

蜡状芽孢杆菌群中规律成簇间隔短回文重复序列的生物信息学分析

王 琰¹ 喻 婵¹ 王阶平¹ 邱 宁¹ 何 进¹ 孙 明¹ 张青叶^{1,2,*}

(1 农业微生物学国家重点实验室 微生物农药国家工程研究中心 华中农业大学生命科学技术学院 武汉 430070)

(2 华中农业大学理学院 武汉 430070)

摘要 规律成簇间隔短回文重复序列 (clustered regularly interspaced short palindromic repeats, CRISPR) 是最近发现针对噬菌体等外源遗传物质的获得性和可遗传性的新型原核生物防御系统。通过 BLAST、多序列比对、RNA 二级结构预测等生物信息学方法对已经完成全基因组测序的蜡状芽孢杆菌群 24 个菌株进行 CRISPR 的系统分析, 结果表明: 42% 的菌株含有该结构; 8 个 CRISPR 座位的正向重复序列可以形成 RNA 二级结构, 提示正向重复序列可能介导外源 DNA 或 RNA 与 CAS 编码蛋白的相互作用; 31% 的间区序列与噬菌体、质粒、蜡状芽孢杆菌群基因组序列具有同源性, 进一步验证间区序列很可能来源于外源可移动遗传因子。由于大部分蜡状芽孢杆菌群菌株含有多个前噬菌体和质粒, 通过对蜡状芽孢杆菌群 CRISPR 的分析, 为揭示其对宿主菌与噬菌体, 以及宿主菌与质粒间的关系奠定基础。

关键词 蜡状芽孢杆菌群 规律成簇间隔短回文重复序列 前导序列 重复序列 间区序列
中图分类号 Q52

规律成簇间隔短回文重复序列 (clustered regularly interspaced short palindromic repeats, CRISPR) 由一段不连续的正向重复序列 (repeats, R) 和插入其中的非重复间区序列 (spacers, S) 组成。它与 CRISPR 相关蛋白 (Cas 蛋白) 组成的免疫系统可以帮助原核生物抵御外源遗传物质 (如噬菌体和质粒) 的入侵^[1]。CRISPR/Cas 系统依赖于转录的小 RNAs, 监测并消除入侵的外源遗传物质, 保护细胞基因组的完整性, 是一种针对特定外源遗传物质入侵的获得性和可遗传性的新型微生物防御系统, 是理解宿主和病毒间相互作用的一个重要新发现^[2]。

重复序列 (R) 一般由 28 ~ 40 个碱基组成, 在一个 CRISPR 中几乎完全一致^[3]。间区序列 (S) 是 CRISPR 中高度可变的组分, 在同一个 CRISPR 中几乎找不到两个相同的间区序列^[4]。普遍认为间区序列来源于外源可移动遗传因子, 其中一部分来自噬菌体或质粒等染色体外序列^[5], 是外源遗传物质入侵的免疫“记录器”。

CRISPR 最为显著的特征是重复序列可以转录, 且形成 RNA 二级结构。在 CRISPR 的 5' 端通常含有一段富含 AT 的前导序列 (leader, L)^[6]。存在于非编码区内, 在相同的原核物种中, 前导序列具有高度的相似性^[6], 研究证明在 CRISPR 的前导序列中包含它的转录启动子^[7]。cas 基因只存在于含有 CRISPR 的基因组中, 一般位于 CRISPR 座位 (locus) 附近^[6], Haft 等将其分为 45 个 cas 基因家族, 根据 Cas 蛋白的序列同源性、组成和功能不同, 进一步分为 8 个亚型^[8], Cas 蛋白具有核酸酶、解旋酶、整合酶和聚合酶活性^[2]。

在已完成基因组测序的原核生物中, 约 40% 的细菌和 90% 的古菌基因组中都发现了 CRISPR 结构^[6], 对它的研究主要集中在嗜热链球菌 (*Streptococcus thermophilus*) 和一些古菌 (Archaeal) 中, 而在芽孢杆菌中对 CRISPR 的研究还很少。蜡状芽孢杆菌群 (*Bacillus cereus* group) 由进化关系相近且与人类活动密切相关的 6 种芽孢杆菌组成, 包括苏云金芽孢杆菌 (*B. thuringiensis*, BT), 炭疽芽孢杆菌 (*B. anthracis*, BA), 蜡状芽孢杆菌 (*B. cereus*, BC), 蕈状芽孢杆菌

