

鼠疫耶尔森菌的基因组学和后基因组学研究进展*

万成松,江凌晓 综述 俞守义 审校

(第一军医大学热带军队卫生学系微生物学教研室,广州 510515)

摘要:鼠疫耶尔森菌能导致致病性极强的鼠疫,可作为生物武器使用。鼠疫耶尔森菌的基因组全长 4.65Mb,GC 含量为 47.6%,含有 3 个重要质粒 pFra/pMT1、pPst/pPCP1 和 pYV1/pCD1,其中质粒 pFra/pMT1 和 pPst/pPCP1 为鼠疫耶尔森菌独有。鼠疫耶尔森菌基因组富含大量的插入序列和假基因,存在频繁的基因内重组,以水平转移的方式获得外源基因。鼠疫耶尔森菌基因组结构和功能的研究为鼠疫的发生、流行、暴发、致病机制研究和筛选鼠疫耶尔森菌治疗药物、研制疫苗提供理论基础和科学依据。

关键词:鼠疫耶尔森菌;基因组学和后基因组学;进展

中图分类号:R37 **文献标识码:**A **文章编号:**1005-5673(2003)-01-0031-06

鼠疫(Plague)是一种自然疫源性的烈性传染病。历史上曾记载过 3 次世界范围的灾难性人间鼠疫流行,分别发生于公元 6~8 世纪,14~17 世纪,19 世纪末~20 世纪初,死亡人数过亿。20 世纪前半叶(1900~1949 年),我国先后发生了 6 次较大的鼠疫流行。此后,大规模的人间鼠疫已基本控制,但散发病例或小规模流行仍有报道。

人类鼠疫多由鼠疫耶尔森菌(*Yersinia pestis*)引起,主要经鼠蚤叮咬而传播,传染性强,病死率高,危害极大。耶尔森菌属至少包括 11 种细菌,其中鼠疫耶尔森菌、假结核耶尔森菌(*Yersinia pseudotuberculosis*)和小肠结肠炎耶尔森菌(*Yersinia enterocolitica*)与人类致病关系较为密切。假结核耶尔森菌和小肠结肠炎耶尔森菌经食物和水源传播,在人类可引起肠道疾病。

鼠疫菌历来被外军列入重要生物战剂范畴,也是国际恐怖分子采用的廉价而方便的资源,威胁国家和社会稳定。

1 鼠疫耶尔森菌基因组学

1998 年,英国 Sanger 中心 Julian Parkhill 和 Brendan Wren 领导的实验室从一位接触了感染鼠疫耶尔森菌的猫的肺鼠疫病人体内分离了一株鼠疫耶尔森菌 CO92 型,并对其全基因组进行测序。最近,完成了全基因组序列测定工作^[1]。鼠疫耶尔森菌的基因组全长 4.65Mb(GeneBank 存取号:NC-003143),GC 含量为 47.6%,含有 3 个重要质粒

pFra/pMT1、pPst/pPCP1 和 pYV1/pCD1。在此之前,Perry^[2]、Lindler^[3]等也公布了鼠疫耶尔森菌 KIM5 株的质粒 pYV1/pCD1、pFra/pMT1 和 pPst/pPCP1 的 DNA 序列,pYV1/pCD1 质粒为鼠疫耶尔森菌、假结核耶尔森菌和小肠结肠炎耶尔森菌等耶尔森菌共有,质粒 pFra/pMT1 和 pPst/pPCP1 为鼠疫耶尔森菌独有^[4,5]。

