

# 嗜热链球菌 CRISPR 序列的检测及原间隔序列预测

李婉, 王娜娜, 张丹青, 霍贵成

(东北农业大学, 乳品科学教育部重点实验室, 黑龙江哈尔滨 150030)

**摘要:** CRISPR 基因座中的间隔序列能够提供特异性免疫使宿主得以对抗那些携带相同序列的入侵元件, 本研究旨在预测实验室 3 株嗜热链球菌的原间隔序列, 为探究其抗噬菌体机制奠定基础。本文检测了 3 株嗜热链球菌的 CRISPR 序列, 预测嗜热链球菌 CRISPR 基因座的活性, 构建重复序列的系统发育树, 并对间隔序列进行同源性分析。结果显示供试菌 S0 含有 3 个 CRISPR, S4 仅含 1 个 CRISPR, 79 含有 4 个 CRISPR, CRISPR1 基因座活性最高, CRISPR2 基因座最不活跃。重复序列与标准菌株的相应重复序列高度保守, 不同 CRISPR 的间隔序列在内容及数目上高度可变, 它们的原间隔序列绝大部分来源于噬菌体, 少数来源于质粒或染色体序列。供试菌 79 与嗜热链球菌 MN-BM-A02 及嗜热链球菌 ASCC 1275 的 4 个 CRISPR 序列均高度一致, 推测其菌株同源性较高。验证了间隔序列是由宿主遭受噬菌体等外源基因元件侵染而获得的。

**关键词:** 嗜热链球菌; 同源性; 重复序列; 间隔序列

文章编号: 1673-9078(2016)10-252-258

DOI: 10.13982/j.mfst.1673-9078.2016.10.038

## CRISPR Detection and Protospacer Prediction in *Streptococcus thermophilus*

LI Wan, WANG Na-na, ZHANG Dan-qing, HUO Gui-cheng

(Key Lab of Dairy Science, Ministry of Education, Northeast Agricultural University, Harbin 150030, China)

**Abstract:** The spacers in clustered regularly interspaced short palindromic repeats (CRISPR) loci can provide specific immunity to invasive elements that carry the same or a similar sequence. In the present study, the protospacers of three *S. thermophilus* strains were predicted to provide a foundation for exploring the phage resistance mechanism. The CRISPR sequences of three *S. thermophilus* strains were determined, the activity of CRISPR loci in several *S. thermophilus* strains was predicted, and a phylogenetic tree for repeated sequences was constructed. In addition, repeat or spacer homology was analyzed. The results showed that the tested bacterium S0 contained three CRISPR loci, S4 only had one CRISPR locus, and strain 79 had four CRISPR loci. We hypothesized that the highest and lowest activities would be observed in CRISPR1 and CRISPR2, respectively. The repeated sequences and corresponding repeated sequences in the standard strain were highly conserved, and different CRISPR spacer sequences were highly dynamic in terms of content and number. Most original spacer sequences were derived from bacteriophages, whereas a few were derived from plasmid or chromosomal sequences. The CRISPR sequence of strain 79 was highly consistent with the four CRISPR sequences of *S. thermophilus* MN-BM-A02 and *S. thermophilus* ASCC 1275, suggesting that the strains are highly homologous. The results suggest that the spacer sequences were acquired from invasion of the host by the bacteriophage and other exogenous genetic elements.

**Key words:** *Streptococcus thermophilus*; homology; repeat; spacer

环境中无所不在的噬菌体在工业环境中对发酵过程产生反复的负面影响, 在大发酵罐中, 乳酸菌的噬菌体感染一直是发酵乳产业面临严峻挑战, 噬菌体会导致产酸缓慢甚至是发酵失败, 造成巨大的经济损失

收稿日期: 2015-11-08

基金项目: 国家自然科学基金项目(31171717); 国家“863”科技计划(2012AA022108)