收稿日期: 2011-04-02 修回日期: 2011-05-10

* 国家自然科学基金资助项目 (30930004)

** 通讯作者, 电子信箱: zqy@mail.hzau.edu.cn

(*B. mycoides*, BM), 假真菌样芽孢杆菌(*B. pseudomycoides*, BP)和韦氏芽孢杆菌(*B. weihenstephanensis*, BW), 是当前研究的热点。已有蜡状芽孢杆菌群 20 个菌株的全基因组序列, 本课题组完成了苏云金芽孢杆菌 4 个菌株 YBT-1520(待发表)、CT-43(待发表)、YBT-020^[9]和 BMB171^[10]的全基因组测序。由于大部分蜡状芽孢杆菌菌株含有多个前噬菌体和质粒, 本文通过对蜡状芽孢杆菌群 CRISPR 的分析, 为揭示其对宿主菌与噬菌体, 以及宿主菌与质粒间的关系奠定基础。

1 材料与方法

1.1 实验数据

苏云金芽孢杆菌 YBT-1520、CT-43、YBT-020 和 BMB171 的全基因组测序由喻子牛课题组完成, 蜡状芽孢杆菌群 20 个菌株全基因组序列从 GenBank (<http://www.ncbi.nlm.nih.gov/>) 下载。

1.2 实验方法

苏云金芽孢杆菌 YBT-1520、CT-43、YBT-020 和 BMB171 的 CRISPR 信息通过 CRISPRFinder 软件^[11]分析获得。蜡状芽孢杆菌群 20 个菌株的 CRISPR 信息从 CRISPR database 数据库^[12]下载获得。使用 Glimmer 软件^[13]对菌株 YBT-1520、CT-43、YBT-020 和 BMB171 的全基因组进行 ORF 预测, 所得结果与其它 20 株菌的全部基因共同通过 BLAST 与 45 个 *cas* 基因家族序列比对, 寻找 *cas* 基因(*e* 值为 $1e-10$)。间隔序列通过 NCBI BLASTN 比对, 寻找外源同源序列。前导序列通过 Clustal X 软件^[14]进行多序列比对获得。RNA 二级结构预测通过 RNAfold 软件^[15]获得。由于 CRISPR 通常位于非编码区内, 通过 CRISPR database 数据库和 CRISPRFinder 软件得到的所有结果应去除存在于编码区内的结果^[12]。对于含有较少重复序列和缺少 *cas* 基因的 CRISPR 座位则定义为疑似座位。

2 结果与分析

2.1 蜡状芽孢杆菌群 24 个菌株的 CRISPR

蜡状芽孢杆菌群 24 个菌株中共有 10 个菌株含有 CRISPR 结构, 所占比例为 42%, 与报道的比例相一致^[6]。这 10 株菌中共有 19 个 CRISPR 座位(包括疑似座位), 这些座位的详细信息见表 1。蜡状芽孢杆菌群中各 CRISPR 座位的重复序列数都较少(<12 个), 而

古菌相比含有较多的重复序列数(>20 个)。

BC 菌株 G9842 有 1 个 CRISPR 座位, subsp. *cytotaxis* NVH 391-98 有 3 个 CRISPR 座位, 它们的 CRISPR 结构差异较大[表 1, 图 1(d)、图 1(g)], 推测这 2 株蜡状芽孢杆菌曾因生活环境的差异而遭受过不同外源遗传物质的入侵。在菌株 subsp. *cytotaxis* NVH 391-98 中, 座位 Bccy2 和 Bccy3 的重复序列相同, 且数量相对较多(6 个和 9 个), 尤为重要, 有 7 个 *cas* 基因位于两座位之间。