1.1 鼠疫耶尔森菌 CO92 株全基因组结构

鼠疫耶尔森菌 CO92 株全基因组结构^[1]见图 1。图示的最外层为鼠疫耶尔森菌基因组的碱基数(单位:Mb),按从外到内的顺序,第 1、2 圈为正负双链;第 3、4 圈为假基因;第 5、6 圈为插入序列,其中蓝色表示 IS1661,黑色表示 IS285,红色表示 IS1541,绿色表示 IS100;第 7 圈为 G+C 含量(值高向外);第 8 圈为 GC 误差(G-C/G+C),黄褐色表示值大于 1,紫色表示值小于 1。其中不同颜色所示为功能基因,黑蓝色为致病性基因,黑色为能量代谢基因,红色为信息传递基因,黑绿色为表面联合基因,蓝绿色为大分子降解基因,红紫色为小分子降解基因,黄色为中间代谢基因,蓝白色为调节基因,橙色为保守的假基因,褐色为假基因,粉红色为噬菌体插入序列,绿色为未知基因,灰白色为混杂基因。

鼠疫耶尔森菌的一般生物学特征见表 1:

1.2 鼠疫耶尔森菌 pYV1/pCD1 的结构

Perry 等^[2]对鼠疫耶尔森菌 KIM5 株的低钙反应(low-Ca²⁺-response,LCR)质粒 pCD1 DNA 进行

* 基金项目:国家“863”计划课题(2001AA223061)

作者简介:万成松(1965-),男,硕士,副教授,主要从事病原微生物基因诊断、结构与功能研究。gzwcs@fimmu.edu.cn

了序列分析,质粒 pCD1 DNA 总长度为 70 559bp (GeneBank 存取号: NC-001972), GC 含量为 44.8%,含有 60 个开放读码框(Open reading frame,

ORF)、3 种 IS (IS100、IS1616 和 IS1617) 和大量的缺失或转位元件,其中 47 个 ORF 被证实与 LCR 的结构组成有关。具体结构见图 2。

