



## 副溶血性弧菌 CRISPR 的检测及其结构分析

刘伟奇<sup>1</sup>, 董旭日<sup>1</sup>, 汤荣<sup>2</sup>, 陈雯静<sup>2</sup>, 洪庆<sup>2</sup>, 刘海泉<sup>1,3,4</sup>, 潘迎捷<sup>1,3,4</sup>, 赵勇<sup>1,3,4\*</sup>

<sup>1</sup>上海海洋大学食品学院, 上海 201306

<sup>2</sup>上海市第一人民医院检验科, 上海 201620

<sup>3</sup>上海水产品加工及贮藏工程技术研究中心, 上海 201306

<sup>4</sup>农业部水产品贮藏保鲜质量安全风险评估实验室, 上海 201306

**摘要:**【目的】检测副溶血性弧菌(*Vibrio parahaemolyticus*, 简称 VP)中规律成簇间隔的短回文序列(Clustered regularly interspaced short palindromic repeats, CRISPR), 并对不同来源的 VP 中 CRISPR 位点的结构多样性进行分析。【方法】根据 CRISPR DB 数据库中公布的 VP 中确定的 CRISPR 结构序列 CRISPR-1 及文献中新发现的疑似 CRISPR 结构序列 CRISPR-2 设计引物, 对不同来源的 79 株 VP 进行 PCR 扩增。利用 CRISPR Finder 分析 CRISPR 结构, 采用生物信息学方法对不同来源 VP 的 CRISPR 位点结构多样性进行比较分析。【结果】79 株 VP 中 CRISPR-1 的检出率为 92.41%, CRISPR-2 的检出率为 96.20%, 同时具有这 2 个位点的菌株占总数的 89.87%, 只有 1 株菌被检出不含有任何位点。分别比较不同来源的菌株 CRISPR-1、CRISPR-2 位点的重复序列发现不存在序列差异, 而临床菌株的这 2 个 CRISPR 位点在间隔序列上比环境分离菌株存在更多的变异。2 个 CRISPR 位点根据间隔序列的不同在 VP 中一共组成 8 种 CRISPR 谱型(编号 A-H), 除 F 谱型外, A-E、G 谱型均只在临床分离菌株中发现, 而在环境分离菌中还发现不含任何位点的 H 型。【结论】CRISPR 在 VP 中普遍存在。环境分离菌株与临床分离菌株中 CRISPR 的结构存在差异。

**关键词:** 不同来源, 副溶血性弧菌, CRISPR, 结构

规律成簇间隔的短回文重复序列(clustered regularly interspaced short palindromic repeats, CRISPR)是近年来发现的广泛分布于细菌和古细菌中的一类特殊的 DNA 重复序列<sup>[1]</sup>, 主要是由同

向重复序列(repeats, R)和间隔序列(spacers, S)间隔排列组成。研究发现<sup>[2]</sup>, CRISPR 中的同向重复序列和间隔序列在细菌进化中均存在变异现象, 使得 CRISPR 结构具有多态性, 并且在同一物种

基金项目: 国家自然科学基金(31571917, 31671779); 上海市科技兴农重点攻关项目(2015 第 4-8 号 2016 第 1-1 号); 上海市教育委员会科研创新计划; 上海市教委“曙光计划”(15SG48)

\*通信作者。Tel: +86-21-61900503; E-mail: yzhao@shou.edu.cn

收稿日期: 2016-11-19; 修回日期: 2017-01-05; 网络出版日期: 2017-02-13

的不同菌株间存在差异, 这种差异为细菌分型溯源、进化研究提供了更加可靠、准确的依据, 目前已在食源性致病菌特别是沙门氏菌的分型<sup>[3]</sup>中得到了广泛的应用; 另外 CRISPR 结构的多态性还被发现与细菌的耐药性存在一定的关系<sup>[4]</sup>。CRISPR 结构的侧翼区常存在 CRISPR 相关蛋白基因(CRISPR-associated, *cas*), 这些基因与 CRISPR 共同构成 CRISPR/*cas* 系统, 为原核生物提供对噬菌体、质粒等外源基因的获得性免疫能力, 从而抵御外源基因的入侵, 限制基因的水平转移<sup>[5-6]</sup>。研究发现 CRISPR/*cas* 系统除了免疫防御功能外, 还可能具有调控细菌生理活动等其他功能<sup>[7]</sup>。目前, 有关 CRISPR 在食源性致病菌中的功能及应用研究已成为热点<sup>[8-12]</sup>。