在 BT 菌株 Al Hakam、YBT-1520、YBT-020 和 CT-43 中分别发现了 5 个、4 个、2 个和 6 个 CRISPR 座位, 是蜡状芽孢杆菌群 10 个菌株中含有 CRISPR 座位最多的。在菌株 YBT-1520 和 CT-43 中, 分别有 2 个 CRISPR 座位位于各自的大质粒 pBMB293 和 pCT281 上, 且结构非常相似, 分别有 10 个和 9 个 *cas* 基因位于各自的 2 个座位之间[表 1, 图 1(i)、图 1(h)], 其中座位 Btyb3 和 Btct4, 座位 Btyb4 和 Btct5 完全一致(表 1)。座位 Btyb3 和 Btct4 重复序列的 3' 端序列为 GAAAT, 仅与保守的 GAAA(C/G) 序列最后一个碱基不同。

2.2 前导序列

通过对 19 个座位上游的 700bp 核苷酸序列进行比对分析, 寻找前导序列。使用 ClustalX 软件^[14]对这些序列进行多轮比对, 最后在 9 株菌的 10 个 CRISPR 座位上游发现了一段富含 AT、不太保守、具有回文性质的序列。含有前导序列的 CRISPR 座位分别为 Bcg1、Btal4、Baa1、Bast1、Baam1、Baa2、Btyb3、Btct4、Bccy2 和 Bwkb2(图 2)。这些前导序列具有以下共同特征: ①含有多个连续的“A”和多个连续的“T”; ②距离第一个重复序列约 50bp, 全长 55 个碱基; ③AT 含量高达 75% ~ 93%。进一步分析表明, 座位 Bccy2、Btyb3 和 Btct4 的上游序列虽含有前导序列, 但其下游的座位 Bccy3、Btyb4 和 Btct5 两侧都不含有前导序列, 因此推测 *cas* 基因簇上下游的 CRISPR 座位很可能共用同一个前导序列, 但具体情况还有待于进一步研究。

2.3 重复序列

根据以上对前导序列的分析, 将所有 19 个 CRISPR 座位分为两类: 一类上游含有富含 AT 的前导序列, 10 个具有前导序列的座位重复序列长度为 23 ~ 42bp, 数量为 2 ~ 11 个。重复序列保守性分析表明, 除了座位 Baam1、Baa2 和 Bast1 的重复序列相同, 座位 Btyb3 和 Btct4 的重复序列相同外, 其它各重复序列并不具有明显的保守性。另一类则不含前导序列。

表 1 蜡状芽孢杆菌群 10 个菌株的 CRISPR
Table 1 The CRISPR loci in 10 *Bacillus cereus* group genomes

Name	Number of CRISPR loci	CRISPR locus name	CRISPR Loci	Number Of repeats	Repeat size	Spacer size /bp (min-max)	<i>cas</i> genes
BA str. ‘Ames Ancestor’	2	Baam1 ^(c)	Chr	4	23	18, 42, 46	No
		Baam2 ^(a)	Chr	3	23	31, 34	No
BA str. A0248	3	Baa1 ^(c)	Chr	2	25	59	No
		Baa2 ^(c)	Chr	4	23	18, 42, 46	No
		Baa3 ^(a)	Chr	3	23	31, 34	No
BA str. Sterne	2	Bast1 ^(c)	Chr	4	23	18, 42, 46	No
		Bast2 ^(a)	Chr	3	23	31, 34	No
BC G9842	2	Bcg1 ^(c)	Chr	2	36	45	No
		Bcg2 ^(a)	Chr	3	23	31, 34	No
BC subsp. <i>cytotoxis</i> NVH 391-98	3	Bccy1 ^(b)	Chr	2	38	53	No
		Bccy2 ^(c)	Chr	6	29	36(35-37)	7
		Bccy3 ^(c)	Chr	9	29	37(36-38)	7
BT str. Al Hakam	5	Btal1 ^(a)	Chr	2	25	59	No
		Btal2 ^(b)	Chr	2	26	56	No
		Btal3 ^(b)	Chr	2	26	35	No
		Btal4 ^(c)	Chr	2	42	39	No
		Btal5 ^(a)	Chr	3	23	31, 34	No
BT YBT-1520	4	Btyb1 ^(b)	Chr	3	31	30	No
		Btyb2 ^(a)	Chr	3	23	31, 34	No
		Btyb3 ^(c)	plasmid pBMB293	11	32	34(32-35)	10
		Btyb4 ^(c)	plasmid pBMB293	6	24	42(41-43)	10
BT CT-43	6	Btet1 ^(a)	Chr	2	32	52	No
		Btet2 ^(a)	Chr	2	27	57	No
		Btet3 ^(a)	Chr	3	23	31, 34	No
		Btet4 ^(c)	plasmid pCT281	11	32	34(32-35)	9
		Btet5 ^(c)	plasmid pCT281	6	24	42(41-43)	9
		Btet6 ^(b)	plasmid pCT83	3	39	24, 26	No
BT YBT-020	2	Btyt1 ^(a)	Chr	2	28	26	No
		Btyt2 ^(b)	Chr	2	26	56	No
BW KBAB4	2	Bwkb1 ^(a)	Chr	3	23	31, 28	No
		Bwkb2 ^(c)	plasmid pBWB404	2	28	36	No

