



嗜热链球菌CRISPR-Cas系统的检测

李婉, 梁宏彰, 张丹青, 王娜娜, 唐雅茹, 李柏良, 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: 【目的】CRISPR-Cas系统为嗜热链球菌抵抗噬菌体等外源基因元件提供获得性免疫, 分析NCBI中已公开发表全基因组序列的9株嗜热链球菌所含CRISPR-Cas系统的数目和类型, 对实验室相应菌株的CRISPR-Cas系统进行检测。【方法】利用生物信息学方法对NCBI中9株已测序嗜热链球菌所含CRISPR-Cas系统进行分析, 根据其Cas基因序列设计引物, 对实验室嗜热链球菌菌株的Cas基因进行扩增、测序, 分析实验室6株嗜热链球菌的CRISPR-Cas系统情况。【结果】9株标准菌株均含不同数目的CRISPR-Cas系统, 其类型主要为II-A型、III-A型和I-E型, 各类型的标志Cas基因高度保守。6株供试菌中, S4仅含Cas9基因, 其它5株均含有Cas9基因、Cas10基因和Cas9*基因, 79和KLDS3.0207还含有Cas3基因。【结论】可根据标准菌株高度保守的Cas基因设计引物, 预测未知嗜热链球菌所含CRISPR-Cas系统的数目和类型。S4仅含1个II-A型CRISPR-Cas系统, 其它5株均含有2个II-A型CRISPR-Cas系统和1个III-A型CRISPR-Cas系统, 此外, 79和KLDS3.0207均含有1个I-E型CRISPR-Cas系统。

关键词: 嗜热链球菌, 规律成簇的间隔回文重复(CRISPR), CRISPR-Cas系统, Cas, 生物信息学分析

随着社会经济的不断发展, 消费者的健康意识逐步提高, 发酵乳制品行业发展迅速。乳品工业的发展高度依赖于微生物发酵系统, 嗜热链球菌作为关键的发酵剂菌种, 在牛奶酸化及多种发酵乳产品的质感形成中发挥重要作用^[1]。然而, 在工业生产中, 嗜热链球菌易受噬菌体的侵染, 导致发酵失败, 造成巨大的经济损失。因此, 如何选育具有噬菌体抗性的乳酸菌作为发酵剂, 是当下发酵乳产业亟待解决的问题。

乳酸菌噬菌体对乳品发酵造成巨大威胁, 反过来, 细菌也进化了多种机制以抵抗病毒入侵, 其中包括CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated genes)系统^[2], 它通过获得性免疫抵抗外源DNA元件入侵, 并能将这种抗性传递给后代^[3-6]。CRISPR基因座与噬菌体特异获得性免疫之间的联系最初在嗜热链球菌中得到证明^[7], 有趣的是, 近年来嗜热链球菌又被证明是探究CRISPR基因座

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

*通信作者。Tel: +86-451-55191807; Fax: +86-451-55190340; E-mail: gchuo@vip.0451.com

收稿日期: 2015-07-20; 修回日期: 2015-09-09; 网络出版日期: 2015-09-15

与噬菌体相互作用的最佳特征模型系统^[8]。近年来,随着全基因组测序技术的不断成熟,在约46%的细菌和87%的古细菌中均发现了CRISPR基因座^[9],这些高度可变的区域由重复间隔单元及Cas (CRISPR-associated)基因组成,Cas基因位于CRISPR序列侧翼^[10],编码Cas蛋白。Cas蛋白具有核酸酶、解旋酶、聚合酶及多核苷酸结合蛋白等功能性结构域^[5],一些Cas蛋白参与间隔区的获得,其他的Cas蛋白提供CRISPR编码的噬菌体抗性并干扰侵略性的基因元件^[5],是宿主抵抗噬菌体侵染的关键因素。若一个CRISPR基因座中的Cas基因失去活性,即使同一染色体上的其它CRISPR序列与Cas基因同时存在,该CRISPR基因座也会丧失抵抗噬菌体及整合间隔区的能力^[3]。以Cas1、Cas2为核心基因,根据Cas蛋白在结构和功能上的差异可将CRISPR-Cas系统分为3大类型,其标志基因分别为Cas3、Cas9 (Csn1)和Cas10 (Csm1)^[11]。