近年研究发现常见的食源性致病菌副溶血性弧菌(*Vibrio parahaemolyticus*, 简称 VP)中存在 CRISPR 结构, 并且已有研究表明 VP 中 CRISPR 结构与细菌进化之间并无严格关系, 并不适合用于细菌分型<sup>[11]</sup>, 而 VP 的 CRISPR 却与其毒力基因 *tdh* 之间存在显著的相关性<sup>[12]</sup>, 这可能表明含有 CRISPR 结构的 VP 存在潜在的致病性。VP 中 CRISPR/*cas* 系统如何对外源基因进行免疫防御, 是否具有其他生理功能还未可知。本文通过对 44 株临床分离菌株及 35 株环境分离菌株进行 CRISPR 结构的检测与分析, 以期对 VP 中 CRISPR 结构的功能机制研究提供基础信息。

## 1 材料和方法

### 1.1 菌株

42 株临床分离菌株由上海市第一人民医院馈赠, C5 由中国科学院微生物研究所馈赠, C39 由美国疾病与控制中心馈赠, 35 株环境分离菌株由

本实验室从河虾中分离, 具体信息见表 1。

### 1.2 CRISPR 检测

根据 CRISPR DB 数据库(<http://crispr.i2bc.paris-saclay.fr/crispr/>)中公布的 VP 中确定的 CRISPR 结构序列 CRISPR-1, 以及位于 *kdgR*、*kduD* 基因之间的疑似 CRISPR 结构序列<sup>[11]</sup>CRISPR-2 设计引物(引物序列见表 2), 其中 CRISPR-1 与 *cas* 基因组成 CRISPR/*cas* 系统, CRISPR-2 为孤立位点。以试剂盒提取的 VP 基因组 DNA 为模板, 对 VP 菌的 CRISPR 进行扩增和序列测定。扩增体系为 50  $\mu$ L, 其中  $2\times Taq$  PCR Mix 25  $\mu$ L, ddH<sub>2</sub>O 20  $\mu$ L, 10 mmol/L 上下游引物各 2  $\mu$ L, DNA 模板 1  $\mu$ L。PCR 反应条件为: 94  $^{\circ}$ C 5 min; 94  $^{\circ}$ C 30 s, 53  $^{\circ}$ C 60 s, 72  $^{\circ}$ C 60 s, 35 个循环; 72  $^{\circ}$ C 10 min。取 5  $\mu$ L PCR 产物于 2% 的琼脂糖凝胶电泳进行初步分析, 所得阳性结果送至生工生物工程(上海)股份有限公司进行测序, 利用 CRISPR Finder<sup>[13]</sup>, 确认 CRISPR 的存在和结构, 利用序列比对软件 ClustalX 2.1 及 DNAMAN 对检出重复序列和间隔序列进行比对。利用在线软件 WebLogo (<http://weblogo.berkeley.edu/logo.cgi>)对重复序列进行保守性分析, 所有间隔序列都通过 BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>)在线比对分析比对, 以寻找与其同源的 DNA 序列。

## 2 结果和分析

### 2.1 79 株 VP 的 CRISPR 位点分析

**2.1.1 CRISPR 的检测结果:** 以 79 株 VP 基因组 DNA 为模板, PCR 扩增 CRISPR-1、CRISPR-2 (图 1), 采用 CRISPR Finder 分析查找 CRISPR, 结果见表 3, 在 79 株 VP 中大部分菌株含有 2 个 CRISPR 位点, 只有 1.27% 的菌株未检出任何 CRISPR 位

表 1. 实验中所用菌株

Table 1. The strains used in this study

Isolate	Genotype	Source of isolation	Isolate	Genotype	Source of isolation
C1	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	C97	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical
C2	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	C100	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical
C15	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	C5	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical
C16	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	C39	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical
C17	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S1	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C18	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S2	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C19	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S4	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C20	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S6	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C21	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S8	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C22	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S9	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C25	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S10	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C26	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S11	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C27	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S12	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C28	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S13	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C29	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S14	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C32	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S15	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C33	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S16	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C34	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S17	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C35	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S18	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C36	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S19	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C37	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S20	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C38	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S21	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C40	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S22	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C41	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S23	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C42	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S24	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C43	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S25	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C44	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S26	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C45	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S27	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C46	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S28	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C47	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S29	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C48	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S30	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C49	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S31	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C50	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S32	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C51	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S33	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C54	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S34	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C55	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S35	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C85	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S36	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C89	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S37	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C90	<i>tdh</i> <sup>+</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Clinical	S38	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>-</sup> / <i>tlh</i> <sup>+</sup>	Shrimp
C94	<i>tdh</i> <sup>-</sup> / <i>trh</i> <sup>+</sup> / <i>tlh</i> <sup>+</sup>	Clinical			