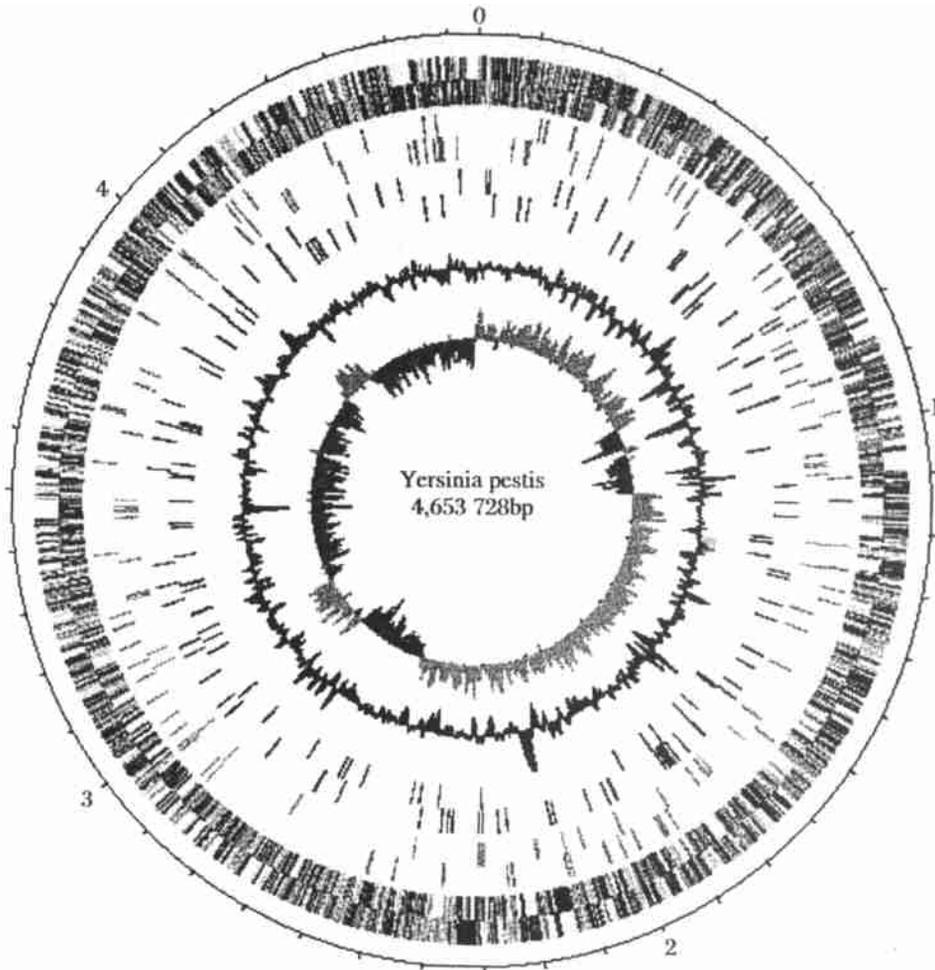


图 1 鼠疫耶尔森菌 CO92 株全基因组结构图

表 1 鼠疫耶尔森菌的一般生物学特征

	染色体	pPst/pPCP1	PYV1/pCD1	pFra/pMT1
拷贝数		186	4.3	1.8
大小	4 653 728bp	9 612bp	70 559bp	96 210bp
G+C 含量	47.64 %	45.27 %	44.84 %	50.23 %
编码序列	4 012	9	97	103
假基因	149	0	8	3
编码率	83.8 %	57.2 %	81.4 %	86.8 %
基因平均长度	998 bp	611bp	643bp	835bp
rRNA	6			
tRNA	70			
其他 RNA	6			

