

# 基于 CRISPR/Cas 的大肠埃希菌分子标志物的监测研究

梁文娟 张荣光 段广才 洪丽娟 张冰 郝园林 杨海燕 陈帅印  
娄婷叶 赵永新

450001 郑州大学公共卫生学院流行病与卫生统计学系(梁文娟、张荣光、段广才、洪丽娟、张冰、郝园林、杨海燕、陈帅印); 453003 新乡医学院分子诊断与医学检验技术河南省协同创新中心(梁文娟、段广才); 453100 新乡医学院第一附属医院检验科(娄婷叶); 453003 新乡医学院第三附属医院检验科(赵永新)

通信作者:段广才, Email:gcduan@zzu.edu.cn

DOI: 10.3760/cma.j.issn.0254-6450.2016.08.005

**【摘要】** 目的 探讨基于 CRISPR/Cas 的大肠埃希菌分子标志物的监测研究。方法 通过 BLAST 收集 GenBank 数据库中 135 株全基因组测序大肠埃希菌、203 株鸟枪法测序大肠埃希菌的 CRISPR/Cas 和 PCR 扩增、测序获得本实验室保存 361 株大肠埃希菌(包括 38 株大肠埃希菌 O157:H7)的 CRISPR 序列,应用 CRISPR Finder 在线软件分析 CRISPR 特征、DNAMAN 软件进行间隔序列的比对,使用 Clustal X 进行 *cas* 多序列比对和 Mega 5.1 软件构建系统进化树。结果 本研究以全新的视角对大肠埃希菌的 CRISPR/Cas 位置进行描述;结果显示,135 株全基因组测序、203 株鸟枪法测序和 361 株本实验室测序的大肠埃希菌中分别有 77.04%、100.00% 和 75.62% 的大肠埃希菌具有 CRISPR1, 分别有 74.81%、100.00% 和 92.24% 的大肠埃希菌具有 CRISPR2, 分别有 11.85%、0 和 1.39% 的大肠埃希菌具有 CRISPR3 和 CRISPR4; GenBank 数据库下载的全基因组测序的 1 株和本实验室测序的 2 株大肠埃希菌存在 4 个 CRISPR 位点;缺少 *cas* 的 CRISPR1 下游有插入序列存在。在 699 株大肠埃希菌中,8 株 O55:H7、180 株 O157:H7、8 株 O157:HNM、40 株 O104:H4、4 株 O145:H28 有独特的 CRISPR;间隔序列的缺失可发生在 CRISPR 中间;依据 I-E 和 I-F 的 *cas* 构建系统发育树,均可分为两类。结论 大肠埃希菌的 CRISPR/Cas 可能作为鉴定强毒株大肠埃希菌或者新型菌株的分子标志物。间隔序列的缺失或获得可能与噬菌体有关。

**【关键词】** 大肠埃希菌;成簇的规律间隔短回文重复序列;分子标志物

**基金项目:** 国家科技重大专项(2013ZX10004607); 新乡医学院分子诊断与医学检验技术河南省协同创新中心(XTCX-2015-PY4)

**A surveillance study on CRISPR/Cas molecular biomarker in *Escherichia coli*** Liang Wenjuan, Zhang Rongguang, Duan Guangcai, Hong Lijuan, Zhang Bing, Xi Yuanlin, Yang Haiyan, Chen Shuaiyin, Lou Tingye, Zhao Yongxin

Department of Epidemiology, College of Public Health, Zhengzhou University, Zhengzhou 450001, China (Liang WJ, Zhang RG, Duan GC, Hong LJ, Zhang B, Xi YL, Yang HY, Chen SY); Henan Innovation Center of Molecular Diagnosis and Laboratory Medicine, Xinxiang Medical University, Xinxiang 453003 (Liang WJ, Duan GC); Laboratory Department, The First Affiliated Hospital of Xinxiang Medical University, Xinxiang 453100, China (Lou TY); Laboratory Department, The Third Affiliated Hospital of Xinxiang Medical University, Xinxiang 453003, China (Zhao YX)

Corresponding author: Duan Guangcai, Email: gcduan@zzu.edu.cn

**【Abstract】** **Objective** A new method related to molecular biomarker with CRISPR/Cas (clustered regularly interspaced short palindromic repeats-cas) in *Escherichia (E.) coli* was developed and used for surveillance programs. **Methods** CRISPR/Cas sequence that containing 135 strains with complete sequence and 203 strains with whole genome shotgun sequence of *E. coli* in GenBank by BLAST and 361 strains of *E. coli* (including 38 strains of *E. coli* O157:H7) in laboratory were identified by PCR and analyzed with the CRISPR Finder. Spacers were compared with DANMAN and