表 2. CRISPR 的引物序列

Table 2. Primers of CRISPR for PCR amplification

Genes	Primers	Sequences (5'→3')	PCR products/bp
CRISPR-1	F	AAGAAGATAGCAGCAGTGGCAATAA	556
	R	TTGGCTGAGCCTGGAGATGTAA	
CRISPR-2	F	AACGCCGATGGCTTTCTGTT	486
	R	TGTTGGTTCGCTGATATTTACGC	

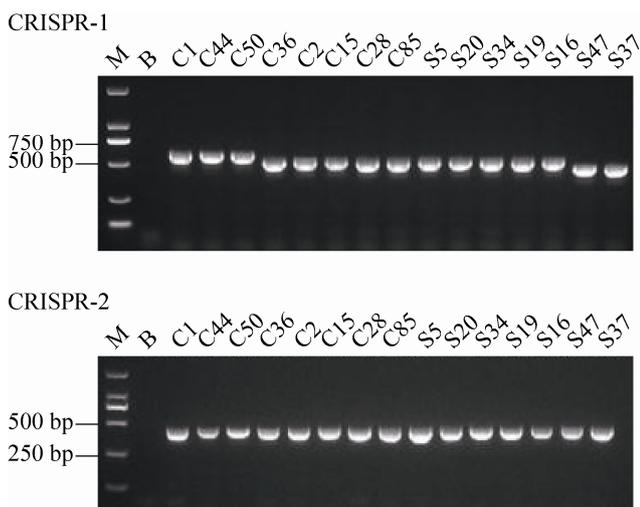


图 1. 2 个 CRISPR 位点 PCR 扩增产物的电泳图  
Figure 1. Gel electrophoresis of the two CRISPR locus. M: DL2000 DNA marker; B: blank; C1-S37: different types of strains.

表 3. 79 株 VP CRISPR 的分布  
Table 3. Distribution of CRISPR in 79 VP strains

Genes	Strain number	Percentage/%
CRISPR-1+CRISPR-2	71	89.87
CRISPR-1	2	2.53
CRISPR-2	5	6.33

表 4. 79 株 VP CRISPR 的检出率和间隔序列数目

Table 4. The positive rate and spacer number of CRISPR in 79 VP strains

Genes	Spacer number	Clinical strain number	Environmental strain number	Percentage/%
CRISPR-1	2	33	31	92.41
	3	9	0	
CRISPR-2	1	43	33	96.20

点, 单独含有 CRISPR-1、CRISPR-2 位点的菌株分别占总数的 2.53%、6.33%。在 79 株 VP 中确定的 CRISPR 结构 CRISPR-1 的总检出率较高(表 4), 其中含 2 个间隔序列的 CRISPR 位点检出率为 81.01%, 11.39% 的菌株含有 3 个间隔序列, 通过比较分析临床分离菌株与环境分离菌株 CRISPR-1 的检出率发现, 临床分离菌株 CRISPR-1 的检出率高于环境分离菌株, 并且只在临床分离菌株中发现含有 3 个间隔序列。疑似 CRISPR 结构 CRISPR-2 在 79 株 VP 中普遍存在, 并且都只含有 1 个间隔序列, 而在这个疑似位点的周围没有 *cas* 基因的存在<sup>[11]</sup>, 推测可能是一个退化的 CRISPR 位点。

**2.1.2 CRISPR 的重复序列:** CRISPR Finder 识别到的重复序列共 2 种, 并未发现有新的重复序列存在。CRISPR-1 的重复序列与 CRISPR DB 数据库中公布的一致, 长度均为 28 bp, 通过 WebLogo 在线软件分析所有检出 CRISPR-1 的重复序列的保守性(图 2-A)发现, 在重复序列的末端位点存在较多的变异。疑似位点 CRISPR-2 的重复序列长度均为 25 bp, 通过序列保守性分析(图 2-B)发现, 该位点的重复序列较保守, 仅有 1 个位点发生变异。

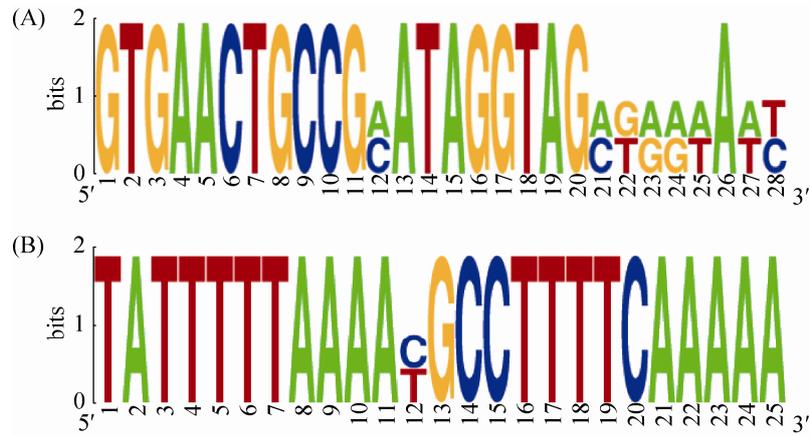


图 2. VP 所有 CRISPR 位点的重复序列

Figure 2. Sequence logo for all repeats in the CRISPR loci of VP. The height of the letters showed the relative frequency of the corresponding nucleotide at that position.