作者简介: 李婉(1990-), 女, 硕士研究生, 研究方向: 乳酸菌生物技术

通讯作者: 霍贵成(1958-), 男, 教授, 博士生导师, 研究方向: 营养与食品安全

失。乳制品发酵易受噬菌体侵染主要有以下几方面原因: 巴氏消毒的奶并不是无菌的, 它或许含有噬菌体; 噬菌体污染在液态奶中易于传播; 在无防腐剂的加工条件下重复使用指定的培养基从而为噬菌体传播提供了恒定的宿主<sup>[1]</sup>。相应地, 乳酸菌进化出多种抗噬菌体机制, 但这些先天的防御系统不能提供特异性免疫, 而 CRISPR (Clustered regularly interspaced short palindromic repeats)-Cas (CRISPR-associated) 系统能够为绝大多数古细菌和多数细菌提供获得性免疫以抵抗裂解性噬菌体或接合质粒等外源基因元件入侵<sup>[2,3]</sup>。

CRISPR-Cas 基因座由一系列重复-间隔排列及多种 *cas* 基因共同组成, *cas* 基因编码的 Cas 蛋白在免疫响应中起重要作用。在同一 CRISPR 基因座中, 重复序列高度保守, 间隔序列则高度可变。即使在种源关系很近的菌株间, 间隔序列的成分也是高度可变且进化迅速的。CRISPR-Cas 系统在古细菌中更为常见, 且在嗜热微生物中比在嗜温微生物中更丰富、间隔序列数目更多<sup>[4]</sup>。部分间隔序列与噬菌体等入侵元件的片段(原间隔序列)完全相同, 这一创新性的发现开启了对 CRISPR-Cas 系统独特的作用机制的探究。随后又发现间隔序列可能来源于质粒, 甚至在一些 CRISPR 基因座中, 间隔序列来源于自身染色体<sup>[5]</sup>。

间隔序列的获得依赖于 Cas1-Cas2 蛋白复合物, 它能从入侵元件中剪切原间隔序列并整合到 CRISPR 序列中<sup>[6-8]</sup>, Cas1-Cas2 蛋白复合物被证实是催化 CRISPR 基因座获得间隔序列的最小结构<sup>[9]</sup>。原核细胞中一个或多个间隔序列被由 Cas 蛋白组成的 Cascade 复合物转录加工成小的 CRISPR RNA (crRNA) 分子。Cascade 复合物随后在 crRNA 与入侵核酸分子中的同源原间隔序列之间形成双工, 触发后者降解<sup>[10,11]</sup>, 从而终止感染。

CRISPR-Cas 系统能够持续获得新的间隔序列这一事实部分重建了噬菌体等外源基因元件的感染历史<sup>[12,13]</sup>。每一个间隔序列都能够提供特异性免疫使宿主得以对抗那些携带相同序列的入侵元件, 独特的间隔序列被认为是宿主防御入侵的噬菌体或质粒 DNA 的一种免疫记忆系统<sup>[14]</sup>。本研究旨在查找实验室 3 株嗜热链球菌的原间隔序列, 分析其受噬菌体、质粒等外源基因元件感染的历史, 预测其噬菌体抗性范围, 为后期探究其抗噬菌体机制奠定基础。

## 1 材料与方法

### 1.1 实验菌株

嗜热链球菌 S0、S4、79 均从东北农业大学教育部重点实验室工业微生物菌种保藏中心(KLDS-DICC)获得。培养基采用 M17 肉汤培养基。

### 1.2 试剂与仪器

#### 1.2.1 试剂

细菌基因组提取试剂盒, 北京天根生物技术有限公司; 溶菌酶, Solabiro 公司; TIANGEN 2×PCR MasterMix, 天根生化科技有限公司; 引物, 深圳华大基因科技有限公司; M17 肉汤, 青岛高科园海博生物技术有限公司。

#### 1.2.2 仪器

LDZF-50KB-II 立式蒸汽灭菌器, 上海申安医疗器械厂; GL-20G-II 离心机, 上海安亭科学仪器厂; DYY-10C 电泳仪, 北京六一仪器厂; UVP 凝胶成像系统, 美国 UVP 公司; PCR 扩增仪, Applied Biosystems; DHP-927 型电热恒温培养箱, 上海一恒科技有限公司; CJ-20 超净工作台, 天津市泰斯特仪器有限公司。