Note: (a) Not real CRISPR locus; (b) Questionable CRISPR locus; (c) Has leader conservative motif around.

研究发现,重复序列的核苷酸序列可以转录为 RNA 并形成 RNA 二级结构^[3,16],这种现象也已经在古菌中通过实验得到验证^[16]。这种结构对于 CRISPR 行使功能起着关键的作用。由于每个 CRISPR 座位的重复序列几乎一样,故分别选取 10 个具有前导序列的座位中的代表性重复序列进行 RNA 二级结构分析(图 3)。经预测,座位 Baa1 和 Bccy2 的重复序列不能形成 RNA 二级结构,其它 8 个 CRISPR 座位的重复序列所形成的 RNA 二级结构主要分为两类:一类以图 3(a)为代表,形成以茎为主,两端各有一个小环的 RNA 二级结构;第二类包括图 3(b)、图 3(c)、图 3(d)和图 3(e),

形成以环为主,茎的一端为一个大环或多个环的 RNA 二级结构。Kunin 等^[3]对 195 个微生物基因组中的 CRISPR 重复序列进行了分析,其中的一些群具有稳定和高度保守的 RNA 二级结构,在重复序列转录为 RNA 的过程中,间区序列与外源 DNA 或 RNA 发生作用通常是通过相邻的两个重复序列互补完成,但当重复序列形成以“环”为主的 RNA 二级结构时,是通过单个“重复序列-间区序列”单元完成,在该过程中推测重复序列可能介导外源遗传组分(DNA 或 RNA)与 CAS 编码蛋白的相互作用。本研究得到的第二类 RNA 二级结构结果与文献报道一致,故作相同推测。



图 1 蜡状芽孢杆菌群 10 个菌株的 CRISPR 座位结构

Fig. 1 Graphic representation of CRISPR loci in 10 *Bacillus cereus* group

The red and black box arrows represent the genes in different strand; The green box represents the leader motif; the blue diamonds indicate the CRISPR loci in repeat-spacer units. The words below each gene show the name of that one. (a-j) CRISPR loci in BA str. ‘Ames Ancestor’, A0248 and Sterne; BC G9842; BT str. Al Hakam; BW KBAB4; BC subsp. *cytotoxicus* NVH 391-98; BT CT-43, YBT-1520 and YBT-020.



图 2 蜡状芽孢杆菌群 10 个 CRISPR 座位前导序列比对分析

Fig. 2 Alignment of the 10 conserved upstream sequences and the AT rich structure (in black box) of 9 *Bacillus cereus* group species The asterisk shows the conserved base in all the 10 CRISPR loci

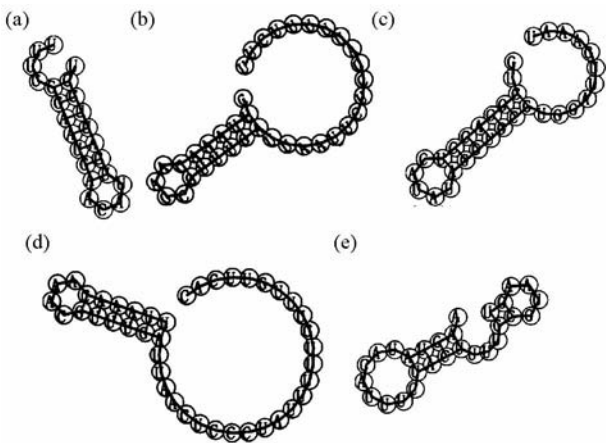


图3 8个CRISPR座位中重复序列的RNA二级结构预测
Fig.3 The predicted RNA secondary structure of 8 repeats