**2.1.3 CRISPR 的间隔序列:** CRISPR Finder 识别到的间隔序列共 231 条, 其中确定位点 CRISPR-1 中检测到 3 种不同类型的间隔序列, 共 155 条, 有 2 种类型的间隔序列与 CRISPR DB 数据库中公布的一致, 另一种间隔序列为新发现的间隔序列 spacer1, 共 10 条。疑似位点 CRISPR-2 都只含有

1 种间隔序列。在同一 CRISPR 位点中间隔序列存在差异(图 3), CRISPR-1 中 3 种类型的间隔序列之间同源性较低, spacer1、spacer2、spacer3 这 3 种类型的间隔序列在一些位点存在碱基的突变其中 spacer2、spacer3 这两种类型间隔序列的多个碱基位点发生不同的突变。CRISPR-2 的间隔序列,

Spacer1.1[1]	TTGTTGGCTCAGGGAGTTTTACCCCTGACAT	31
Spacer1.2[9]	TTGTTGGCTCAGT GAGTTTTACCCCTGACAT	31
Consensus	t t g t t g g c t c a g g a g t t t t a c c c c t g a c a t	
Spacer2.1[1]	GAGATACCACATGCTCAAGCAGATGCTAACA	31
Spacer2.2[16]	GAGATACCACAGGCTCAAGCAGATGCTAACA	31
Spacer2.3[55]	GAGATACCACAAGCTCAAGCAGATGCTAACA	31
Spacer2.4[1]	GAGATACCACAAGCTCAAGCACATGCGAACA	31
Consensus	g a g a t a c c a c a g c t c a a g c a a t g c t a a c a	
Spacer3.1[1]	AGTCGGTCAACTGATTAATACGTTGTTGCCA	30
Spacer3.2[70]	AGTCGGTCAACTGAGAATACGTTGTTGCCA	30
Spacer3.3[1]	AGTCGGTCAACTGAGAATACATTGTTGCCA	30
Consensus	a g t c g g t c a a c t g a a a t a c t t g t t g c c a	
Spacer4.1[2]	TTTGAAACCGCGCAATTAAGAATCCAGAAAAATAAGATCGCAATCA	47
Spacer4.2[69]	TTTGAAACCGCGCAATTAAGAATCCAGAAAAACAAGATCGCAATCA	47
Spacer4.3[1]	TTTGAAACCGCGCAATTAAGAATCCAGAAAAACAAGATCGCAATCA	47
Spacer4.4[2]	TTTGAAACCGCGCAATTAAGAATCCAGAAAAACTAGATCGCAATCA	47
Spacer4.5[1]	TTTGAAACCGCGCAATTAAGAATCCAGAAAAACTAGATCGCAATCA	47
Spacer4.6[1]	TTTGAAACCGCGCAATTAAGAATCCAGAAAAACAAGATCGCAATCA	47
Consensus	t t t g a a a c c g c g c a a t t a a g a a t c c a a a a a a g a t c g c a a t c a	

图 3. CRISPR 的间隔序列比对

Figure 3. Sequence alignment of CRISPR spacers. Messages in bracket represented the number of strains with corresponding spacer. Blue and pink respectively represented 100%, 75% homology.

spacer4 多样性较高, 主要集中在 3 个位点发生碱基的突变, 共 6 种亚型。将所有检出的间隔序列在 NCBI 数据库中进行 BLAST 分析, 并未发现与之高度同源的质粒或噬菌体序列, 但新获得的间隔序列 spacer1.2 与数据库中 VP 自身基因组上的假定蛋白的基因序列高度同源, 这表明 spacer1.2 靶向了自身基因序列。

## 2.2 不同来源的 VP 间 CRISPR 位点的比较分析

### 2.2.1 环境菌株与临床菌株 CRISPR 位点的比较分析: 79 株 VP 中, 除 S37 外, 每株至少含有 1 个

CRISPR 位点, 根据 CRISPR-1 位点间隔序列的多样性, 共组成了 8 种不同的 CRISPR 谱型, 结合 CRISPR-2 位点间隔序列的多样性, 8 种谱型又包括不同的亚型, 这些位点在 79 株 VP 中的分布情况见图 4。通过 CRISPR 位点分布图发现, 大部分 VP 菌株主要分布在 F-5 亚型, 比较临床分离菌株与环境分离菌株发现, 环境分离菌株主要分布在 F 谱型, 只有菌株 S37 为 H 谱型, 而在剩余其他谱型中未发现有环境分离菌株, CRISPR-2 位点的多样性赋予了环境分离菌株 CRISPR 位点的多样性,