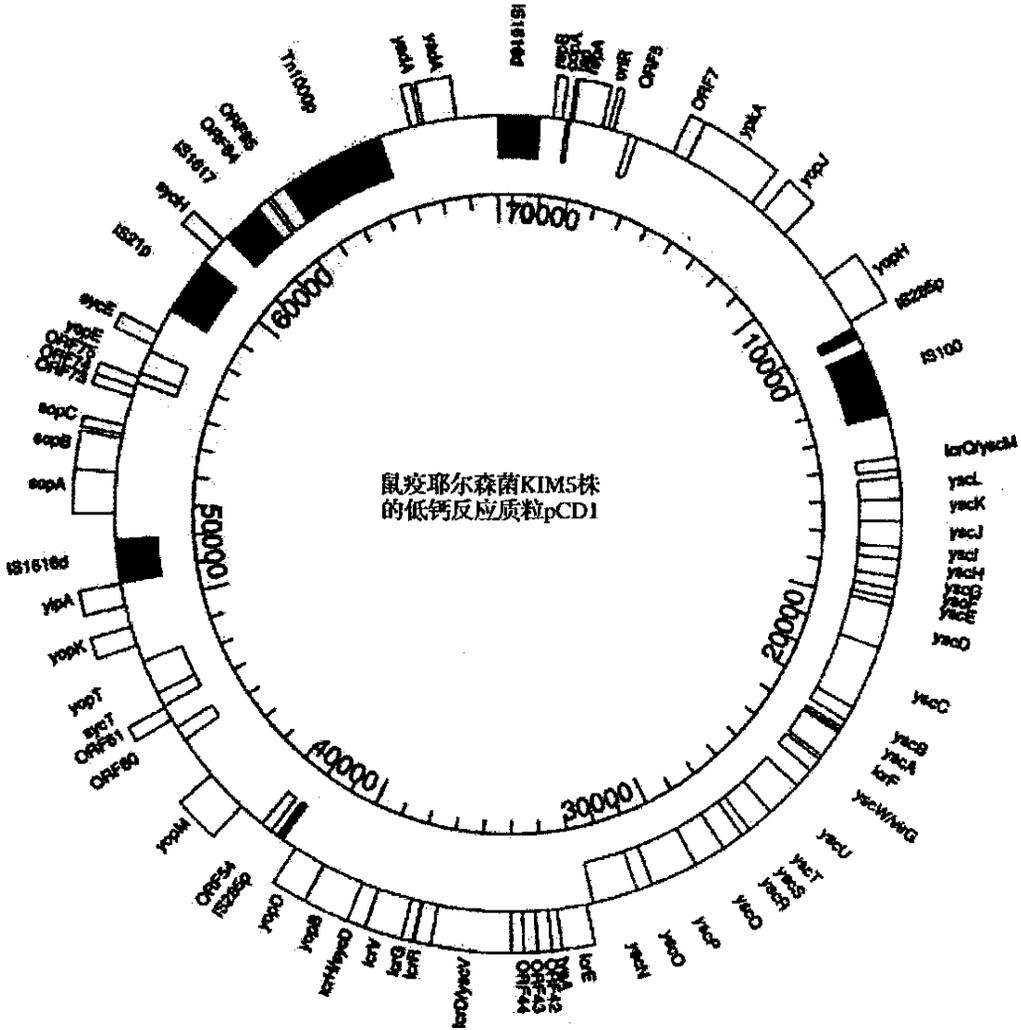


图 2 鼠疫耶尔森菌质粒 pYVI/ pCD1 的基因组成结构

1.3 质粒 pPst/pPCP1 的结构

质粒 pPst/pPCP1 DNA 的总长度为 9 610bp (GeneBank 存取号: NC-001881), GC 含量为 45.3%^[5]。整个质粒含有 5 个潜在 ORF: 一个 IS100 插入序列、鼠毒素 (Pesticin) 基因、鼠毒素免疫蛋白基因、纤维蛋白溶酶原激活因子基因和一个与 *E. coli* 高度同源的 CoIE1 复制子 (3 119 ~ 3 899bp), 分别编码鼠毒素、鼠毒素免疫蛋白和纤维蛋白溶酶原激活因子等。具体结构见图 3。

1.4 质粒 pFra/pMT1 的结构

质粒 pFra/pMT1 DNA 总长度为 100 984bp (GeneBank 存取号: AF053947), GC 含量为 50.2%^[7]。经过 GeneQuest 软件分析发现整个质粒含有 145 个潜在 ORF, 具体结构见图 4。

2 鼠疫耶尔森菌的后基因组学

随着人类基因组计划的完成和相关模式生物基因组测序工作的不断深入, 生命科学已经进入了后基因组 (Post-genome) 时代。人们在鼠疫耶尔森菌基因组测序的同时就开始了鼠疫耶尔森菌的后基因组学研究。

