

· 现场调查 ·

中国鼠疫菌种资源遗传特征分析

俞东征 海荣 董兴齐 李敏 夏连续 史献明 魏建春 崔百忠 王鹏
孙连芝 张志凯 胡源 张恩民

【摘要】目的 建立《中国甲类传染病菌种资源库》,选择适当的指标,以反映中国鼠疫菌种的遗传背景。方法 使用传统的和分子生物学方法,测定中国鼠疫菌种代表菌株的遗传特征。并对遗传特征的分布及其意义进行分析。结果 利用 6 项指标,将中国的鼠疫菌划分成 15 种遗传型,每一型占据一片相对独立的而又自相延续的地理空间。分析遗传型之间的亲缘关系,并根据菌株间的差异,追溯鼠疫耶尔森菌作为一个生物种形成后的进化过程。结论 中国各鼠疫自然疫源地存在着不同特征的鼠疫菌株,但为达到对菌株认同的目标,还需要确定更多的遗传指征。

【关键词】鼠疫菌;资源;遗传特征

Genetic analysis of *Yersinia pestis* strains isolated in China YU Dong-zheng*, HAI Rong, DONG Xing-qi, LI Min, XIA Lian-xu, SHI Xian-ming, WEI Jian-chun, CUI Bai-zhong, WANG Peng, SUN Lian-zhi, ZHANG Zhi-kai, HU Yuan, ZHANG En-min. *Institute for Infectious Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

【Abstract】Objective The strains of *Yersinia pestis* isolated in different period and different natural foci in China were analyzed. Methods Traditional and molecular biological methods were used. Rhamnose fermentation, rRNA gene copy number, nitrite reduction, and the glycerol fermentation were important characters for typing, and pulse field gel electrophoresis (PFGE) and random amplified polymorphic DNA (RAPD) profile could reflect the genetic distance between the strains. Results The strains could be divided into 15 genetic types by those 6 characters with each of them covered an isolated geographical territories. Conclusion The characters of strains were described; the genetic relationship of different types, their evolution, and the forming and shift of plague natural foci were analyzed.

【Key words】*Yersinia pestis*; Resource; Genetic characteristic

2001 年发表了第一组鼠疫菌全基因组序列,自此鼠疫研究进入后基因组时代。鼠疫菌全基因组序列,阐明了以往观察到的许多现象的 DNA 结构根源,但也提出了更多的问题^[1,2]。这些问题,需要通过鼠疫菌与其肠杆菌祖先之间的比较,以及大数量不同来源的鼠疫菌株之间的比较才能解决。我国自建国以来,坚持了全面的鼠疫预防控制和监测工作,并积累了大量的鼠疫菌株,构成了目前世界上最大的一份菌种资源。这份资源无论在其来源的广泛性、代表性和系统性方面,都是其他国家所无法比拟的。然而,这些鼠疫菌种只有在确定了遗传特征,并与分离时的背景资料相联系,才可能真正成为资源,用于追踪传染来源、理解鼠疫致病机理,以及理解鼠

疫发生和发展规律,并最终为控制鼠疫服务。在 1970 年代,我国曾在纪树立的组织和指导下,开展过全国性的鼠疫菌分型及鼠疫自然疫源地研究工作,并在 1983 年和 1984 年分别发表了这两项研究的结果(以下简称为 1983 年分型),确定了我国 8 种类型的鼠疫自然疫源地、鼠疫菌的 17 个生态型,以及鼠疫菌遗传特征因自然疫源地而异这一基本规律。本项研究系在上述分型研究的基础上,增加了近年来发展的分子生物学方法,为确立稳定的菌株识别特征,并在这些特征的基础上,建立起我国的鼠疫菌种资源库。

本项研究采用糖醇酵解试验、亚硝酸盐生成、营养需求和色素沉着试验等传统方法,以及 rRNA 指纹图、脉冲场电泳图、随机引物扩增图、甘油等基因测定与 IS100 插入位置测定等分子生物学方法,测定了不同年代分离自我国不同鼠疫自然疫源地的鼠疫菌株的遗传特征。这项工作共形成以下研究结果:中国鼠疫耶尔森菌的分子生物学特征与遗传学