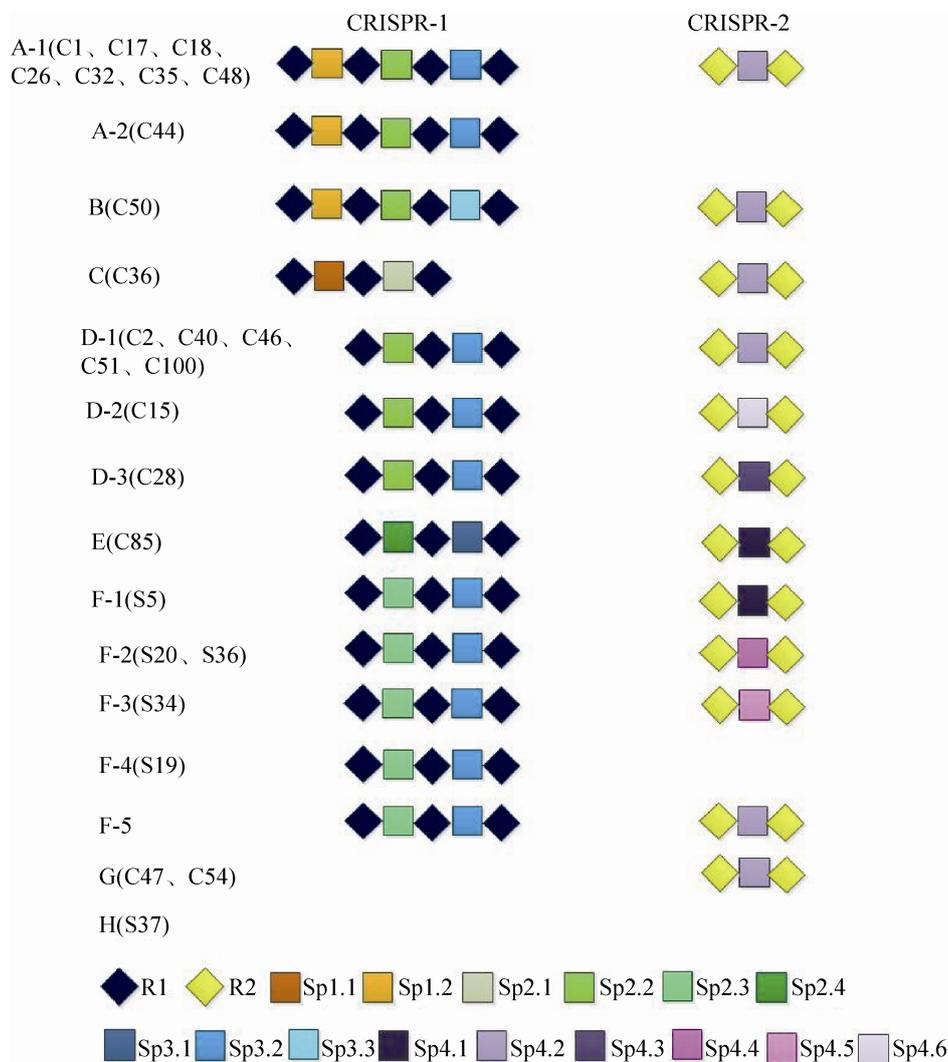


图 4. 79 株 VP CRISPR 的结构分布

Figure 4. The CRISPR structures in the 79 VP strains A→H represented the 8 CRISPR spectral patterns. Messages in bracket represented the strains with corresponding CRISPR spectral pattern.

环境分离菌株的变化影响比较单一。而临床分离菌株分布较分散,除 H 谱型外,其他谱型中均存在临床分离菌株,在临床分离菌株中 CRISPR-1、CRISPR-2 位点的间隔序列均发生不同程度的碱基突变,多个位点的碱基突变赋予了临床分离菌株较环境分离菌株更多的多样性。在临床分离菌株 C44、C47、C54 及环境分离菌株 S19、S37 中均有 CRISPR 位点缺失的发生,而在临床菌株 C36 中发现该菌的 CRISPR-1 位点缺失较古老的间隔序列,新获得的间隔序列 spacer1 及间隔序列 spacer2 也较 A、B 型存在碱基的变化。这些都表明临床分离菌株与环境分离菌株 CRISPR 位点存在差异,也证明了在不同的生存环境下 VP 的 CRISPR 发生了微进化。