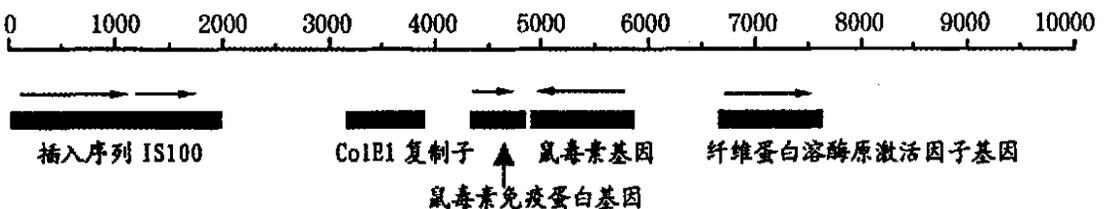


图 3 鼠疫耶尔森菌质粒 pPst/ pPCP1 的基因组成结构

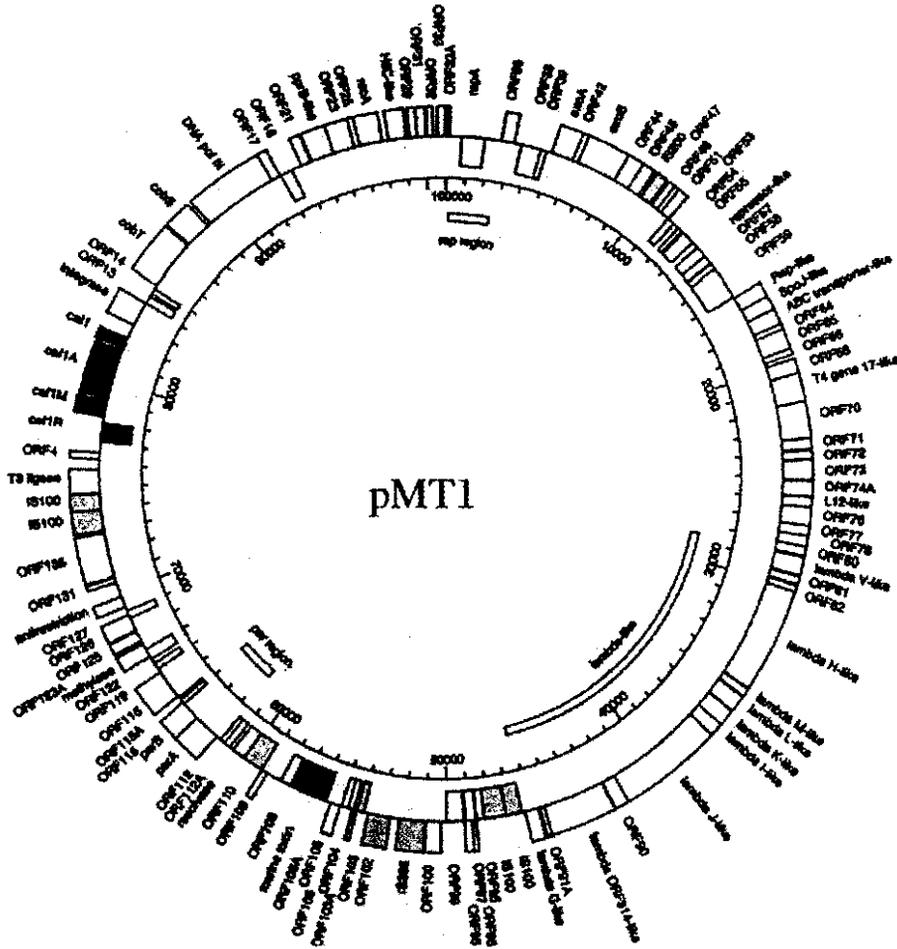


图 4 鼠疫耶尔森菌质粒 pFra/pMT1 的基因组成结构

2.1 鼠疫耶尔森菌染色体上基因的功能

从生物进化和群体遗传学的角度^[6],比较鼠疫耶尔森菌和假结核耶尔森菌的看家基因和历史上出现的 3 次世界范围的人间鼠疫大流行情况,发现鼠疫耶尔森菌和假结核耶尔森菌在鼠疫暴发和间歇期交互出现,其可能为同一菌种的 2 个亚种。对这 2 种耶尔森杆菌的比较基因组学研究表明,鼠疫耶尔森菌和假结核耶尔森菌的 DNA 同源率为 83%,认为鼠疫耶尔森菌是 1500~2000 年前从假结核耶尔森菌的 O:1b 血清型进化而来的。

获得外源基因在鼠疫耶尔森菌的进化过程中是非常重要的。临床分离的一些致病性鼠疫耶尔森菌内总存在一个 100kb 左右的外来质粒。人们自然推测鼠疫的暴发可能与 pFra/pMT1 质粒和 pPst/pPCP1 质粒在鼠疫耶尔森菌和假结核耶尔森菌内的跳转和水平转移有关,该质粒上可能存在某些毒力基因。

在鼠疫耶尔森菌染色体上显示一些特征岛,其中有些基因似乎来源于昆虫病原体。编码鼠毒素

Ymt 的质粒就是跳蚤生活所必需^[8]。昆虫体内一些编码杀虫毒素复合物的基因 *tcaA/tcaB/tcdA*, *tcaC/tcdB/tccC* 等不同程度地在鼠疫耶尔森菌中存在^[9]。更有趣的是,PCR 结果显示,*tca* 杀虫毒素基因也在假结核耶尔森菌的 IP32953 中存在,这说明在鼠疫耶尔森菌出现之前,假结核耶尔森菌已同环境中的昆虫有一定的联系。

序列分析显示^[1],在鼠疫耶尔森菌基因组内存在一些编码表面抗原、纤毛或粘附素的基因,已知 *caf* 基因编码 F1 抗原,还有 8 个和 *psa* 基因和 *caf* 基因操纵子相似的结构。大量编码不同纤毛或粘附素的基因可能有利于其侵入宿主。