### 1.3 实验方法

#### 1.3.1 菌株的培养及基因组 DNA 的提取

将冷冻保存的实验菌株转接至 M17 肉汤培养基中, 37 °C 活化培养两代, 每代 24 h。菌株总 DNA 的提取参照细菌基因组提取试剂盒说明书。

#### 1.3.2 CRISPR 序列扩增与测序

以 NCBI (<http://www.ncbi.nlm.nih.gov/>) 上已公布全基因组序列的 9 株嗜热链球菌为标准菌株, 根据其 CRISPR 上下游序列的保守性, 利用 Primer Premier 6.0 软件针对不同类型 CRISPR-Cas 系统设计引物, 对 3 株供试嗜热链球菌菌株进行 CRISPR 序列扩增。PCR 反应体系如下, 反应体系: 25 μL, 其中 DNA 模板: 1 μL; 正向引物: 1 μL; 反向引物: 1 μL; 2×MasterMix: 12.5 μL; ddH<sub>2</sub>O: 9.5 μL。正、反向引物序列见表 1。

PCR 反应程序: 94 °C 预变性 5 min, 35 个循环包括: 94 °C 变性 30 s, 54 °C 退火 30 s, 72 °C 延伸 1 min。最后 72 °C 再延伸 7 min。PCR 产物进行琼脂糖凝胶电泳检测, 随后送华大科技公司对目的片段序列进行测定。

#### 1.3.3 CRISPR 基因座活性分析

利用 CRISPR Finder 对 NCBI 中已公布全基因组序列的 9 株嗜热链球菌标准菌的 CRISPR 序列进行查找, 统计标准菌株及供试菌株中每个 CRISPR 所含的独特间隔序列的数目, 利用 GraphPad Prism 5 软件制作散布图。

#### 1.3.4 CRISPR 序列同源性分析

测序结果利用 CRISPR Finder (<http://crispr.u-psud.fr/Server/>) 进行 CRISPR 序列查找, 对重复序列和间隔序列进行整理。将供试菌株与嗜热链球菌的重复序列利用 MEGA5 软件下的邻接法(Neighbor-Joining Method), 自展值(bootstrap value)分析设置重复次数为 1000 次构建系统发育树, 对重复序列进行系统发生学分析。间隔序列进行 NCBI blast, 查找同源性序列, 预测原间隔序列。

## 2 结果与讨论

### 2.1 CRISPR 序列扩增与测序

根据 NCBI 中已公布全基因组序列的嗜热链球菌为标准菌株设计 CRISPR 序列的扩增引物如表 1。PCR 结果如图 1，图 1a 中 3 株嗜热链球菌在 2500 bp 附近均有清晰条带；图 1b 中 S4 无扩增结果，S0 在

300 bp 处有清晰条带，79 在 500 bp 处有清晰条带，且扩增效率高、特异性强；图 1c 中 S4 仍无扩增条带，S0 在 1500 bp 处有清晰条带，79 在 800 bp 处有清晰条带，在 400 bp 处有不清晰条带，初步判断为干扰带；图 1d 中仅 79 在 1000 bp 处有清晰条带。

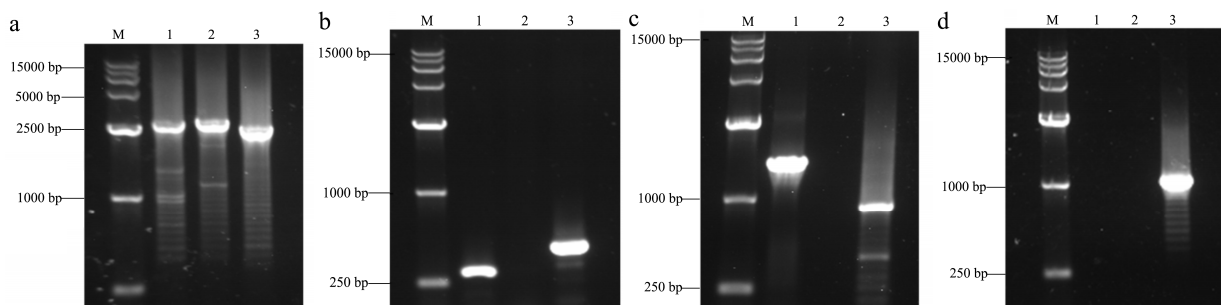


图 1 CRISPR 序列扩增结果

Fig.1 Amplification results of CRISPR sequence

注：图 1a、1b、1c 和 1d 分别为 CRISPR1、CRISPR2、CRISPR3 和 CRISPR4 序列电泳图。M: DNA Marker; Lane 1: S0; lane 2: S4; lane 3: 79。