**2.2.2 79 株 VP 与数据库中全基因组测序的 VP 的 CRISPR 位点的比较分析:**通过 CRISPR Finder 查找 NCBI 数据库中 12 株全基因组测序的 VP (RIMD 2210633、BB22OP、FDA\_R31、CDC\_K4557、UCM-V493、FORC\_008、FORC\_006、FORC\_004、ATCC17802、FORC\_014、CHN25、FORC\_023) 中的 CRISPR 位点,由于 CRISPR 为非编码序列,由 CRISPR Finder 检测到的位于编码区的重复间隔序列不在本研究范围内。在 12 株菌中, FDA\_R31 未检测到任何 CRISPR 序列, RIMD 2210633 为上述 F-5 亚型, BB22OP 为 D-1 亚型,其余 9 株 VP 均属于 G 型,在这 9 株菌中新发现了 3 种不同类型的疑似位点 CRISPR-2。通过检测全基因组测序的 VP 发现, CRISPR 位点只出现在 VP 的第 2 条染色体中,并未发现菌株的质粒及染色体 1 上具有 CRISPR 位点,已有研究推测 VP 的 CRISPR 结构可能是以水平基因转移的方式整合到细菌染色体上<sup>[11]</sup>,而为何 CRISPR 位点只整合到染色体 2 中,这其中的获得机制还有待进一步研究。由于 12 株菌中检测到的确

定位点 CRISPR-1 很少,而 CRISPR-1 的多样性主要出现在临床分离菌株中,通过比对 NCBI 数据库中全基因组测序的临床分离菌株<sup>[11]</sup>并未发现新的重复序列及间隔序列,VP 的 CRISPR-1 结构在不同年代分离的菌株中相近,这表明 VP 的 CRISPR 进化较缓慢,侧面说明了 VP 的 CRISPR/cas 系统活力较低。

### 3 讨论

CRISPR/cas 是大多数细菌和古细菌内一种防御异源片段入侵的免疫系统,而 CRISPR 结构中的间隔序列在免疫过程中发挥着重要作用,间隔序列可以靶向外源基因,CRISPR/cas 系统通过获得外源基因的一段标志基因为新的间隔序列,从而抵御这一外源基因的再次入侵<sup>[5-6]</sup>。而本研究中只有 CRISPR-1 位点周围存在 cas 基因,具有免疫防御功能,在检出的 CRISPR-1 位点中 1 株菌最多只含有 3 个间隔序列,大部分菌株只有 2 个间隔序列,并且这些间隔序列的序列相近,而这些间隔序列也并未发现与任何质粒或噬菌体基因同源,而通过本实验室之前对菌株耐药性分析发现,临床分离菌株都具有多重耐药基因<sup>[14]</sup>,这些现象似乎与 CRISPR/cas 系统的功能不符,一个高活性的 CRISPR/cas 系统不仅具有多样性较多的间隔序列,并且应有效地抵御外源基因的入侵。近期已有研究表明,在 VP 中发现由移动元件表达的蛋白 AcrF9 是一类 anti-CRISPR 蛋白,并且具有很高的活性,可以有效地抑制 I-F 型 CRISPR 系统的活性,不仅是抑制 VP 自身 CRISPR 系统(I-F 型<sup>[12]</sup>)活性,而且能够抑制其他 I-F 型细菌的 CRISPR 系统活性,尽管这些菌株的 I-F 型 CRISPR 系统的 Cas 蛋白的同源性较低,但抑制活性仍然很高<sup>[15]</sup>。这或许是 VP 中水平基因转移频繁,CRISPR 系统活性低的原因。

在比对 VP 间隔序列同源性时, 我们发现 spacer1.2 靶向 VP 基因组蛋白, 其中的功能还未可知, 这种现象同样在其他菌株中也有所发现, 而在一些菌中 CRISPR 的间隔序列靶向一些重要的蛋白家族, 这些蛋白家族在菌株的生理调节活动中发挥一定的作用<sup>[16]</sup>, 这或许可以推测 VP 的 CRISPR 也很可能参与细菌生存方式、毒力等多项生理活动调节。在本文研究中临床分离菌株与环境分离菌株的 CRISPR 存在一定的差异, 临床分离菌株相对于环境分离菌株在间隔序列上存在较多的变异, 甚至有一些菌株的 CRISPR 位点缺失, 已有研究<sup>[17]</sup>表明毒力或抗生素抗性等的选择压力很可能影响 CRISPR/cas 系统的存在, 在不同生存环境的选择压力下, 为何临床分离菌株 CRISPR 表现出较多变异性, 这其中的机制还有待进一步的研究。