the phylogenetic trees of *cas* gene were constructed under Clustal X and Mega 5.1. **Results** With new perspective, a descriptive method was developed targeting on the position of CRISPR/*cas* in *E. coli*. The CRISPR1 was detected in 77.04%, 100.00% and 75.62% and the CRISPR2 was detected in 74.81%, 100.00% and 92.24% and the CRISPR3 and CRISPR4 were detected in 11.85%, 0 and 1.39% for 135 strains with complete sequence, 203 strains with whole genome shotgun sequence and 361 strains in the laboratory, respectively. One strain downloaded in GenBank with whole genome sequencing and 2 strains in the our laboratory were identified that containing four CRISPR locus. The other *E. coli* strain was with insertion sequence in downstream of the non-*cas* CRISPR1. The unique CRISPR was found in 8 strains of O55 : H7, in 180 strains of O157: H7, in 8 strains of O157 : HNM, in 40 strains of O104 : H4, in 4 strains of O145 : H28, in all the 699 *E. coli* strains. The phylogenetic tree could be divided into two groups—*cas* with type I-E or type I-F. **Conclusions** CRISPR/*Cas* might be used as a valuable molecular biomarker in epidemiological surveillance studies to identify the high virulent strains or new strains of *E. coli*. Phage might be related to the missing or obtaining of spacers.

**【Key words】** *Escherichia coli*; Clustered regularly interspaced short palindromic repeats/*cas*; Molecular biomarker

**Fund programs:** National Science and Technology Major Project of China (2013ZX10004607); Henan Provincial Innovation Center of Molecular Diagnosis and Laboratory Medicine (XTCX-2015-PY4)

基因的水平转移是生物进化的一种重要机制,涉及到新的生物性状和新型菌株的形成<sup>[1]</sup>。近年来发现引起疫情暴发的新型菌株肠出血性大肠埃希菌 O157 : H7 和 O104 : H4 分别是肠致病性大肠埃希菌 O55 : H7 和肠黏附性大肠埃希菌 55899 接受可移动元件的转移形成<sup>[2-3]</sup>。因此,探讨基因的水平转移机制和鉴定、判断强毒株大肠埃希菌或者新型菌株,追溯其感染来源是分子流行病学的重要任务。成簇的规律间隔短回文重复序列 (clustered regularly interspaced short palindromic repeats, CRISPR) 主要由重复序列 (repeat) 和间隔序列 (spacer) 组成,其与 CRISPR 相关基因 (CRISPR-associated genes, *cas*) 共同构成 CRISPR/*Cas* 系统。一个细菌中可有多个 CRISPR 位点,在每个 CRISPR 位点中重复序列几乎完全一致,间隔序列高度可变,来源于外源噬菌体和质粒等可移动遗传因子<sup>[4]</sup>; *cas* 基因可表达多种酶功能的 CAS 蛋白,发挥 CRISPR/*Cas* 系统的作用<sup>[5]</sup>。本课题组前期对 CRISPR/*Cas* 系统在志贺菌中的分布特征<sup>[6]</sup>, CRISPR 与毒力<sup>[7]</sup>、耐药<sup>[8]</sup>的关系等进行了研究,结果显示,CRISPR 中的间隔序列对于识别不同菌株特征具有一定意义。在大肠埃希菌中常见有 4 个 CRISPR 位点,包含两种类型<sup>[9]</sup>。已有研究应用基于 CRISPR2 位点 qPCR 的方法,特异性检测大肠埃希菌 O26 : H11、O45 : H2、O103 : H2、O111 : H8、O121 : H19、O145 : H28、O157 : H7 和 O104 : H4<sup>[10-11]</sup>。本研究通过对实验室保存大肠埃希菌 CRISPR 进行检测,并与 GenBank 数据库中的大肠埃希菌 CRISPR/*Cas* 进行分析,探讨大肠埃希菌中 CRISPR/*Cas* 的分布和分子特征,及其作为大肠埃希菌分子

鉴定的价值和病原监测的意义。

## 材料与方法

1. 菌株: GenBank 数据库中 135 株全基因组测序大肠埃希菌和 203 株鸟枪法测序大肠埃希菌,本实验室大肠埃希菌 361 株 (含 38 株大肠埃希菌 O157 : H7), 共 699 株细菌。

2. 主要仪器和试剂: 大肠埃希菌 O157 胶体金检测试剂条盒购于郑州万泰生物科技有限公司; 大肠埃希菌 O157 : H7 诊断血清购于宁波天润生物药业有限公司; API20E 生化鉴定板条购于法国生物梅里埃股份有限公司; PTC-100 型基因体外扩增仪购于美国 MJRESEARCH 公司。各种固体培养基及液体培养基在本实验室按配方配制。

3. 全基因组测序大肠埃希菌 CRISPR/*Cas* 位置和结构: 在 NCBI 上查找已测序的大肠埃希菌全基因组序列, 根据大肠埃希菌的重复序列信息, 通过 BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 获得大肠埃希菌的 CRISPR, 选取前后 30 kb 作为 CRISPR/*Cas* 候选范围。根据 *cas* 基因确定 CRISPR/*Cas* 系统的类型, 绘制 CRISPR/*Cas* 系统位置及结构图。