表 1 CRISPR 序列的扩增引物

Table 1 Amplification primers of CRISPR sequences

Primer name	Primer sequence(5'→3')	Annealing temperature/°C
CRISPR1	F: TGCTGAGACAACCTAGTCTCTC	56
	R: TAAACAGAGCCTCCCTATCC	
CRISPR2	F: TTAGCCCCTACCATAGTGCTG	55
	R: TTAGTCTAACACTTCTGGAAGC	
CRISPR3	F: AGTTCTAAGAGATGTCAACCGA	56
	R: AGCTGAAGTCATGCTGAGATTA	
CRISPR4	F: GCTTACAACGGCAATACTTA	56
	R: GAGATATGTCAAGGATGAGAAG	

表 2 嗜热链球菌 CRISPR 概况

Table 2 General information of *Streptococcus thermophilus* CRISPR

CRISPR	CRISPR Length /bp	Number of Spacer	DR Length/bp	DR
S0				
CRISPR1	2409	36	36	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC
CRISPR2	101	1	30	GATATAAACCTAATTACCTCGAGAGGGGAC
CRISPR3	1292	20	36	GTTTTGGAACCATTCGAAACAACACAGCTCTAAAC
S4				
CRISPR1	2744	41	36	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC
79				
CRISPR1	2146	32	36	GTTTTTGTACTCTCAAGATTTAAGTAACTGTACAAC
CRISPR2	258	3	36	GATATAAACCTAATTACCTCGAGAGGGGACGGAAAC
CRISPR3	827	12	36	GTTTTGGAACCATTCGAAACAACACAGCTCTAAAC
CRISPR4	762	12	28	GGATCACCCCGCGTGTGCGGAAAAAC

注：DR 为重复序列。

## 2.2 嗜热链球菌 CRISPR 序列的检测

测序结果经 CRISPR-Finder 进行 CRISPR 序列查找, 除图 1c 中 79 在 400 bp 处的条带非 CRISPR 序列外, 其余扩增条带均鉴定为 CRISPR 序列。CRISPR

序列信息如表 2, 供试菌 S0 含有 3 个 CRISPR, 供试菌 S4 仅含 1 个 CRISPR, 供试菌 79 含有 4 个 CRISPR。其中 CRISPR 序列最长为 2744 bp, 含有 41 个间隔序列, 最短仅 101 bp, 仅含 1 个间隔序列。重复序列最长为 36 bp, 最短为 28 bp。

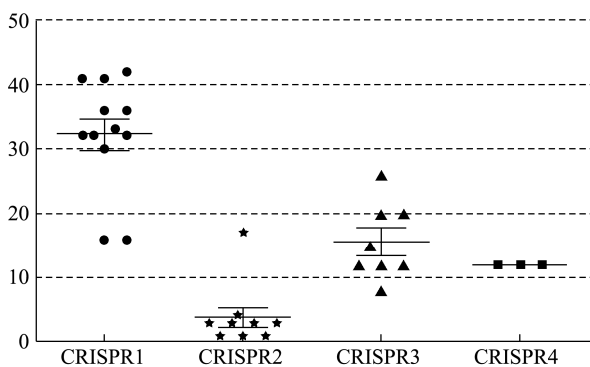


图 2 嗜热链球菌 CRISPR 间隔序列的数目分布

Fig.2 Distribution of the number of CRISPR spacer sequences in *Streptococcus thermophilus*