## 参考文献

- [1] Karginov FV, Hannon GJ. The CRISPR system: small RNA-guided defense in bacteria and archaea. *Molecular Cell*, 2010, 37(1): 7–19.
- [2] Wang J, Qiu SF, Song HB, Sun YS. Progress in CRISPR research on bacteria typing and evolution. *Letters in Biotechnology*, 2013, 24(3): 414–417. (in Chinese)  
王建, 邱少富, 宋宏彬, 孙岩松. CRISPR 在细菌分型和进化中的研究进展. *生物技术通讯*, 2013, 24(3): 414–417.
- [3] Cai YQ, Li QC, Tao J, Jiang DJ, Pan ZM, Jiao XA. Advances in structure and function of CRISPR in *Salmonella*. *Progress in Veterinary Medicine*, 2015, 36(2): 97–101. (in Chinese)  
蔡银强, 李求春, 陶静, 蒋道军, 潘志明, 焦新安. 沙门菌 CRISPR 的结构与功能研究进展. *动物医学进展*, 2015, 36(2): 97–101.
- [4] Wang LL, Wang YF, Duan GC, Xue ZR, Guo XJ, Wang PF, Xi YL, Yang HY. Detection of CRISPR and its relationship to drug resistance in *Shigella*. *Acta Microbiologica Sinica*, 2015, 55(4): 476–483. (in Chinese)  
王琳琳, 王颖芳, 段广才, 薛泽润, 郭向娇, 王鹏飞, 郝园林, 杨海燕. 志贺菌 CRISPR 的检测及其与耐药的关系. *微生物学报*, 2015, 55(4): 476–483.
- [5] Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, 2008, 322(5909): 1843–1845.
- [6] Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, van der Oost J, Koonin EV. Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, 2011, 9(6): 467–477.
- [7] Medina-Aparicio L, Rebollar-Flores JE, Gallego-Hernández AL, Vázquez A, Olvera L, Gutiérrez-Ríos RM, Calva E, Hernández-Lucas I. The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *Journal of Bacteriology*, 2011, 193(10): 2396–2407.
- [8] Liu WQ, Wang X, Liu HQ, Pan YJ, Zhao Y. Advances in structures and functions of CRISPR in common foodborne pathogenic bacterium. *Food Science*, <http://www.cnki.net/kcms/detail/11.2206.TS.20160607.1718.132.html>. (in Chinese)  
刘伟奇, 王旭, 刘海泉, 潘迎捷, 赵勇. 常见食源性致病菌 CRISPR 系统结构与功能研究进展. *食品科学*, <http://www.cnki.net/kcms/detail/11.2206.TS.20160607.1718.132.html>.
- [9] Fricke WF, Mammel MK, McDermott PF, Tartera C, White DG, Leclerc JE, Ravel J, Cebula TA. Comparative genomics of 28 *Salmonella enterica* isolates: evidence for CRISPR-mediated adaptive sublineage evolution. *Journal of Bacteriology*, 2011, 193(14): 3556–3568.
- [10] Di HL, Yan H, Shi L. Analysis of CRISPR regions in food-borne *Listeria monocytogenes*. *Modern Food Science and Technology*, 2014, 30(8): 64–69, 237. (in Chinese)  
狄慧玲, 闫鹤, 石磊. 食源性单核细胞增生李斯特菌 CRISPR 结构的研究. *现代食品科技*, 2014, 30(8): 64–69, 237.
- [11] Ge CY, Lu HX, Qu DF, Zhu JL, Yuan DX, Feng LF. Comparative analysis of clustered regularly interspaced short palindromic repeats (CRISPR) in the genomes of *Vibrio parahaemolyticus* group. *Genomics and Applied Biology*, 2015, 34(8): 1712–1722. (in Chinese)  
葛彩云, 陆海霞, 曲道峰, 朱军莉, 袁冬霞, 冯立芳. 副溶血性弧菌群中规律成簇间隔短回文重复序列的比较分析. *基因组学与应用生物学*, 2015, 34(8): 1712–1722.
- [12] Sun HH, Li YH, Shi XL, Lin YM, Qiu YQ, Zhang JJ, Liu Y, Jiang M, Zhang Z, Chen QC, Sun Q, Hu QH. Association of CRISPR/Cas evolution with *Vibrio parahaemolyticus* virulence factors and genotypes. *Foodborne Pathogens and Disease*, 2015, 12(1): 68–73.
- [13] Grissa I, Vergnaud G, Pourcel C. CRISPRfinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, 2007, 35(Suppl 2): W52–W57.
- [14] Li H, Zhang ZH, Tang R, Lou Y, Zhao L, Chen WJ, Hong Q, Pan YJ, Zhao Y. Comparison of antimicrobial resistance of pathogenic *Vibrio parahaemolyticus* in seafood and clinical samples. *Chinese Journal of Zoonoses*, 2016, 32(11): 1006–1012. (in Chinese)  
李欢, 张昭寰, 汤荣, 娄阳, 赵莉, 陈雯静, 洪庆, 潘迎捷, 赵勇. 食品与临床分离的致病性副溶血性弧菌耐药性比较. *中国人兽共患病学报*, 2016, 32(11): 1006–1012.
- [15] Pawluk A, Staals RHJ, Taylor C, Watson BNJ, Saha S, Fineran PC, Maxwell KL, Davidson AR. Inactivation of CRISPR–Cas systems by anti-CRISPR proteins in diverse bacterial species. *Nature Microbiology*, 2016, 1(8): 16085.
- [16] Pleckaityte M, Zilnyte M, Zvirbliene A. Insights into the CRISPR/Cas system of *Gardnerella vaginalis*. *BMC Microbiology*, 2012, 12(1): 301.
- [17] Palmer KL, Gilmore MS. Multidrug-resistant enterococci lack CRISPR–cas. *mBio*, 2010, 1(4): e00227-10.