### 4. CRISPR 的识别和同源比对:

(1) 实验室菌株: 大肠埃希菌采用煮沸法提取 DNA, 根据参考文献扩增大肠埃希菌 CRISPR (引物序列见表 1) 序列<sup>[12]</sup>, 所有引物均由北京赛百盛基因技术有限公司合成。PCR 扩增条件, 体系均参见文献<sup>[12]</sup>。PCR 扩增产物交由生工生物工程 (上海) 股份有限公司进行测序。

(2) GenBank 数据库菌株: 根据大肠埃希菌的重

表1 CRISPR引物序列

CRISPR 基因	引物	序列(5' ~ 3')	PCR 扩增产物长度(bp) <sup>a</sup>
1	F	GTTATGCGGATAATGCTACC	497 ~ 1 829
	R	CGTAYYCCGGTRGATTTGGA	
2	F	AAATCGTATGAAGTGATGCAT	579 ~ 2 318
	R	GTCGATGCAAACACATAAATA	
3	F	GCGCTGGATAAAGAGAAAAAT	836 ~ 1 786
	R	GCCCACCATTCACCTGTA	
4	F	CTGAACAGCGACTGATTTA	1 024 ~ 1 425
	R	GTACGACCTGAGCAAAG	

注: <sup>a</sup>R=G或A, Y=T或C; <sup>b</sup>PCR产物长度根据双向测序结果进行拼接

复序列信息, 通过BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 获得大肠埃希菌的CRISPR, 选取前后2 000 bp 作为CRISPR 候选范围。

(3) 通过CRISPR Finder 在线软件 (<http://crispr.u-psud.fr/Server/>) 分析获得细菌的重复序列及数目, 间隔序列及数目。DNAMAN 进行所有间隔序列的比对, 使用Clustal X 进行 *cas* 多序列比对和Mega 5.1 软件构建 *cas* 系统发育树。

5. *cas* 基因的识别以及同源树构建: 大肠埃希菌的 *cas* 基因有 I-E 型和 I-F 型两种类型, I-E 型位于 CRISPR1 的下游, I-F 型位于 CRISPR3 的下游。为了识别大肠埃希菌中的 *cas*, 选择公开可用的大肠埃希菌完整基因组的 *cas*, I-E 和 I-F 的 *cas* 基因序列分别来自于大肠埃希菌 K-12 substr. MG1655、O157 : H7 str. EDL933 和 B7A, O83 : H1 str. NRG 857C。依据 *cas* 基因序列对大肠埃希菌进行 BLAST, 获得每个菌株中存在的 *cas* 基因簇, 选取代表性的菌株构建大肠埃希菌 *cas* 基因的同源树。

结 果

1. 建立大肠埃希菌 CRISPR/Cas 定位新的描述方法: 对 135 株 GenBank 中全基因组测序大肠埃希菌研究发现, CRISPR1 位点位于锌依赖外肽酶 M28 (Zn-dependent exopeptidase M28) 和 *cas2* 之间,

CRISPR2 位点位于(糖)激酶(sugar kinase)和 7-羧基-7-去氮鸟嘌呤合成酶(7-carboxy-7-deazaguanine synthase)之间, CRISPR3 和 CRISPR4 位于 ATP 依赖 Clp 蛋白酶 ATP 结合亚单位 ClpA (ATP-dependent Clp protease ATP-binding subunit ClpA) 和翻译起始因子 IF-1 (translation initiation factor IF-1) 之间, 见表 2。

表2 135株全基因组测序大肠埃希菌CRISPR侧翼序列基因

项目	基 因	百分比(%)
CRISPR2	F: 锌依赖外肽酶 M28	93.3(97/104) <sup>a</sup>
	R: <i>cas2</i>	93.8(92/96) <sup>b</sup>
CRISPR2	F: (糖)激酶	100.0(101/101) <sup>c</sup>
	R: 7-羧基-7-去氮鸟嘌呤合成酶	100.0(101/101)
CRISPR3-4	F: ATP 依赖 Clp 蛋白酶 ATP 结合亚单位 ClpA	100.0(135/135) <sup>d</sup>
	R: 翻译起始因子 IF-1	100.0(135/135) <sup>d</sup>

注: <sup>a</sup>104株大肠埃希菌存在 CRISPR1; <sup>b</sup>96株大肠埃希菌存在 *cas*; <sup>c</sup>101株大肠埃希菌存在 CRISPR2; <sup>d</sup>135株大肠埃希菌存在 CRISPR3-4 或 CRISPR3 和 CRISPR4

I-E 型的 *cas*: 135 株全基因组测序大肠埃希菌中, ①104 株具有 CRISPR1 位点, 其中 89.4% (93/104) 具有完整的 *cas* (依次为 *cas2-cas1-casE-casD-casC-casB-casA-cas3*), 只有 1 个 spacer (图 1A)。②有 3 株的 8 个 *cas* 不完整: 有 2 株缺少 *cas3*, 1 株只有 *cas2* 和 *cas3*。③8 株完全没有 *cas*, 其下游紧邻转座酶, 其中 1 株在 CRISPR 与转座酶之间有一个插入序列 IS186(图 1B)。