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(俞东征、海荣、魏建春、张志凯、胡源、张恩民);云南省地方病防治研究所(董兴齐、王鹏);青海省地方病防治研究所(李敏、崔百忠);内蒙古地方病防治研究所(夏连续、孙连芝);河北省鼠疫防治研究所(史献明)

意义、固体平皿法对中国鼠疫各生态型菌株生化测定的实验研究、我国鼠疫菌营养类型及其流行病学意义、我国鼠疫耶尔森菌 Pgm 特征的研究、鼠疫耶尔森菌脉冲场凝胶电泳型分析、中国鼠疫菌随机引物扩增多态性指纹图谱、中国鼠疫菌核糖体型地理分布、布氏田鼠鼠疫菌 102 kb 基因座结构研究、鼠疫菌 102 kb pgm 基因座一端 IS100 的缺失与 Pgm⁺ 稳定性的研究、中国鼠疫菌 pMT 质粒的结构研究。在这些研究结果中搜集原始数据和图象资料,进行统计学分析,并结合鼠疫在我国的分布、致病严重程度、流行特征等流行病学资料进行遗传特征分析。

一、形成鼠疫菌种 F 型的突变

目前认为,鼠疫菌原为肠道菌,由假结核耶尔森菌 1b 血清型进化而来。在鼠疫菌的这种进化过程中,出现的第一种具有地区分布特征的突变,是丧失了利用鼠李糖的能力。鼠疫菌最近的祖先——假结核耶尔森菌具有酵解鼠李糖而产酸的能力,历史上曾长期将这一差异,作为区分鼠疫菌与假结核菌的指标。然而,在我国的鼠疫防制与监测中,确实发现了部分地区的鼠疫菌仍保留着这种古老的特征。目前,具有这种特征的鼠疫菌分布在 4 个相互孤立的区域:锡林格勒高原布氏田鼠疫源地、川青高原青海田鼠疫源地(依照 1983 年分型时确定的命名原则,为新发现的鼠疫自然疫源地)、天山山地灰旱獭疫源地和位于新疆洛浦县的一小片喜马拉雅旱獭疫源地。

由于插入序列 IS100 的转位,形成了鼠疫菌的 102 kb 基因岛,是与鼠李糖利用能力丧失密切相关的特征。在布氏田鼠、青海田鼠及天山的灰旱獭与长尾黄鼠的所有菌株中,均保留鼠李糖利用能力的同时,其 102 kb 基因岛仅一端具有 IS100 的插入。目前保存的新疆洛浦的鼠疫菌株也能利用鼠李糖,但均为 Pgm 阴性,且均可确定为 102 kb 基因岛完整删除所致。依此推断这种菌株应当具有 102 kb 基因岛另一端的 IS100 插入。在我国其他地区流行的鼠疫菌,属于具有完整的 102 kb 基因岛,且不能利用鼠李糖的菌株。其间也散在少数 102 kb 基因岛失去了一侧的 IS100,但仍表现为鼠李糖阴性的菌株,这些菌株应视为在 102 kb 基因岛形成后,又重新剪切失去了一端的 IS100,属于偶然的突变,不能构成独立的型别。

由于 IS100 插入造成的鼠疫菌染色体节段逆转,导致一个拷贝 rRNA 基因簇丧失,是鼠疫菌另一

种具有地区分布特征的突变。鼠疫菌在由肠杆菌进化产生时,也像肠杆菌一样,具有 7 拷贝 rRNA 基因簇。这些较为古老的菌株,目前在我国被分割成两大片及一个孤立的岛状地区:河西走廊以东的黄鼠、沙鼠和布氏田鼠鼠疫地区,自帕米尔延续到喜马拉雅及冈底斯山脉的旱獭鼠疫地区,以及川青交界处的青海田鼠鼠疫地区。在这类鼠疫菌株中,由于基因簇近侧限制酶位点的突变,长爪沙鼠疫源地中的鼠疫菌呈现独特的 rRNA 基因指纹图形,可以与其他 7 拷贝 rRNA 基因簇的鼠疫菌相区别。