本研究通过生物信息学分析发现同一类型CRISPR-Cas系统的标志Cas基因保守性较高,针对不同类型设计通用引物,对实验室6株未测序嗜热链球菌是否含有CRISPR-Cas系统及CRISPR-Cas系统类型进行预测。该方法成本较低,且能方便快捷地预测实验菌株CRISPR-Cas系统的类型。

1 材料和方法

1.1 材料

1.1.1 菌株与培养基:嗜热链球菌S0、S4、79、883、KLDS3.0206和KLDS3.0207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。

1.1.2 主要试剂和仪器:M17肉汤培养基,北京天根生物技术有限公司;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; 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; 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。

1.2 CRISPR序列的鉴定

在CRISPR Database (<http://crispr.u-psud.fr/crispr/>) 网站上下载嗜热链球菌菌株确定为CRISPR的序列信息,并对其重复序列DRs进行整理。在NCBI上查找并下载CRISPR Database上未收录的嗜热链球菌菌株全基因组序列,利用CRISPR Finder^[12] (<http://crispr.u-psud.fr/Server/>)对CRISPR序列进行查找。这些菌株包括:嗜热链球菌CNRZ1066 (*Streptococcus thermophilus* CNRZ1066, 登录号为NC_006449),嗜热链球菌JIM 8232 (*Streptococcus thermophilus* JIM 8232, NC_017581),嗜热链球菌LMD-9 (*Streptococcus thermophilus* LMD-9, NC_008532),嗜热链球菌LMG 18311 (*Streptococcus thermophilus* LMG 18311, NC_006448),嗜热链球菌MN-ZLW-002 (*Streptococcus thermophilus* MN-ZLW-002, NC_017927),嗜热链球菌ND03 (*Streptococcus thermophilus* ND03, NC_017563),嗜热链球菌ASCC 1275 (*Streptococcus thermophilus* ASCC 1275, NZ_CP006819.1),嗜热链球菌SMQ-301 (*Streptococcus thermophilus* SMQ-301, CP011217),嗜热链球菌MN-BM-A02 (*Streptococcus thermophilus* strain MN-BM-A02, CP010999)。

1.3 Cas基因的预测

在CRISPR序列上下游20 kb范围内查找编码区,利用HUMMER网络服务器下的TIGRFAM蛋白家族数据库(<http://hmmer.janelia.org/search/hmmscan>)^[13]、UniProt蛋白质数据库下的BLAST (<http://www.uniprot.org/blast/>)及NCBI的BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 预测编码区

所编码基因的功能, 查找*Cas*基因。根据*Cas*基因的不同确定CRISPR-*Cas*系统的类型并绘制CRISPR-*Cas*系统结构图, 分类依据参照Kira S. Makarova^[11]。

1.4 不同类型CRISPR-*Cas*系统标志*Cas*基因同源性分析

在NCBI上下载嗜热链球菌不同类型CRISPR-*Cas*系统其标志*Cas*基因的氨基酸序列, 利用BLASTn 2.2.29+建立序列相似性。

1.5 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2代。菌株总DNA的提取参照细菌基因组提取试剂盒说明书。

1.6 *Cas*基因扩增与测序

以NCBI上已公布的9株嗜热链球菌为标准菌株, 根据其*Cas*基因序列的保守性, 利用Primer Premier 6.0软件针对不同类型CRISPR-*Cas*系统的

标志*Cas*基因设计引物, 对实验室未测序嗜热链球菌菌株进行*Cas*基因扩增。PCR产物进行琼脂糖凝胶电泳检测, 并对目的片段进行测序。测序结果利用MEGA 5构建进化树, 用DNAMAN进行多序列比对。