## 2.3 CRISPR 基因座活性分析

9 株标准嗜热链球菌及 3 株供试菌株各个 CRISPR 基因座所含的独特间隔序列数目散布图如图 2。CRISPR1 基因座中间隔序列最多 42 个, 平均约 32 个; CRISPR2 基因座中间隔序列最多为 17 个, 但其他不超过 5 个, 平均约 4 个; CRISPR3 基因座中间隔序列最多为 26 个, 平均约 16 个; 3 个 CRISPR4 基因座的间隔序列均为 12 个。

每个独特的间隔序列都是从入侵的外源基因元件(如烈性噬菌体、接合质粒等)中获得的, 因此一个 CRISPR 基因座独特间隔序列的数目可在一定程度上反映该 CRISPR 基因座的活跃程度。由图 2 可以看出, CRISPR1 基因座中间隔序列数目的最大值和平均值均为最高, 因此推测嗜热链球菌中 CRISPR1 基因座最为活跃, CRISPR3 基因座次之。而 CRISPR2 基因座中间隔序列数目的平均值最低, 且除 1 个 CRISPR2 基因座含 17 个间隔序列外, 其他均不超过 5 个, 因此推测嗜热链球菌中 CRISPR2 基因座活性最低。

## 2.4 重复序列同源性分析

将供试菌株的重复序列与 NCBI 中已公开发表全基因组序列的 9 株嗜热链球菌 CRISPR 重复序列构建发育树如图 3, 其中 DR1、DR2、DR3、DR4 分别对应 CRISPR1、CRISPR2、CRISPR3、CRISPR4 的重复序列。该发育树主要包含 4 个分支, 3 株菌株的 DR1 与 9 株标准菌株的 DR1 位于同一分支, 供试菌株与标准菌株的 DR2 位于同一分支, 说明其同源性较高, S0、79 的 DR3 与标准菌株 *St* LMD-9、*St* ND03、*St* MN-ZLW-002 及 *St* ASCC 1275、*St* MN-BM-A02、*St* SMQ-301 的 DR3 位于同一分支上, 79 与 *St* ASCC 1275、*St* MN-BM-A02 的 DR4 位于同一分支, 说明其进化关系密切。

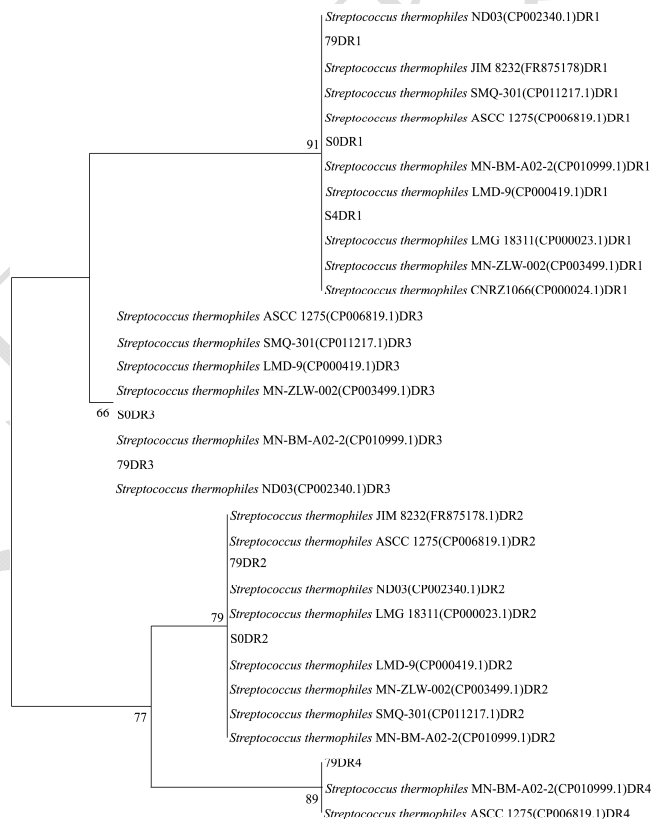


图 3 重复序列系统发育树

Fig.3 Phylogenetic tree of direct repeats

## 2.5 间隔序列同源性分析

将 3 株供试菌的间隔序列逐一进行 NCBI blast, 查找与其完全一致的原间隔序列, 并预测原间隔序列的活性。Blast 分析发现 157 个间隔序列中仅 39 个比对上完全一致的原间隔序列, 比对率低于 25%, 这可能是由于相应的噬菌体数据库不完善造成的。而这 39 个间隔序列却比对上 72 个原间隔序列, 暗示 1 个间隔序列可能抵抗多个外源基因元件入侵。其中, S0 的原间隔序列主要来源于噬菌体, 这些原间隔序列所在编