I-F 型的 *cas*: 135 株全基因组测序大肠埃希菌中, ①16 株大肠埃希菌具有 CRISPR3 和 CRISPR4, 均存在 I-F 型类型的 *cas* (*cas1-cas3-csy1-csy2-csy3-csy4*) (图 2A)。②119 株全基因组测序大肠埃希菌存在 CRISPR3-4, 此类型不存在 *cas* (图 2B) 且变化较小, 后文不再描述。

2. 大肠埃希菌的 CRISPR 特点: 通过 BLAST 获得 GenBank 数据库下载大肠埃希菌的 CRISPR 和本实验室 PCR 扩增、测序获得大肠埃希菌的 CRISPR, 利用 CRISPR Finder 在线软件分析 CRISPR 获取间

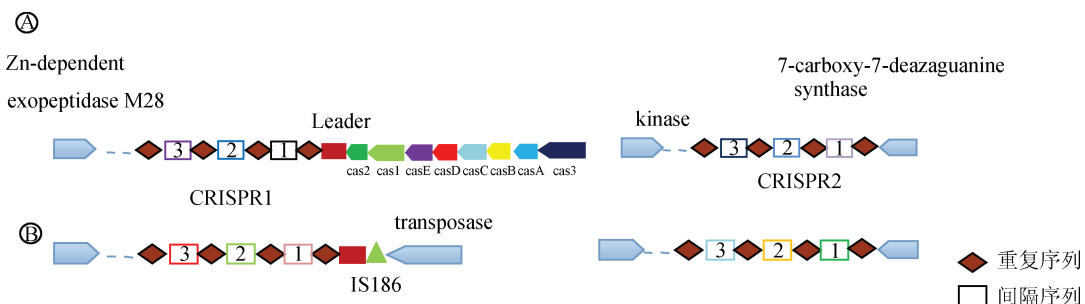


图1 I-E型的CRISPR/Cas结构

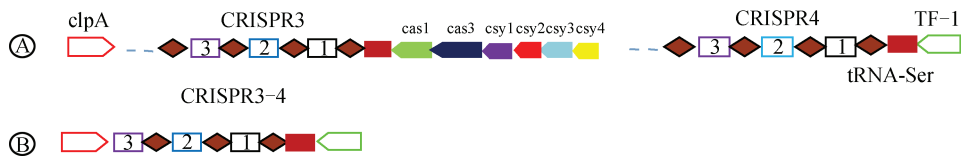


图2 I-F型的CRISPR/Cas结构

隔序列及数目,结果显示,数据库下载的1株和本实验室测序的2株大肠埃希菌中存在4个CRISPR位点。在大肠埃希菌中,由于CRISPR3和CRISPR4在大肠埃希菌存在较少,所以研究主要以CRISPR1和CRISPR2为分子靶点(表3)。

表3 699株大肠埃希菌的CRISPR分布(%)

大肠埃希菌	CRISPR1	CRISPR2	CRISPR3	CRISPR4
GenBank下载				
全基因组测序	77.04 (104/135)	74.81 (101/135)	11.85 (16/135)	11.85 (16/135)
鸟枪法测序	100.00 (203/203)	100.00 (203/203)	0 (0/203)	0 (0/203)
本实验室测序	75.62 (273/361)	92.24 (333/361)	1.39 (5/361)	1.39 (5/361)

3. 大肠埃希菌的CRISPR分布特点:通过DNAMAN对大肠埃希菌的间隔序列进行比对。

(1)O55:H7:数据库下载的338株大肠埃希菌中有10株O55:H7,其中8株具有完全一致的CRISPR1(3条spacer)和完全一致的CRISPR2(3条spacer)(图3)。另2株具有不同的CRISPR1(分别有13和7条spacer)和CRISPR2序列(3和4条spacer)。本实验室测序的361株菌中无O55:H7。

(2)O157:H7:699株菌中,包含数据库下载的153株和本实验室测序的38株O157:H7,其中数据库下载的144株和本实验室测序的36株大肠埃希菌均具有完全一致的CRISPR1(3条spacer)和完全一致的CRISPR2(1条spacer),CRISPR1的第3条spacer和CRISPR2的spacer分别与O55:H7的CRISPR1的第3条spacer和CRISPR2的第3条spacer完全一致(图3)。

另有数据库下载的9株O157:H7具有不同的CRISPR1(2~7条spacer)和不同的CRISPR2(3~5条spacer),本实验室测序的2株大肠埃希菌O157:H7未扩增出CRISPR1和具有不同的CRISPR2(4条spacer)。

(3)O157:HNM:699株菌中,数据库下载的9株O157:HNM中8株均具有完全一致的CRISPR1(4条spacer)和CRISPR2(5条spacer)(图3),CRISPR1的第4条spacer和CRISPR2的第4条spacer分别与大肠埃希菌O55:H7的CRISPR1的第3条spacer和CRISPR2的第3条spacer完全一致。另1株大肠埃希菌O157:HNM具有不同的CRISPR1(2条spacer)和CRISPR2(1条spacer)。本实验室测序的361株菌中无O157:HNM。