假结核耶尔森菌的基因组测序工作正在美国加利福尼亚 Lawrence Livermore 国家实验室 (<http://bbrp.llnl.gov/bbrp/html/microbe.html>) 进行,通过基因组学比较表明,鼠疫耶尔森菌以水平基因获得的方式从假结核耶尔森菌获得大量基因。鼠疫耶尔森菌基因组 GC 含量较高,且含有大量的 IS,说明在鼠疫耶尔森菌染色体中可能存在着频繁的基因内重

组,许多基因簇包含的粘附素基因、分泌系统基因和毒素基因似乎是从其它细菌获得。

鼠疫耶尔森菌在水平获得外源基因的同时,也存在自身基因的丢失和突变而形成假基因^[1]。基因组包含大约 150 个假基因,其中 51 个假基因是被 IS 所中断形成的,58 个假基因是由于漂移突变造成的,32 个是由于缺失所致,一些是在其进化过程中

适宜于小肠内的残留基因。假结核耶尔森菌能特异地粘附在小肠表面并侵入细胞内,其 *yadA* 和侵入素基因起了重要作用^[10],而这 2 个基因在鼠疫耶尔森菌中则以假基因的形式存在。假结核耶尔森菌具有运动功能,而鼠疫耶尔森菌没有,通过分析发现,在鼠疫耶尔森菌的 2 个鞭毛基因簇和趋化基因簇中,鞭毛基因簇存在 6 个基因突变(图 5)。

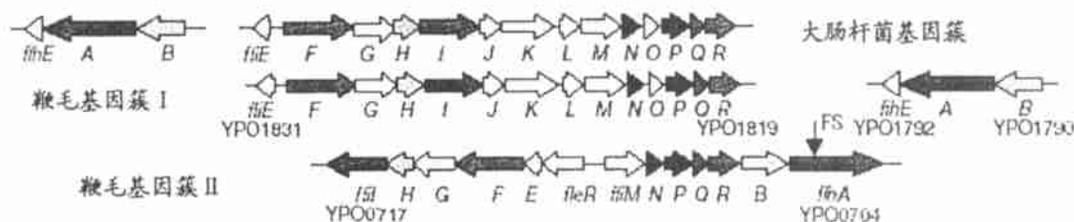


图 5 鼠疫耶尔森菌的鞭毛基因簇上的突变

水平基因获得和基因流动的事实证明了鼠疫耶尔森菌是一个正在进行大规模基因流动的病原,也正是在这种基因流动中进化为一种新的强毒病原菌。

2.2 质粒 pFra/pMT1 上基因的功能

质粒 pFra/pMT1 可高频出现并整合到鼠疫耶尔森菌的染色体上,质粒 pFra/pMT1 具有很强的毒性。有实验证明丢失该质粒毒株,其致病性降低。

使用 BLAST 对质粒 pMT1 进行分析^[5],145 个潜在 ORF 中发现 115 个 ORF 是可编码区,其中含有 2 个 IS100、1 个 IS1541、1 个 IS285、2 个转座子,已知的 5 个编码 F1 荚膜抗原、F1 荚膜锚定(Anchoring)蛋白、Caf1M 抗原、Caf1R 调节蛋白和鼠毒素(Murine toxin, MT)的基因也在 pMT1 质粒上。这 5 个重要基因位于一个大小为 18kb 的被称为毒力岛(Pathogenicity islands)区域(67 669 ~ 85 595)^[5,11]。并且 115 个 ORF 中有 30 个与一些细菌的已知基因有同源性,分别相似于编码 *E. coli* 的鞭毛蛋白和噬菌体宿主特异性蛋白的基因、噬菌体 p7 *parA* 和 *parB* 操纵子、假单胞菌 *cobS* 和 *cobT* 基因等。通过软件 MotifFinder 在 PROSITE 数据库中发现了 14 个基序,一些基序还包含有 ATP/GTP 结合定点(P 环)、细胞附着位点(Cell attachment site)、ABC(ATP binding cassette)转座子、54 相互作用域(Domain)。此外,通过 GeneMark 软件预测,还有 44 ORF 在目前相应的数据库中找不到已知的或假设的同源性序列,大量的 ORF 还没得到认识,它的结构与功能的关系引起人们极大的兴趣。