码区主要编码衣壳蛋白、末端酶、受体结合蛋白等。此外,其 CRISPR1 的 spacer32 与乳酸乳球菌 4.0325 染色体的两段序列完全匹配,后经 PHAST 网络服务器 (<http://phast.wishartlab.com/index.html>) 对乳酸乳球菌 4.0325 的原噬菌体进行在线预测,发现这两段序列均位于乳酸乳球菌 4.0325 的原噬菌体上,其所在的编码区均编码末端酶。spacer32 与乳球菌噬菌体 P335 序列完全匹配,说明原间隔序列也可能来源于非同属的噬菌体。

S4 的 41 个间隔序列中仅 7 个匹配上原间隔序列,其中 Spacer38 匹配上嗜热链球菌质粒 pND103、p2992、p2980、pt38、pt39 的热激蛋白序列,其余间隔序列均与噬菌体序列匹配。供试菌 79 有 18 个间隔序列能够完全匹配原间隔序列,CRISPR1 的 Spacer23

的原间隔序列位于副乳房链球菌 KCTC 11537 的基因组上,对乳房链球菌 KCTC 11537 的原噬菌体进行预测发现该原间隔序列位于原噬菌体的卷尺蛋白上,CRISPR1 的 Spacer23 与嗜热链球菌 *repA* 基因相匹配,该基因编码复制蛋白。其余间隔序列均匹配噬菌体的解旋酶、切除酶、引发酶等基因序列。供试菌 79 的 CRISPR2 和 CRISPR3 的间隔序列均匹配噬菌体序列,这些原间隔序列所在编码区可编码引发酶、裂解酶等。

比对上的原间隔序列中有 66 个均来源于噬菌体,5 个来源于质粒,1 个来源于嗜热链球菌染色体。若宿主将染色体序列整合为间隔序列,则会引起自身免疫而损伤染色体,这种自身免疫在噬菌体含有类似 CRISPR 相关 Cas4 蛋白时可能发生<sup>[15]</sup>。

表 3 原间隔序列特点

Table 3 Features of protospacers

S0 CRISPR1	Strains	Activity	Identities
Spacer8	<i>Streptococcus</i> phage Abc2	major capsid protein	100% (30/30)
Spacer29	<i>Streptococcus thermophilus</i> bacteriophage 7201	MPL-7201	100% (30/30)
	<i>Lactococcus lactis</i> subsp. <i>lactis</i> KLDS 4.0325(977275-977304)	terminase	100% (30/30)
Spacer32	<i>Lactococcus lactis</i> subsp. <i>lactis</i> KLDS 4.0325(1733989-1734018)	terminase	100% (30/30)
	<i>Lactococcus</i> phage P335	large terminase subunit	100% (30/30)
S0 CRISPR3	Strains	Activity	Identities
Spacer3	<i>Streptococcus</i> phage TP-778L	unnamed protein	100% (31/31)
	<i>Streptococcus</i> phage ALQ13.2	receptor-binding protein	100% (30/30)
Spacer15	<i>Streptococcus thermophilus</i> bacteriophage 2972	antireceptor	100% (30/30)
S4 CRISPR1	Strains	Activity	Identities
Spacer5	<i>Streptococcus</i> phage TP-J34	minor structural protein	100% (30/30)
	<i>Streptococcus thermophilus</i> bacteriophage Sfi11	minor structural protein	100% (30/30)
	<i>Streptococcus</i> phage pC2S4	tail-host specificity protein	100% (30/30)
Spacer20	<i>Streptococcus thermophilus</i> bacteriophage Sfi11	putative anti-receptor	100% (30/30)
Spacer25	<i>Streptococcus thermophilus</i> bacteriophage Sfi21	phage anti-receptor	100% (30/30)
Spacer29	<i>Streptococcus thermophilus</i> bacteriophage Sfi11	putative anti-receptor	100% (30/30)
Spacer30	<i>Streptococcus thermophilus</i> bacteriophage Sfi11	minor structural protein	100% (30/30)
Spacer38	<i>Streptococcus thermophilus</i> plasmid pND103, p2992, p2980, pt38, pt39	heat shock protein	100% (30/30)
79 CRISPR1	Strains	Activity	Identities
Spacer8	<i>Streptococcus</i> phage 20617 complete prophage genome	putative excisionase	100% (30/30)
Spacer10	<i>Streptococcus</i> phage 20617 complete prophage genome	hypothetical protein	100% (30/30)
	<i>Streptococcus</i> phage 20617 complete prophage genome	helicase	100% (30/30)
Spacer13	<i>Streptococcus</i> phage TP-J34 complete genome	hypothetical protein	100% (30/30)
	<i>Streptococcus</i> phage 5093	DnaC-like protein	100% (30/30)
Spacer14	<i>Streptococcus thermophilus</i> bacteriophage kappa3	putative primase	100% (29/29)