(a) Predicted secondary structure of direct repeat sequence of Baam1, Baa2 and Bast1 (b) Predicted secondary structure of direct repeat sequence of Bcg1 (c) Predicted secondary structure of direct repeat sequence of Btyb3 and Btct4 (d) Predicted secondary structure of direct repeat sequence of Btal4 (e) Predicted secondary structure of direct repeat sequence of Bwkb2

2.4 间区序列和cas基因

在19个CRISPR座位中共含有64个间区序列。其中BT菌株的座位Btyb3和Btct4间区序列数最多,各有10个,且完全相同,它们下游座位Btyb4和Btct5的间区序列也完全一致。以上具有相同间区序列的CRISPR座位重复序列也相同,因此,其可作为菌株分型的分子标记。BC菌株 subsp. cytotoxis NVH 391-98的cas基因簇两侧各具有1个CRISPR座位,座位Bccy2和Bccy3的重复序列虽然相同,但间区序列完全不同。

进一步对全部64个间区序列进行同源性搜索,其中16%与噬菌体、21%与质粒、3%与蜡状芽孢杆菌群基因组序列具有同源性(表2)。BT菌株YBT-1520和CT-43的座位Btyb3和Btct4的10个间区序列中,有6个与噬菌体或质粒具有很高同源性,表明在进化中这些噬菌体或质粒很可能侵染过宿主,并在宿主中留下印记,使宿主获得了针对它们的特异性免疫防御功能,进一步的验证工作本课题组正在进行。

表2 Spacer序列分析
Table 2 Spacer sequence analysis

CRISPR locus	Spacer sequence	The possible resource of spacer sequence
Baam1 & Baa1&Bast1	CTCGTTCATTAAATTTACTTTTTAGCTCTATTAAACTATGGA	Cyanothecae plasmid Cy782203
Bccy2	TCCACATAGAAGCATTTGTGAATCATAATTTTTGCTGT CTTTTTTGGTCTCGCAGCTTTCCACATCATGAATTT	Streptococcus phage phi3396 Bacillus phage AP50
Bccy3	TTTCTAACAATGTCCAAAAACCATTGAATACATTGG CCGTGTTGGTGGCGTTGGTGCTGCATTGCAAACCGT	Salmonella enterica plasmid pHCM2 Enterobacteria phage 285P
Btal3	TTTTAAGTATCGGTGCGATAAAATGTGAAAATCGCC	BC subsp. Cytotoxis NVH 391-98 genome
Btal4	TCTTGAGCTGCAACAATTTCTCCCTACCGAACATATCCC	BC subsp. Cytotoxis NVH 391-98 genome
Btyb3& Btct4	ACAAGATACGAGAATTTAATGACGTCACCCACAT ACAATGGGAAAAGTGTAGGTTATGCACGTTTTAAA GAAGATAGTAGAAGAAATCAAAAAAAGTGAT TAGGAGTTATTAATGGCTTTAGATGTTAGACCA TTTGCAACTAGCTCAATTTGTTTTCCGTTTTCA ATAATATCACAAATTTTGACAAAATTCAGTCATG	BT phage GIL16c BT plasmid DNA Vibrio harveyi plasmid pVH1 BT phage Bam35c and GIL16c, Enterobacter cloacae plasmid pECLA BC AH187 plasmid pAH187_45 Bacillus phage 250 and IEBH, Staphylococcus epidermidis plasmid SAP024A
Btyb4& Btct5	AATGGAACGTTCTTACGGATCAAAAGAAGTCGCACCTTATA	BC plasmid pAH187_45 and pBMB28
Btct6	GAAATAAAATTTGTTATCTTTAGA AAAACCGTTTTTGTATAGATAGGCA	BW plasmid pBWB402 Phage T5

Note: & represents AND.