(4)O104:H4:数据库下载的大肠埃希菌中包含46株O104:H4,其中46株具有完全一致的CRISPR1(1条spacer)、40株具有完整的CRISPR2(13条spacer)。肠黏附性大肠埃希菌55899具有与其完全一致的CRISPR1(1条spacer),其CRISPR2有15条spacer,其11、12位置的spacer与13、14位置的spacer完全一致,缺失了重复的2条spacer,便形成O104:H4的CRISPR2的13条spacer(图4)。基因组分析显示,O104:H4有stx2a prophage的插入。另有6株大肠埃希菌O104:H4具有不同的CRISPR2(8~15个spacer不等)。本实验室测序的361株菌中无

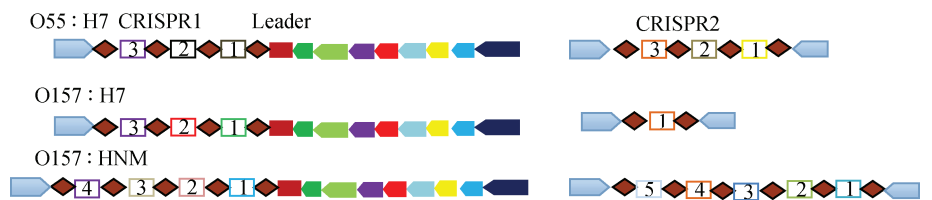


图3 大肠埃希菌O55:H7、O157:H7、O157:HNM的CRISPR特点

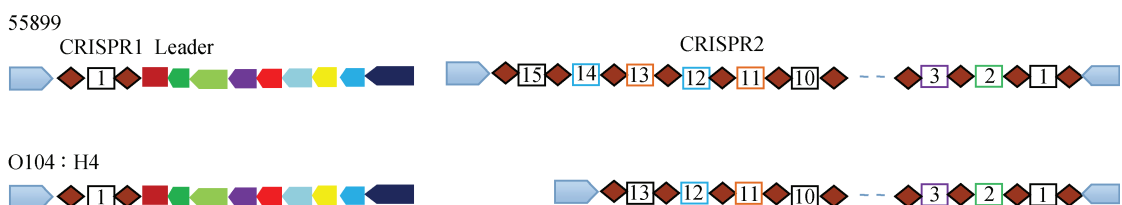


图4 肠黏附性大肠埃希菌55899和大肠埃希菌O104:H4的CRISPR特点

O104:H4。

(5)O145:H28:数据库下载的大肠埃希菌包含4株O145:H28。其中2株分离于2010年美国生菜引起暴发的菌株(RM13514, RM12581)具有完全一致的CRISPR1(6条spacer)和CRISPR2(3条spacer),2株分离于2007年比利时冰激凌引起暴发的菌株(RM13516, RM12761),具有完全一致的CRISPR1(9条spacer)和CRISPR2(2条spacer)(图5)。与比利时大肠埃希菌O145:H28 RM13516和RM12761相比,美国的大肠埃希菌O145:H28 RM13514和RM12581在CRISPR1的2、5、8位置缺失了spacer和CRISPR2的1位置新增加了spacer(图5)。spacer的缺失可发生在CRISPR的中间,发生间隔性的缺失或者相同间隔序列的缺失。基因组分析显示:美国的大肠埃希菌O145:H28和比利时的大肠埃希菌O145:H28具有不同的毒力基因类型,且菌株RM13514存在BsuB I/Pst I甲基转移酶基因。本实验室测序的361株菌中无O145:H28。

4. cas系统进化树:依据I-E和I-F的cas构建系统发育树(图6),I-E和I-F的cas均可分为两个遗传关系紧密的群组。I-E型cas系统进化树显示O104:H4与55989、O55:H7与O157:H7有更近的亲缘关

系,提示cas与CRISPR存在共进化。

### 讨 论

CRISPR/Cas广泛分布于细菌和古生菌中,1987年Ishino等<sup>[13]</sup>对大肠埃希菌的碱性磷酸酶同工酶*iap*所对应的核酸序列进行分析,发现其下游存在29 bp组成的重复序列和非重复间隔序列。2002年Jansen等<sup>[14]</sup>将其正式命名为CRISPR。本研究通过对GenBank数据库中全基因组测序大肠埃希菌135株的CRISPR位点进行检测及分析,结果显示,93.3%以上的大肠埃希菌的CRISPR1位点位于锌依赖外肽酶M28和*cas2*之间,CRISPR2位点位于(糖)激酶和7-羧基-7-去氮鸟嘌呤合成酶之间,CRISPR3和CRISPR4位于ATP依赖Clp蛋白酶ATP结合亚单位ClpA和翻译起始因子IF-1之间。这一位置的描述或许可以作为大肠埃希菌CRISPR/Cas新的定位方法。Diez-Villaseñor等<sup>[9]</sup>将CRISPR1和CRISPR2分别定位于核心基因*iap*和*cysH*、*ycgE*和*ycgF*之间,CRISPR3和CRISPR4定位于核心基因*clpA*和*infA*之间。通过对CRISPR的查找,基因*cysH*、*ycgE*、*ycgF*和*infA*并不在CRISPR的直接相邻处。因此本研究建立新的对大肠埃希菌的CRISPR/Cas位置的描述