在这些分布地区之间的大片区域:从天山山地、祁连山地、大部分青藏高原、滇西山地的野鼠鼠疫地区,直至我国南方的家鼠鼠疫地区,流行的鼠疫菌只拥有 6 拷贝 rRNA 基因簇。在我国鼠疫菌的实际进化过程中,这样的删除可能发生了两次,第一次发生在鼠李糖阳性的菌株中,形成了目前天山山地的菌型,第二次则形成了自祁连山地经青藏高原、滇西山地,直至我国华南地区的鼠疫疫源地。

鼠疫菌还原硝酸盐成为亚硝酸盐,即脱氮能力的丧失,是传统的鼠疫菌分型指标,也是第三种具有地区分布特征的突变。这种突变只产生在拥有 7 拷贝 rRNA 基因簇的菌株之中。国内其他研究的结果表明,这种突变由周质异化硝酸盐还原酶 A (napA) 基因中的点突变引起,并且确实证明了,鼠疫菌这种能力的丧失,为两次独立的、不相同的突变的结果。一次形成了目前在布氏田鼠和青海田鼠中的菌型,而另一次可能发生在昆仑山型的菌株,形成了现在的昆仑山 A 型、黄土高原型以及鄂尔多斯高原型。

最后,在失去了利用鼠李糖能力,失去了一拷贝 rRNA 基因簇,但仍保留着脱氮能力的鼠疫菌株中,产生了在全世界范围内最突出的一次突变:丧失了利用甘油的能力。形成了一个在我国只分布于南方地区,但却蔓延到世界各地的鼠疫菌型。国外研究表明,这一突变可能由甘油利用基因簇中两段不同的删除造成^[3]。本研究的结果表明,我国的甘油阴性鼠疫菌株,只由其中一段长 93 bp 的删除引起。

利用以上 4 种不同类型的指标,可以将我国的鼠疫菌划分成 8 种不同的类型,各占据一定的自然疫源地,并显示遗传特征的联系。

二、鼠疫菌株间的遗传学距离

与 1983 年分型的结果相比较,以上的 8 种鼠疫菌类型中,只有 3 种仅包括单一的鼠疫菌生态型。其余各类型,虽然遗传关系相近,却远隔在距离遥远

的疫源地内。

在本研究中,采用了两种非常敏感的、能够反映整个基因组内突变的方法,即脉冲场电泳图形(PFGE)^[4]与随机引物扩增图形(RAPD)^[5]。使用这些方法,除了在一次连续的流行过程中分离的菌株外,不同菌株间多少都能显示出差别,通过这些差别可以判断菌株间遗传距离的远近。在脉冲场电泳图中,虽然同为布氏田鼠或青海田鼠中分离的菌株,图形并不相同,但在一个菌株所显示的5~6条带中,至少存在4条相同的带,相反,在相毗邻的青海田鼠与喜玛拉雅旱獭中分离的菌株,带型却根本不同。按照菌株之间的相似值,可以定量地判断各鼠疫菌型之间遗传关系的远近;而按照这些相似值间断分布的部位,可以将目前已经研究的鼠疫菌,划分成5个不同的类型。在RAPD中,采用了RAPD1和RAPD2两条不同的引物序列,各能分出11种或6种不同的图形。这些图形之间的差异,也能反映鼠疫菌株之间的遗传距离,但实际的菌株分布表明,只有其中的部分典型图形,才具有明确的分型价值。采用这两种方法,可以分析一些局部的疫源地。

1. 天山山地:在1983年分型中,该地的鼠疫菌被划分成北天山东段、北天山西段A和北天山西段B三种生态型。这些生态型间,只在麦芽糖酵解能力和部分菌株的氨基酸需求方面存在区别。本研究结果表明,这些生态型内的变异十分复杂,惟有尼勒克县境内的鼠疫菌与其他地区分离的菌株,在RAPD1图形存在明显区别。因而,将原来的北天山东段型与北天山西段B型合并为北天山东段型,而只将生存于尼勒克县疫源地的原北天山西段A型称为北天山西段型。由于天山山地山高谷深,疫源地被分割在不同的山谷之中,鼠疫在这些疫源地中的流行历史久远且异常活跃,北天山东段型各菌株的RAPD1图形非常多样,可能由多种不同的亚型甚至型组成。