具有cas基因的3株菌有两个共同特点:①cas基因簇两侧分别含有一个CRISPR座位;②cas基因簇两侧的CRISPR座位中仅一侧的座位具有前导序列。BC菌 cytotoxis NVH 391-98与BT菌株YBT-1520和CT-43的cas基因数量和排列顺序都存在差异,但cas3基因和

cas5基因都保守存在。此外,BT菌株YBT-1520和CT-43都含有多个被截断的csd2基因片段。

3 讨论

CRISPR/Cas系统作为一种抵御外源遗传物质侵

染的新型原核生物防御系统,近年来,对其研究逐渐深入。本研究首次对蜡状芽孢杆菌群的 CRISPR 进行系统的生物信息学分析,发现它们具有重复序列和间区序列较少,重复序列形成的 RNA 二级结构不完全保守等特点。在蜡状芽孢杆菌群 31% 的间区序列中发现了与其具有高同源性的噬菌体、质粒和蜡状芽孢杆菌群基因组序列,进一步证明间区序列可能来源于外源可移动遗传因子,为今后研究宿主菌的获得性免疫防御系统提供数据和线索。

由于蜡状芽孢杆菌群各菌株间具有极高的基因组序列相似性,因而对于一些菌株很难进行分类,特别是一些没有显著特性的菌株。而在已测序的蜡状芽孢杆菌群中,CRISPR 座位分布较广,间区序列可以为菌株的分型提供新的分子标记。此外,CRISPR 座位的作用机制与真核生物的 RNA 干扰非常相似,利用这一特点,我们对 CRISPR 进行改造,插入待研究的基因序列,就可以达到将目标基因进行沉默的目的,为基因敲除提供新思路。

在分析的所有 19 个 CRISPR 座位中有 3 个座位 (Btyb3、Btct4 和 Bccy2) 含有前导序列并与其下游的 *cas* 基因簇组成了完整的 CRISPR/Cas 系统,由于座位 Bccy2 的重复序列不能形成 RNA 二级结构,故推测座位 Btyb3 和 Btct4 的 CRISPR/Cas 系统最有可能具有抵御外源遗传物质入侵,保护细胞基因组完整性的功能。

通过分析发现在 BT 菌株 YBT-1520 和 CT-43 中特有的现象:座位 Btyb4 和 Btct5 的两侧分别含有 6 个和 5 个被截短的 *csd2* 基因片段。据此,我们作出两种猜测:①野生菌株 YBT-1520 和 CT-43 的电转化效率最高,只有 10^3 ,而突变株 BMB171 (仅有 1 个大质粒,且不含 CRISPR 座位) 的电转化效率可高达 10^{10} ,如果 *csd2* 基因片段是菌株在进化中保留的功能性结构域,

推测其发挥抵御外源遗传物质侵染的功能将很可能是两株菌很难进行质粒转化的原因;②全基因组测序结果表明,菌株 YBT-1520^[17] 和 CT-43 (待发表) 都含有 10 个大小不等的质粒,如果 *csd2* 基因片段是菌株进化中残留的不具功能的片段,它的缺失是否与两株菌可以同时容纳多个质粒相关? 进一步研究本课题组正在进行。