转下页

接上页

	<i>Streptococcus thermophilus</i> bacteriophage DT1	putative primase	100% (29/29)
Spacer23	<i>Streptococcus parauberis</i> KCTC 11537	tape measure protein	100% (30/30)
Spacer24	<i>Streptococcus</i> phage 20617 complete prophage genome	hypothetical protein	100% (30/30)
	<i>Streptococcus</i> phage 5093	hypothetical protein	100% (30/30)
Spacer29	<i>Streptococcus thermophilus</i> partial repA gene	replication protein	100% (30/30)
79CRISPR3	Strains	ORF/Activity	Identities
Spacer2	<i>Streptococcus</i> phage 20617 complete prophage genome	hypothetical protein	100% (30/30)
	<i>Streptococcus</i> phage ALQ13.2	primase	100% (30/30)
	<i>Streptococcus thermophilus</i> bacteriophage kappa3	putative primase	100% (30/30)
	<i>Streptococcus thermophilus</i> bacteriophage 2972	primase	100% (30/30)
	<i>Streptococcus thermophilus</i> bacteriophage Sfi11	putative primase	100% (30/30)
	<i>Streptococcus thermophilus</i> bacteriophage DT1	putative primase	100% (30/30)
Spacer6	<i>Streptococcus thermophilus</i> bacteriophage SFi18	lysin	100% (30/30)
	<i>Streptococcus</i> phage 20617 complete prophage genome	hypothetical protein	100% (30/30)
Spacer9	<i>Streptococcus</i> phage 20617 complete prophage genome	phage lysin	100% (30/30)
Spacer11	<i>Streptococcus</i> phage 20617 complete prophage genome	Activity	Identities
79 CRISPR4	Strains	Activity	Identities
Spacer10	<i>Streptococcus thermophilus</i> temperate phage and flanking sequences, isolate DSM20617	unnamed protein	100% (33/33)
	<i>Streptococcus thermophilus</i> temperate phage and flanking sequences, isolate SK778	unnamed protein	100% (33/33)
	<i>Streptococcus</i> phage 20617 complete prophage genome	hypothetical protein	100% (33/33)
	<i>Streptococcus</i> phage TP-778L complete genome	unnamed protein	100% (33/33)

间隔序列是高度可变的,即使在亲源关系比较近的菌株中,间隔序列的成分也是极具动态且快速进化的,该性质被用于结核分枝杆菌、鼠疫耶尔森菌、沙门氏菌及白喉杆菌等菌株的分型<sup>[16-19]</sup>。在这种情况下,79的4个CRISPR序列与标准菌株嗜热链球菌MN-BM-A02及嗜热链球菌ASCC 1275的4个CRISPR序列完全相同,由此得出供试菌79与这两株标准菌同源性较高。嗜热链球菌MN-BM-A02具有高产酸速率和弱后酸化的特点,而嗜热链球菌ASCC 1275含有一个新的由重复单元组成的胞外多糖基因簇,可生成荚膜多糖和黏液多糖,是目前为止产胞外多糖量最高(~1029 mg/L)的嗜热链球菌<sup>[20]</sup>。这为供试菌79在产酸、产胞外多糖及后酸化能力等方面的探究提供了视角。