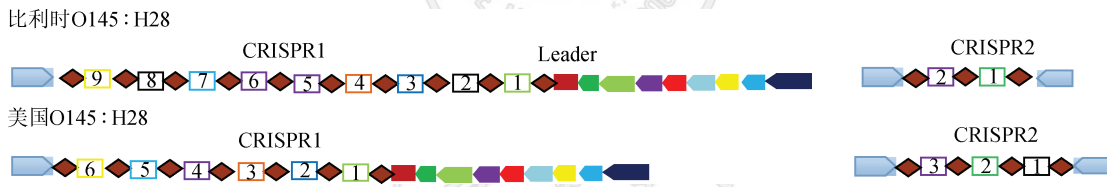


图5 比利时和美国大肠埃希菌O145:H28的CRISPR特点

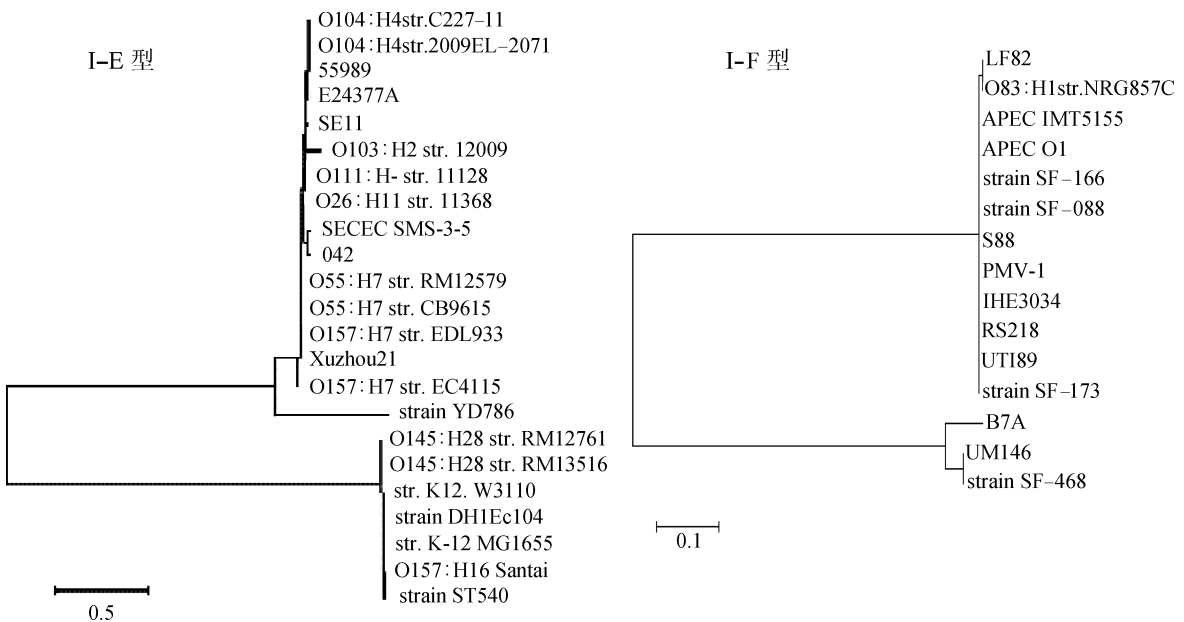


图6 大肠埃希菌cas系统进化树

方法,对CRISPR的定位更加准确和直观。

Touchon和Rocha<sup>[5]</sup>研究显示,在大肠埃希菌中没有发现超过3个CRISPR位点菌株的存在,本研究显示,大肠埃希菌B7A具有完整的I-E和I-F系统——4个CRISPR位点,实验室保存大肠埃希菌2014029、2014112也具有4个CRISPR位点,这一结果丰富了大肠埃希菌关于CRISPR/Cas的结构类型。研究显示,*cas*是CRISPR/Cas的活性体现,*cas*的存在与CRISPR中间隔序列数目有关。但是在本研究中,即使CRISPR中只有一个间隔序列仍有*cas*存在,即使存在22个间隔序列的CRISPR1,其*cas*仍有不完整的现象或者有4个间隔序列的CRISPR1的下游不存在*cas*。在1株缺失*cas*的菌株中,CRISPR1(4条间隔序列)的下游有插入序列IS186,也许*cas*的丢失和插入序列的存在有关。

