of Georgia [J]. *BMC Microbiol*, 2011, 11: 139.
- [7] Wang YH, Hai R, Zhang ZK, et al. Genetic relationship between *Francisella tularensis* strains from China and from other countries [J]. *Biomed Environ Sci*, 2011, 24(3): 310-314.
- [8] Johansson A, Farlow J, Larsson P, et al. Worldwide genetic relationships among *Francisella tularensis* isolates determined by multiple-locus variable-number tandem repeat analysis [J]. *J Bacteriol*, 2004, 186(17): 5808-5818.
- [9] Larsson P, Svensson K, Karlsson L, et al. Canonical insertion-deletion markers for rapid DNA typing of *Francisella tularensis* [J]. *Emerg Infect Dis*, 2007, 13(11): 1725-1732.
- [10] Ahlinder J, Öhrman C, Svensson K, et al. Increased knowledge of *Francisella* genus diversity highlights the benefits of optimised DNA-based assays [J]. *BMC Microbiol*, 2012, 12: 220.
- [11] Fujita O, Uda A, Hotta A, et al. Genetic diversity of *Francisella tularensis* subspecies *holarctica* strains isolated in Japan [J]. *Microbiol Immunol*, 2008, 52(5): 270-276.
- [12] Vogler AJ, Birdsell D, Wagner DM, et al. An optimized, multiplexed multi-locus variable-number tandem repeat analysis system for genotyping *Francisella tularensis* [J]. *Lett Appl Microbiol*, 2009, 48(1): 140-144.
- [13] Wang YH, Peng Y, Hai R, et al. Diversity of *Francisella tularensis* subsp. *holarctica* lineages, China [J]. *Emerg Infect Dis*, 2014, 20(7): 1191-1194.
- [14] Keim P, Johansson A, Wagner DM. Molecular epidemiology, evolution, and ecology of *Francisella* [J]. *Ann N Y Acad Sci*, 2007, 1105: 30-66.
- [15] Farlow J, Wagner DM, Dukerich M, et al. *Francisella tularensis* in the United States [J]. *Emerg Infect Dis*, 2005, 11(12): 1835-1841.
- [16] Staples JE, Kubota KA, Chalcraft LG, et al. Epidemiologic and molecular analysis of human tularemia, United States, 1964-2004 [J]. *Emerg Infect Dis*, 2006, 12(7): 1113-1118.
- [17] García Del Blanco N, Dobson ME, Vela AI, et al. Genotyping of *Francisella tularensis* strains by pulsed-field gel electrophoresis, amplified fragment length polymorphism fingerprinting, and 16S rRNA gene sequencing [J]. *J Clin Microbiol*, 2002, 40(8): 2964-2972.

(收稿日期:2015-07-02)

(本文编辑:王岚)