2. 昆仑山地:昆仑山地虽然是青藏高原的北部边缘,但生活在昆仑山地西段的鼠疫菌却与青藏高原型有较大的区别。自和田分离的昆仑A型中,部分菌株的RAPD图形接近帕米尔高原型,另一部分则为一个独特的型,昆仑山地便是这种类型菌株的主要集中地。分离自洛浦的昆仑B型,按照上一节的4项指标即可划分为一个独立的型,其遗传背景与帕米尔型接近。在若羌县分离的菌株虽被划归昆仑A型,但其遗传背景已经接近青藏高原型。

3. 帕米尔高原与冈底斯山脉:按照1983年分型,有4个生态型的菌株具有7拷贝rRNA基因簇,鼠李糖阴性,甘油阳性和脱氮阳性的特征。分别为松辽平原A、松辽平原B、帕米尔高原和冈底斯山型。帕米尔高原型是一个遗传背景十分相近的型,RAPD1的第5型菌株主要集中在这一地区。相反,冈底斯山型的遗传背景却非常复杂,在这一地区里几乎可以发现所有的RAPD图形。冈底斯山似乎也与天山山地相似,疫源地细碎分割,流行历史久远而且异常活跃,这一地区中也同样可能存在多数不同的型与亚型。

4. 祁连山地与青藏高原:在1983年分型中,以柴达木盆地界,祁连山地与柴达木盆地北缘的鼠疫菌被划归祁连山型,而柴达木南缘至昆仑山以上被划归青藏高原型。这两个生态型分布的自然环境非常接近,均以喜玛拉雅旱獭为主要储存宿主,而且除了部分菌株的麦芽糖酵解能力外,二者几乎没有区别。本研究表明,以祁连山分水岭为界,祁连山北麓肃南县疫源地的鼠疫菌才真正具有独立的遗传学特征,而其他原划归祁连山型的鼠疫菌,与青藏高原型并无本质区别。因而,只将肃南疫源地的菌株仍命名为祁连山型,其他如甘肃的肃北、阿克塞、青海的祁连、门源和柴达木东北边缘地区分离的菌株则转入青藏高原型菌株。

5. 布氏田鼠鼠疫与青海田鼠鼠疫:布氏田鼠鼠疫是一种非常特殊的鼠疫类型,按照脉冲场电泳图形所显示的差异,其鼠疫菌与其他类型鼠疫菌的区别,甚至可以达到亚种水平。青海田鼠鼠疫是一种新发现的鼠疫类型,其特征与布氏田鼠鼠疫有许多相似之处。本研究表明,虽然两片疫源地相距遥远,但两种鼠疫菌的遗传距离却十分接近,可能为其主要储存宿主选择所致。二者间差别在于麦芽糖酵解的程度不同,青海田鼠的鼠疫菌仅在液体发酵管中呈阳性反应,而在固体培养基上不能显示明显阳性;二者的RAPD2扩增图形也不相同。因此,本研究仍将它们划归两个不同的型。依照1983年确定的命名原则,将青海田鼠中的鼠疫菌定名为青川高原型。

6. 黄土高原:令人惊奇的是,黄土高原型与昆仑山A型的分布地区相距遥远,自然条件与宿主类型也根本不同,但其遗传背景却非常接近,利用上述指标几乎无法区分。它们之间的区别,只在于1983年分型中确定的营养需求及蛋白质电泳图形的差异。在本研究中,仍将它们列为不同的型。但黄土高原

A 型与 B 型之间,似乎难以形成两个独立的型,本研究将它们归并成为一个型。

7. 松辽平原:松辽平原的鼠疫菌遗传背景相互接近,可以归并成为同一个型。

8. 滇西纵谷与华南居民区:滇西纵谷型与青藏高原原型作为分型标志的 4 种突变完全相同,但无论从 PFGE 还是 RAPD 图形判断,二者的遗传距离均相距甚远。因此,鼠疫菌由旱獭进入滇西的小型野生啮齿动物,必定已经经历了非常久远的年代。相反,云南野鼠与家鼠中的鼠疫菌虽然酵解甘油的能力不同,遗传距离却相当接近,提示鼠疫进入家鼠在进化历史上还是不太久远的事件。家鼠类型的鼠疫菌现在已经侵入广西和贵州,仍有向东扩展的趋势。因而,本研究将这种类型的鼠疫菌改称华南居民区型。