## Detection and analysis of CRISPR in *Vibrio parahaemolyticus*

Weiqi Liu<sup>1</sup>, Xuri Dong<sup>1</sup>, Rong Tang<sup>2</sup>, Wenjing Chen<sup>2</sup>, Qing Hong<sup>2</sup>, Haiquan Liu<sup>1,3,4</sup>, Yingjie Pan<sup>1,3,4</sup>, Yong Zhao<sup>1,3,4\*</sup>

<sup>1</sup> College of Food Science and Technology, Shanghai Ocean University, Shanghai 201306, China

<sup>2</sup> Shanghai General Hospital, Shanghai 201620, China

<sup>3</sup> Shanghai Engineering Research Centre of Aquatic-Product Processing & Preservation, Shanghai 201306, China

<sup>4</sup> Laboratory of Quality & Safety Risk Assessment for Aquatic Product on Storage and Preservation (Shanghai), Ministry of Agriculture, Shanghai 201306, China

**Abstract:** [Objective] To detect the clustered regularly interspaced short palindromic repeats (CRISPR) in *Vibrio parahaemolyticus*, and analyze its structural diversities in different sources of *V. parahaemolyticus* strains. [Methods] The primers of convincing CRISPR structure CRISPR-1 were designed by using the sequence according to CRISPR database, while the primers of questionable CRISPR structure CRISPR-2 were designed according to the literature. PCR approach was used to detect the CRISPR locus in all 79 strains, and all CRISPR sequences were analyzed using CRISPR Finder. Furthermore, the structural diversities of CRISPR in different sources of *V. parahaemolyticus* were analyzed using bioinformatics methods. [Results] The positive rate of CRISPR-1 was 92.41% and of CRISPR-2 96.20%. The strains possessed with these two locus accounted for 89.87% of the total strains, and only one strain did not contain any locus. There was no difference between the repeats of CRISPR-1 and CRISPR-2 from different sources of *V. parahaemolyticus*, whereas the spacers of those two locus in clinical strains had more variations than those in environmental strains. Therefore, two CRISPR locus formed 8 spectral patterns (numbers A-H) according to the variations of spacers. Except type F, the patterns A-E, G only were found in the clinical strains. And the type H containing none of locus was only found in environmental strains. [Conclusion] CRISPR commonly existed in *V. parahaemolyticus*. There were differences in the structure of CRISPR between environmental and clinical strains.

**Keywords:** different sources, *Vibrio parahaemolyticus*, CRISPR, structure

(本文责编: 李磊)

Supported by the National Natural Science Foundation of China (31571917, 31671779), by the Key Project of Shanghai Agriculture Prosperity through Science and Technology [2015(4-8), 2016(1-1)], by the Innovation Program of Shanghai Municipal Education Commission and by the "Dawn" Program of Shanghai Education Commission (15SG48)

\*Corresponding author. Tel: +86-21-61900503; E-mail: yzhao@shou.edu.cn

Received: 19 November 2016; Revised: 5 January 2017; Published online: 13 February 2017



赵勇, 博士, 教授, 博士生导师, 上海海洋大学食品学院副院长。中国食品科学技术学会青年工作委员会委员、非热加工技术分会理事, 上海市食品学会青年工作委员会主任委员、食品安全专委会副主任委员, 上海市微生物学会微生物耐药防控专委会副主任委员。2005年于南京农业大学获得微生物学理学博士学位, 2006年复旦大学生命科学学院访问学者, 2008年美国佐治亚大学食品科学与技术系访问学者, 2009年英国食品研究院访问学者。近年主持科研项目共19项, 科研总经费达千万以上; 发表科研论文205篇, 其中SCI论文53篇; 获得各类科技奖项10余项。目前主要研究工作为: 食源性致病菌快速检测技术、风险评估技术及高效防控技术的研究。