CRISPR的间隔序列可以作为大肠埃希菌分子分型的靶标。已有研究显示,基于CRISPR2位点qPCR特异性检测大肠埃希菌O26:H1、O45:H2、O103:H2、O111:H8、O121:H19、O145:H28、O157:H7和O104:H4,提示CRISPR与大肠埃希菌的血清型有一定的联系<sup>[10-11]</sup>。本研究显示,CRISPR1和CRISPR2均可以特异性检测大肠埃希菌O145:H28、O157:H7、O104:H4、O55:H7和O157:HNM,且CRISPR1与*cas*直接相连,其变化或许更直接也更多样化。但是对于O26:H11、O45:H2、O103:H2、O111:H8和O121:H19等大肠埃希菌的部分血清型,其CRISPR中间隔序列的丢失和获得比较多样化,能反映它们之间的亲缘关系,但是在分子标志物方面,目前数据不支持。本研究显示,数据库下载的2株大肠埃希菌O55:H7具有不同的CRISPR1和CRISPR2序列;另有数据库下载的9株大肠埃希菌O157:H7具有不同的CRISPR1和不同的CRISPR2,本实验室测序的2株大肠埃希菌O157:H7未扩增出CRISPR1和具有不同的CRISPR2;有1株大肠埃希菌O157:HNM具有不同的CRISPR1和CRISPR2;有6株大肠埃希菌O104:H4的具有不同的CRISPR2,这种多样性的存在也许是因为菌株在进化过程中突变产生的,也许可以作为其进一步亚型分型,但因为样本数量有限,有待后续进一步研究。

CRISPR的间隔序列能反映出大肠埃希菌的亲缘关系。大肠埃希菌O55:H7属于肠致病性大肠埃希菌,全基因组测序显示出血性大肠埃希菌O157:H7与其亲缘关系最近,大肠埃希菌O55:H7和O157:H7的CRISPR1和CRISPR2都含有一条相同的间隔

序列。大肠埃希菌55899属于肠黏附性大肠埃希菌,出血性大肠埃希菌O104:H4全基因组测序与其亲缘关系最近,肠黏附性大肠埃希菌55899与大肠埃希菌O104:H4具有完全一致的CRISPR1(1条spacer),CRISPR2中除却失重复的2条间隔序列,其余间隔序列也完全相同。

CRISPR的间隔序列与噬菌体的关系。研究显示,大肠埃希菌O55:H7丢失O55抗原簇,基因组插入*stx*噬菌体,捕获大质粒pO157后,可以改变O抗原形成O157:H7血清型<sup>[2]</sup>,在此过程中丢失了4个间隔序列,新获得2个间隔序列。已有研究显示,肠黏附性大肠埃希菌55899在进化成大肠埃希菌O104:H4的过程中,有Stx2a prophage的插入<sup>[3]</sup>。Cooper等<sup>[15]</sup>研究显示美国和比利时大肠埃希菌O145:H28具有不同的毒力基因类型,且美国菌株RM13514的*stx2a*噬菌体中存在*BsuB I/Pst I*甲基转移酶基因。Makarova等<sup>[16]</sup>研究显示,间隔序列的丢失从远离*cas*的一端开始。本研究显示,间隔序列的缺失可以发生在CRISPR的中间,发生间隔性缺失或者缺失相同的间隔序列。间隔序列的缺失或许是细菌为了后期更容易获得利于自身生存的其他可移动元件。

依据I-E和I-F的*cas*构建系统发育树显示I-E和I-F的*cas*均可分为两个遗传关系紧密的群组。Wang等<sup>[17]</sup>依据志贺菌*cas1*和*cas2*构建同源树分为两类。大肠埃希菌和志贺菌同属肠杆菌科,其*cas*在进化过程中显示多样性。I-E型*cas*系统进化树显示,O104:H4与肠黏附性大肠埃希菌55899、O55:H7与O157:H7有更近的亲缘关系,提示*cas*与CRISPR存在共进化。

本研究探讨了全基因组测序的大肠埃希菌和临床分离大肠埃希菌中CRISPR/Cas系统的分布特征,结果显示,CRISPR/Cas不仅可以作为大肠埃希菌分子标志物,而且可以成为鉴定和判断强毒株大肠埃希菌或者新型菌株出现的重要分子标志物。

利益冲突 无

#### 参 考 文 献

- [1] Jain R, Rivera MC, Moore JE, et al. Horizontal gene transfer accelerates genome innovation and evolution[J]. Mol Biol Evol, 2003, 20(10): 1598-1602. DOI: 10.1093/molbev/msg154.
- [2] Feng P, Lampel KA, Karch H, et al. Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7[J]. J Infect Dis, 1998, 177(6): 1750-1753. DOI: 10.1086/517438.
- [3] Ahmed SA, Awosika J, Baldwin C, et al. Genomic comparison of *Escherichia coli* O104:H4 isolates from 2009 and 2011 reveals