2.3 质粒 pYV1/pCD1 上基因的功能

质粒 pYV1/pCD1 含有大量的与 LCR 有关的 ORF^[2],其中的 35 个 ORF 连在一起构成 LCR 基因簇。在 37 °C 缺 Ca²⁺ 情况下,鼠疫耶尔森菌的 LCR 基因能强烈表达和分泌毒力蛋白 LcrV 和 Yop。虽然不能将 LCR 基因簇明确地定义为一个毒力岛,但是经过 BLAST 搜索发现质粒 pYV1/pCD1 含有大量与已知毒力基因同源的序列,如 LcrV、Yop、一个类似 IncFIIA 的复制子和一个类似 SopABC 的阻隔片段区。

质粒 pYV1/pCD1 基因组上的一个著名特征是含有 2 个经过漂移突变假基因,一个原是编码粘附素的 *yadA*,一个原是存在于假结核耶尔森菌 pYV 质粒上编码脂蛋白信号肽的 *ylpA*。

质粒 pYV1/pCD1 的 Yop 蛋白、Yop 转位蛋白、Yop 伴随蛋白、V 抗原等和鼠疫耶尔森菌的毒力有关^[4]。

3 鼠疫耶尔森菌的基因组学和后基因组学研究的意义

鼠疫耶尔森菌的基因组结构和功能研究有助于更精确地研究鼠疫耶尔森菌的形态、结构、代谢、遗传变异和进化,寻找致病相关基因和药物作用靶位,为研制特异的诊断试剂、有效药物和保护性强的疫苗提供科学依据。

3.1 建立灵敏、特异和快速的鼠疫耶尔森菌分子生物学诊断技术

鼠疫耶尔森菌基因组测序工作的完成,为研究鼠疫耶尔森菌的基因组结构和功能打下了坚实的基础,更为建立鼠疫耶尔森菌分子生物学检测技术提供了丰富的基因水平的信息。人们可以根据鼠疫耶

尔森菌染色体和相关质粒基因组上的保守序列,设计探针和引物,采用杂交技术和 PCR 技术快速、特异检测鼠疫耶尔森菌,也可将鼠疫耶尔森菌相关基因和蛋白质以微阵列的形式制成基因芯片和蛋白质芯片大规模检测鼠疫耶尔森菌的基因谱和蛋白表达谱,在疫源地鼠疫耶尔森菌快速侦检和反生物恐怖方面具有重要意义。

3.2 研究鼠疫耶尔森菌的致病机制及其与宿主的相互关系

鼠疫耶尔森菌能产生很强的致死性疾病,而假结核耶尔森菌对人体产生的疾病却是温和的,为什么两者在对人类的致病性方面有如此大的差距呢?一般认为,鼠疫耶尔森菌基因组上存在许多毒力基因,编码一些特殊的毒力蛋白质或蛋白酶,它们在人体的代谢过程中发挥着重要作用。那么这些毒力基因在鼠疫耶尔森菌的基因结构和序列有什么特点?各编码什么毒力蛋白质或蛋白酶?这些蛋白如何通过信号传导与宿主细胞的某个受体相互作用而影响细胞的生物学功能?其在宿主细胞中的调控与表达是怎样启动和激活的?这都是人们十分感兴趣的问题。尽管这些疑问需得到实验证实,但是我们认为这些基因对鼠疫的发生和流行至关重要。

3.3 更快捷地设计和开发鼠疫耶尔森菌治疗药物和疫苗

鼠疫耶尔森菌基因组测序工作完成之后,可通过生物信息学的方法寻找抗鼠疫耶尔森菌药物作用的靶位和编码鼠疫耶尔森菌保护性抗原的基因,通过相关生物软件进行药物设计和疫苗设计,可缩短药物和疫苗研制的时间,是未来药物开发和疫苗设计的一种十分可行的方法。