参考文献

- [1] Garneau J E, Dupuis M E, Villion M, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 2010, 468(7320):67-71.
- [2] Al-Attar S, Westra E R, van der Oost J, et al. Review: Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol Chem*, 2011, 392(4):277-289.
- [3] Kunin V, Sorek R, Hugenholtz P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol*, 2007, 8(4):R61.
- [4] Lillestol R K, Redder P, Garrett R A, et al. A putative viral defence mechanism in archaeal cells. *Archaea*, 2006, 2(1):59-72.
- [5] Shah S A, Hansen N R, Garrett R A. Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism. *Biochem Soc Trans*, 2009, 37(Pt 1):23-28.
- [6] Jansen R, Embden J D, Gaastra W, et al. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*, 2002, 43(6):1565-1575.
- [7] Pul U, Wurm R, Arslan Z, et al. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol Microbiol*, 2010, 75(6):1495-1512.
- [8] Haft D H, Selengut J, Mongodin E F, et al. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol*, 2005, 1(6):e60.
- [9] Zhu Y, Shang H, Zhu Q, et al. Complete genome sequence of *Bacillus thuringiensis* serovar. finitimus Strain YBT-020. *J Bacteriol*, 2011, 193(9):2379-2380.
- [10] He J, Shao X, Zheng H, et al. Complete genome sequence of *Bacillus thuringiensis* mutant strain BMB171. *J Bacteriol*, 2010, 192(15):4074-4075.
- [11] Grissa I, Vergnaud G, Pourcel C. CRISPRfinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res*, 2007, 35(Pp):W52-W57.
- [12] Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, 2007, 8:172.
- [13] Delcher A L, Harmon D, Kasif S, et al. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res*, 1999, 27(23):4636-4641.
- [14] Larkin M A, Blackshields G, Brown N P, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*, 2007, 23(21):2947-2948.
- [15] Mathews D H, Sabina J, Zuker M, et al. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol*, 1999, 288(5):911-940.
- [16] Tang T H, Bachellerie J P, Rozhdestvensky T, et al. Identification of 86 candidates for small non-messenger RNAs

[1] Garneau J E, Dupuis M E, Villion M, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid

from the archaeon *Archaeoglobus fulgidus*. Proc Natl Acad Sci USA, 2002, 99(11):7536-7541.
[17] Zhong C, Peng D, Ye W, et al. Determination of plasmid copy

number reveals the total plasmid DNA amount is greater than the chromosomal DNA amount in *Bacillus thuringiensis* YBT-1520. PLoS One, 2011, 6(1):e16025.

Bioinformatics Analysis of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) in the Genomes of *Bacillus cereus* Group

WANG Yan¹ YU Chan¹ WANG Jie-ping¹ QIU Ning¹

HE Jin¹ SUN Ming¹ ZHANG Qing-ye^{1,2}

(1 State Key Laboratory of Agricultural Microbiology and National Engineering Research Center of Microbial Pesticides, College of Life Science, Huazhong Agricultural University, Wuhan 430070, China)
(2 School of Science, Huazhong Agriculture University, Wuhan 430070, China)

Abstract CRISPR is a novel type of microbial defense system, which is unique in that it is invader-specific, adaptive and heritable. It is a recent breakthrough in understanding host-virus interactions. Bioinformatics methods including BLAST, multiple sequence alignment, and RNA structure prediction was used to analyze the CRISPR structures of 24 *Bacillus cereus* group genomes. CRISPR existed in 42% strains. Two types of RNA secondary structures derived from the repeat sequences were predicted, and demonstrated that stem-loop secondary structure might function in mediating the interaction between foreign genetic elements and CAS-encoded proteins. The sequence homologous among 31% spacer, phage, plasmid and the genomes of *Bacillus cereus* group further verified that spacer was likely to come from the exogenous mobile genetic factor. As most of the *Bacillus cereus* group strains contain multiple plasmids and prophages, the CRISPR research in *Bacillus cereus* group by this study would be help to reveal relationship between host strains with plasmid or host strains with phage.

Key words *Bacillus cereus* group CRISPR Leader Repeat Spacer

梅特勒-托利多参加 2011 世界制药机械、 包装设备与材料中国展

2011 世界制药机械、包装设备与材料中国展与 2011 世界生化、分析仪器与实验室设备中国展以及世界制药原料中国展近日在上海召开。梅特勒-托利多在此次会上重点展出了实验室研发和分析检测仪器、在线检测仪器、包装检测仪器和工业称重设备,这些仪器设备将帮助制药行业用户加快研发进程,提升生产效率和产品质量。配合今年中国新版 GMP 的正式实施,梅特勒-托利多全方位的制药行业解决方案,有助于制药行业用户轻松满足 GMP 要求。会上,梅特勒-托利多的高级技术应用顾问刘慧敏博士介绍了 RTCal/iC Safety 实时在线反应量热技术以及 ReactIR 流通池在连续工艺技术中的应用,梅特勒-托利多(美国)的自动化学咨询顾问王建博士讲解了用工艺过程分析技术来提高工艺效率和产品质量。