三、中国鼠疫菌种遗传分型

综合考虑鼠疫菌的以上两部分识别标志,可以将我国的鼠疫菌种划分成 15 种遗传型,每一型占据一片相对独立而又自相延续的地理空间。表 1 列出这些型别与其标识特征。图 1 是中国鼠疫菌遗传类型的地域分布。

四、新产生的亚型

在本研究中使用的大多数分型特征,都没有发现与菌株分离的时间存在明确的关系,说明这些作为分型特征的突变,都发生在我国开展鼠疫监测之前的鼠疫菌进化的年代里。然而,也发现了 3 次与时间存在明确关系的事件。

一次发生在青海省同德县。在 1991 年以前,这一地区流行的菌株,虽被划分在祁连山型,但却具有青藏高原原型的 RAPD 特征。当 2001 年这里再次发生流行时,分离的菌株却全为来自祁连山北麓的祁连山型特征。这次菌株特征的转变只涉及已经存在的菌型,并未产生新的亚型。另一次发生在青海田鼠疫源地。1997 年发现这种类型的疫源地时,绝大多数的菌株为 RAPD1 的 2 型,与青藏高原原型的遗传距离较为接近。而从 2000 年再次调查开始,无论是在四川省石渠县还是次年在青海省境内,绝大多数菌株都表现为 RAPD1 的 3 型。这是一种在这一大

表1 中国鼠疫菌遗传类型

型别	标识特征						
	rRNA 基因数	鼠李糖	甘油	脱氮	RAPD1	RAPD2	PFGE
1北天山东段	6	+	+	+	多种类型		
2北天山西段	6	+	+	+	1		
3祁连山	6	-	+	+	1		2
4青藏高原	6	-	+	+	2		2
5滇西纵谷	6	-	+	+			3
6华南居民区	6	-	-	+			3
7锡林格勒高原	7	+	+	-		1	1
8川青高原	7	+	+	-	2,3	2	1
9昆仑山 B	7	+	+	-	3		
10昆仑山 A	7	-	+	-	5		
11黄土高原	7	-	+	-			
12松辽平原	7	-	+	+			
13帕米尔高原	7	-	+	+	3		
14冈底斯山	7	-	+	+	多种类型		4,5
15鄂尔多斯高原 (7 位点突变)		-	+	-	4		



数字代表表 1 中所列的鼠疫菌遗传类型
图1 中国鼠疫菌遗传类型的地区分布

的地区内没有出现过的遗传特征,因而,应当认为是川青高原型的第二亚型。

不同性质的突变事件发生在云南省。20 世纪 90 年代初在澜沧江下游谷地中开始,发现鼠疫菌获得了一种新的 6 kb 质粒。这种形式的菌株后来播散到省内广大地区,应当认为是华南居民区型的第二亚型。由于质粒可能丢失,不是一种好的分型特征。携有这种新型质粒,固然可以判断为新亚型,但在这种质粒出现后,没有这种质粒的菌株,却不能排除在新亚型之外。

五、本研究在鼠疫菌遗传特征方面的新发现

1983 年的分型工作采用了 12 项分型指标,本研究中省略了只起辅助的描述作用,或操作繁复,结果受实验条件的影响较大的指标。在本研究中,采用了 6 项指标,基本上肯定了 1983 年分型的结果。在传统的生物学分型工作中,早就发现同一菌株可能存在具有不同遗传特征的菌落。因而,在本研究中,糖醇酵解等工作普遍采用了固体平板法。结果表明,菌株的解离只是个别的现象,主要存在于松辽平原和云南省的早期菌株中,且只发生在麦芽糖和甘油等少数分型指征。使用平板法所获得的结果,绝大多数与以往的结果相符。这说明,1983 年分型所采用的主要指标,是稳定可靠的。