3.4 研究鼠疫耶尔森菌毒力质粒(基因)的获得及其进化关系

基因和毒力质粒的跳转和水平转移在细菌的进化过程中非常重要。鼠疫耶尔森菌外源毒力质粒(基因)的获得、漂移和突变,使细菌对新环境产生一定的适应性,在鼠疫耶尔森菌的进化过程中,具有十分重要的意义。历史上几次鼠疫流行的资料表明毒力质粒的获得和毒力基因的出现与鼠疫的发生和流行有一定的相关性。

当假结核耶尔森菌和小肠结肠炎耶尔森菌的基因组测序工作完成之后,可通过比较基因组学的手段,研究鼠疫耶尔森菌、小肠结肠炎耶尔森菌和假结核耶尔森菌的起源、进化和种系发育等,为鼠疫的发生、流行、暴发、致病机制研究和筛选鼠疫耶尔森菌

治疗药物、研制疫苗提供分子水平的理论基础和科学依据。

4 鼠疫耶尔森菌主要相关网站

www.sanger.ac.uk/Projects/Y-pestis/

www.bacteriamuseum.org/species/pestis.shtml

shtml

www.cdc.gov/ncidod/dvbid/plague/

www.medmicro.mds.qmw.ac.uk/yersinia

magpie.genome.wisc.edu/browser/Yersinia

pestis-circle.html

www.ncbi.nlm.nih.gov/cgi-bin/Entrez/

framik?db=Genome&gi=201

参考文献:

- [1] Pakhill J, Wren BW, Thomson NR, et al. Genome sequence of *Yersinia pestis*, the causative agent of plague [J]. *Nature* 2001, 413(6855):523-527.
- [2] Perry RD, Straley SC, Fetherston JD, et al. DNA sequencing and analysis of the low-Ca²⁺ response plasmid pCD1 of *Yersinia pestis* KIM5 [J]. *Infect Immun* 1998, 66(10):4611-4623.
- [3] Lindler LE, Plano GV, Burland V, et al. Complete DNA sequence and detailed analysis of the *Yersinia pestis* KIM5 plasmid encoding murine toxin and capsular antigen [J]. *Infect Immun* 1998, 66(12):5731-5742.
- [4] Cornelis GR, Boland A, Boyd AP, et al. The virulence plasmid of *Yersinia*, an antihost genome [J]. *Microbiol Mol Biol Rev* 1998, 62(4):1315-1320.
- [5] Hu P, Elliott J, McCready P, et al. Structural organization of virulence-associated plasmids of *Yersinia pestis* [J]. *J Bacteriol* 1998, 180(19):5192-5202.
- [6] Achtman M, Zuth K, Morelli G, et al. *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis* [J]. *Proc Natl Acad Sci USA* 1999, 96(24):14043-14048.
- [7] Youngren B, Radnedge L, Hu P, et al. A plasmid partition system of the P1-P7 par family from the pMT1 virulence plasmid of *Yersinia pestis* [J]. *J Bacteriol* 2000, 182(14):3924-3928.
- [8] Hinnebusch J, Cherepanov P, Du Y, et al. Murine toxin of *Yersinia pestis* shows phospholipase D activity but is not required for virulence in mice [J]. *Int. J. Med. Microbiol* 2000, 290:483-487.
- [9] Waterfield NR, Bowen DJ, Fetherston JD, et al. The tc genes of *Photobacterium*: a growing family [J]. *Trends Microbiol* 2001, 9(4):185-191.
- [10] Clarissa C, Heathera J, yasemin H, et al. Invasion of Epithelial Cells by *Yersinia pestis*: Evidence for a *Y. pestis*-Specific Invasin [J]. *Infect Immun*, 1998, 66(12):5731-5742.
- [11] Rakin A, Noelting C, Schropp P, et al. Integrative module of the high-pathogenicity island of *Yersinia* [J]. *Mol Microbiol*. 2001; 29(2):407-415.

(收稿 2002-04-24)