### 3 结论

3.1 原间隔序列绝大多数来源于噬菌体,少数来源于质粒,极少数来源于染色体。

3.2 供试菌79与标准菌株嗜热链球菌MN-BM-A02及嗜热链球菌ASCC 1275具有较高的同源性。

3.3 嗜热链球菌中CRISPR1基因座活性最高,CRISPR3基因座次之,CRISPR2基因座活性最低。

### 参考文献

- [1] Klaenhammer T R, Fitzgerald G F. Bacteriophages and bacteriophage resistance [J]. Springer Netherlands, 1994: 106-168
- [2] Bolotin A, Quinquis B, Sorokin A, et al. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin [J]. Microbiology, 2005, 151(8): 2551-2561
- [3] Horvath P, Romero D A, Coûté-Monvoisin A, et al. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus* [J]. Journal of Bacteriology, 2008, 190(4): 1401-1412
- [4] Weinberger A D, Wolf Y I, Lobkovsky A E, et al. Viral diversity threshold for adaptive immunity in prokaryotes [J]. Mbio., 2012, 3(6): 17
- [5] Stern A, Keren L, Wurtzel O, et al. Self-targeting by CRISPR: gene regulation or autoimmunity? [J]. Trends in Genetics, 2010, 26(8): 335-340
- [6] Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes [J]. Science, 2007, 315(5819): 1709-1712

- [7] Deveau H, Barrangou R, Garneau J E, et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus* [J]. Journal of Bacteriology, 2008, 190(4): 1390-1400
- [8] Li M, Wang R, Zhao D H, et al. Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process [J]. Nucleic Acids Research, 2014, 42(4): 2483-2492
- [9] Nu Ez J K, Lee A S Y, Alan E, et al. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity [J]. Nature, 2015, 519(7542): 193
- [10] Haurwitz R E, Jinek M, Wiedenheft B, et al. Sequence-and structure-specific RNA processing by a CRISPR endonuclease [J]. Science, 2010, 329(5997): 1355-1358
- [11] Semenova E, Jore M M, Datsenko K A, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence [J]. Proceedings of the National Academy of Sciences, 2011, 108(25): 10098-10103
- [12] Stern A, Mick E, Tirosh I, et al. CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome [J]. Genome Research, 2012, 22(10): 1985-1994
- [13] Weinberger A D, Sun C L, Pluciński M M, et al. Persisting viral sequences shape microbial crispr-based immunity [J]. Plos Computational Biology, 2012, 8(4): e1002475
- [14] Datsenko K A, Pougach K, Tikhonov A, et al. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system [J]. Nature Communications, 2012, 3: 945
- [15] Hooton S P, Connerton I F. *Campylobacter jejuni* acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein [J]. Frontiers in Microbiology, 2015, 5(5): 744
- [16] Kamerbeek J, Schouls L, Kolk A, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology [J]. Journal of Clinical Microbiology, 1997, 35(4): 907-914
- [17] Cui Y J, Li Y J, Olivier G, et al. Insight into microevolution of *yersinia pestis* by clustered regularly interspaced short palindromic repeats [J]. Plos One, 2008, 3(7): e2652
- [18] Fenyun L, Subhashinie K, Jayarao B M, et al. Subtyping *Salmonella enterica* serovar enteritidis isolates from different sources by using sequence typing based on virulence genes and clustered regularly interspaced short palindromic repeats (CRISPRs) [J]. Applied & Environmental Microbiology, 2011, 77(13): 4520-4526
- [19] Igor M, Olga N, Elena L, et al. Efficient Discrimination within a *corynebacterium diphtheriae* epidemic clonal group by a novel macroarray-based method [J]. J. Clinical Microbiology, 2005, 43(4): 1662-16
- [20] Wu Q, Tun H M, Leung F C, et al. Genomic insights into high exopolysaccharide-producing dairy starter bacterium *Streptococcus thermophilus* ASCC 1275 [J]. Scientific Reports, 2014, 4(7500): 4974