这些个别的菌株解离现象,也提供了鼠疫流行史的重要信息。国外研究已经表明,甘油酵解能力的缺失,为甘油-3-磷酸脱氢酶基因内部,两种不同的删节所致。本研究提示,甘油利用能力的突变,还存在着第三种形式。可能是以点突变的形式,在滇西纵谷型菌株中即已出现,并具有自发恢复能力,从而引起了菌株解离现象。不仅如此,发现弥渡县也曾分离到这样的菌株,提示滇西纵谷型菌株在历史上并不局限于目前剑川疫源地分布的地区,并可能引起过较大规模的流行。

在鼠疫菌最大的 pMT 质粒上,国外的研究表明存在两处 IS100 插入,其中一处插入的位置可变,在鼠疫菌东方变种菌株中,插入在 DNA 聚合酶基因中。本研究表明,这一位置是否存在 IS100 插入可以作为判别甘油阴性菌株应属于华南居民区型还是滇西纵谷型的辅助指标,而且表明,在滇西纵谷型与作为对照的青藏高原型中,这第二个位置可变的 IS100 并不存在。说明鼠疫菌的东方变种不是由中世纪变种,而是由古老变种直接进化而来的。

国外的研究已经表明,102 kb 基因岛的形成是

鼠疫菌成为一个生物种的重要标志。本研究表明,在我国存在着多个明确的鼠疫疫源地区,鼠疫菌中 102 kb 基因岛一侧的 IS100 尚未插入。这不仅能够成为这些地区鼠疫菌的重要识别标志,而且还与一些鼠疫疫源地中独特的流行特征密切相关。

在目前已经测定的鼠疫菌全基因组序列显示,东方变种的鼠疫菌发生了一拷贝的 rRNA 基因簇删除。本研究表明,这种删除实际上在鼠疫菌进化的非常早期即已发生,遍及古老、中世纪和东方三个不同变种的菌株。这些结果还提示,鼠疫菌的进化过程在我国境内一直在由西向东发展,由此推断,鼠疫可能起源于我国以西地区。

六、今后工作方向

本研究中采用的部分分子生物学方法,还不能与序列资料直接联系,从而明确分型标识的位置。IS100 定位研究在本研究中已经取得初步的成效,串联重复序列多态性分析已开始进行,但还没有体现在本研究中。应该尽快发展这些可以精确定位的方法,以使鼠疫菌种资源的遗传背景具有更高的明确性。

由于非典型菌株的存在,本研究距单个菌株认同的目标还有一定距离。需要在今后的工作中,为每一项分型指标找到多种具有等同效力的标识,才能达到可靠判断每一菌株来源的目的。102 kb 基因岛一侧的 IS100 插入和 pMT 质粒上的第二处 IS100 插入,分别为鼠李糖利用与甘油利用提供了同等效力标识,是这方面工作的良好开端。在今后的工作中,还需要搜索更多的这种特征。

全基因组序列的出现已经为达到这些目标提供了条件,我们需要充分利用我国的菌种资源优势,达到对鼠疫菌遗传结构更深入的认识。

参 考 文 献

- 1 Parkhill J, Wren BW, Thomson NR, et al. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature*, 2001, 413: 523-527.
- 2 Deng W, Burland V, Plunket III G, et al. Genome sequence of *Yersinia pestis* KIM. *J Bacteriol* 2002, 184: 4601-4611.
- 3 Motin VL, Georgescu AM, Elliott JM, et al. Genetic variability of *Yersinia pestis* isolates as predicted by PCR-based IS100 genotyping and analysis of structural genes encoding glycerol-3-phosphate dehydrogenase (glpD). *J Bacteriol*, 2002, 184: 1019-1027.
- 4 Williams JGK, Kubelik AR, Livak KJ. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*, 1990, 18: 6531-6535.
- 5 Schwartz DC, Cantor CR. Separation of yeast chromosome-sized DNA by pulsed field gradient gel electrophoresis. *Cell*, 1984, 37: 67-75.

(收稿日期 2003-06-18)

(本文编辑:张林东)