- plasmid, and prophage heterogeneity, including shiga toxin encoding phage stx2[J]. PLoS One, 2012, 7(11): e48228. DOI: 10.1371/journal.pone.0048228.
- [4] Mojica FJM, Díez-Villaseño C, García-Martínez J, et al. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements[J]. J Mol Evol, 2005, 60(2): 174-182. DOI: 10.1007/s00239-004-0046-3.
- [5] Touchon M, Rocha EPC. The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella* [J]. PLoS One, 2010, 5(6): e11126. DOI: 10.1371/journal.pone.0011126.
- [6] Guo XJ, Wang YF, Duan GC, et al. Detection and analysis of CRISPRs of *Shigella* [J]. Curr Microbiol, 2015, 70(1): 85-90. DOI: 10.1007/s00284-014-0683-8.
- [7] 郭向娇, 王颖芳, 段广才, 等. 临床分离志贺菌中 CRISPR/Cas 系统的分布及其与毒力基因的关系[J]. 微生物学通报, 2015, 42(3): 543-549. DOI: 10.13344/j.microbiol.china.140529.
- Guo XJ, Wang YF, Duan GC, et al. Distribution of CRISPR/Cas system in *Shigella* clinical strains and its relationship with virulence genes [J]. Microbiol China, 2015, 42(3): 543-549. DOI: 10.13344/j.microbiol.china.140529.
- [8] 王琳琳, 王颖芳, 段广才, 等. 志贺菌 CRISPR 的检测及其与耐药的关系[J]. 微生物学报, 2015, 55(4): 476-483. DOI: 10.13343/j.cnki.wsxb.20140322.
- Wang LL, Wang YF, Duan GC, et al. Detection of CRISPR and its relationship to drug resistance in *Shigella* [J]. Acta Microbiol Sinica, 2015, 55(4): 476-483. DOI: 10.13343/j.cnki.wsxb.20140322.
- [9] Díez-Villaseñor C, Almendros C, García-Martínez J, et al. Diversity of CRISPR loci in *Escherichia coli* [J]. Microbiology, 2010, 156(Pt 5): 1351-1361. DOI: 10.1099/mic.0.036046-0.
- [10] Delannoy S, Beutin L, Fach P. Use of clustered regularly interspaced short palindromic repeat sequence polymorphisms for specific detection of enterohemorrhagic *Escherichia coli* strains of serotypes O26 : H11, O45 : H2, O103 : H2, O111 : H8, O121 : H19, O145 : H28, and O157 : H7 by real-time PCR [J]. J Clin Microbiol, 2012, 50(12): 4035-4040. DOI: 10.1128/JCM.02097-12.
- [11] Bennett PM. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria [J]. Br J Pharmacol, 2008, 153 Suppl 1: S347-357. DOI: 10.1038/sj.bjp.0707607.
- [12] Touchon M, Charpentier S, Pognard D, et al. Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements [J]. Microbiology, 2012, 158(Pt 12): 2997-3004. DOI: 10.1099/mic.0.060814-0.
- [13] Ishino Y, Shinagawa H, Makino K, et al. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product [J]. J Bacteriol, 1987, 169(12): 5429-5433.
- [14] Jansen R, van Embden JDA, Gaastra W, et al. Identification of genes that are associated with DNA repeats in prokaryotes [J]. Mol Microbiol, 2002, 43(6): 1565-1575. DOI: 10.1046/j.1365-2958.2002.02839.x.
- [15] Cooper KK, Mandrell RE, Louie JW, et al. Comparative genomics of enterohemorrhagic *Escherichia coli* O145 : H28 demonstrates a common evolutionary lineage with *Escherichia coli* O157 : H7 [J]. BMC Genomics, 2014, 15: 17. DOI: 10.1186/1471-2164-15-17.
- [16] Makarova KS, Haft DH, Barrangou R, et al. Evolution and classification of the CRISPR-Cas systems [J]. Nat Rev Microbiol, 2011, 9(6): 467-477. DOI: 10.1038/nrmicro2577.
- [17] Wang PF, Zhang B, Duan GC, et al. Bioinformatics analyses of *Shigella* CRISPR structure and spacer classification [J]. World J Microbiol Biotechnol, 2016, 32(3): 38. DOI: 10.1007/s11274-015-2002-3.

(收稿日期: 2016-03-21)

(本文编辑: 万玉立)

## 中华流行病学杂志第七届编辑委员会通讯编委名单

(按姓氏汉语拼音排序)

陈曦(湖南)	党少农(陕西)	窦丰满(四川)	高婷(北京)	高立冬(湖南)	还锡萍(江苏)	贾曼红(云南)
金连梅(北京)	荆春霞(广东)	李琦(河北)	李十月(湖北)	李秀央(浙江)	林玫(广西)	林鹏(广东)
刘莉(四川)	刘玮(北京)	刘爱忠(湖南)	马家奇(北京)	倪明健(新疆)	欧剑鸣(福建)	潘晓红(浙江)
彭晓旻(北京)	彭志行(江苏)	任泽舫(广东)	施国庆(北京)	汤奋扬(江苏)	田庆宝(河北)	王丽(北京)
王璐(北京)	王金桃(山西)	王丽敏(北京)	王志萍(山东)	武鸣(江苏)	谢娟(天津)	解恒革(海南)
严卫丽(上海)	阎丽静(北京)	么鸿雁(北京)	余运贤(浙江)	张宏伟(上海)	张茂俊(北京)	张卫东(河南)
郑莹(上海)	郑素华(北京)	周脉耕(北京)	朱益民(浙江)	祖荣强(江苏)		