2 结果和分析

2.1 嗜热链球菌的CRISPR-*Cas*系统的一般结构

按照CRISPR-*Cas*系统在染色体上的位置先后将CRISPR-*Cas*系统按CRISPR1-*Cas*、CRISPR2-*Cas*、CRISPR3-*Cas*、CRISPR4-*Cas*排序, 绘制得到9株标准嗜热链球菌菌株的CRISPR-*Cas*系统结构图(图1)。图1-A为CRISPR1-*Cas*系统; 图1-B为St MN-ZLW-002的CRISPR2-*Cas*系统; 图1-C为其余8株菌的CRISPR2-*Cas*系统, 其中St ND03的CRISPR2-*Cas*系统的重复序列为30 bp; 图1-D为St LMD-9、St SMQ-301、St ND03、St MN-ZLW-

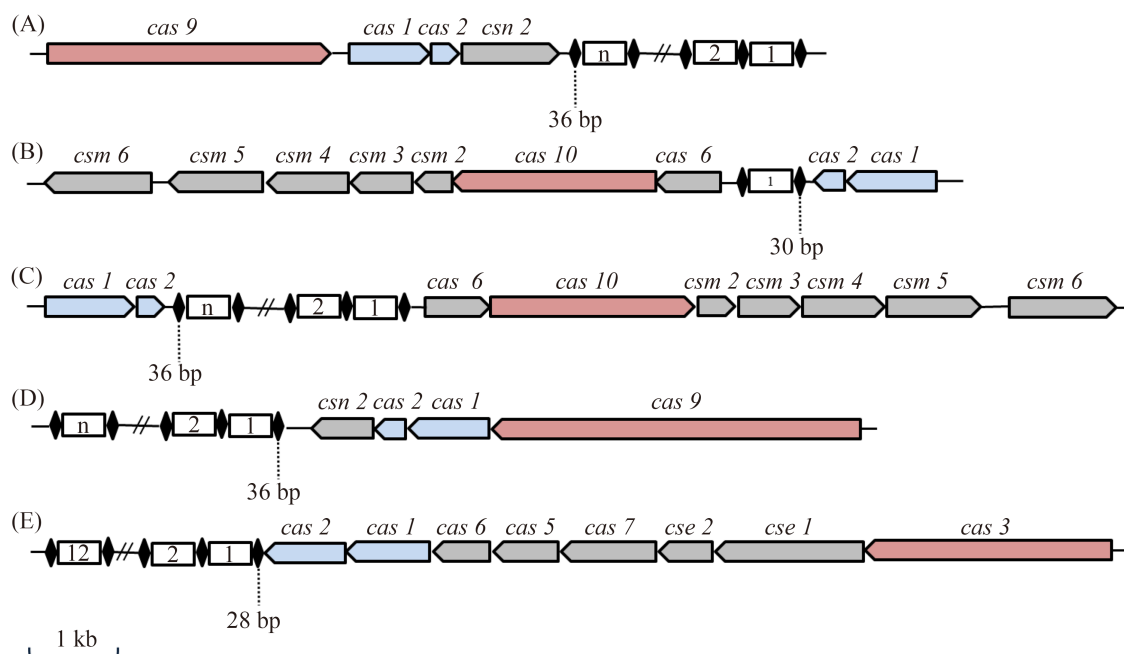


图 1. 嗜热链球菌CRISPR-*Cas*系统结构图

Figure 1. Schematic representation of the CRISPR-*Cas* systems. Genes names according to Kira S. Makarova et al (2011) are indicated above corresponding genes. The black diamonds indicate direct repeats, spacers are numbered in white boxes. The signature genes for CRISPR-*Cas* types are shown with red arrows.

002、St MN-BM-A02和St ASCC 1275的CRISPR3-Cas系统；图1-E为St MN-BM-A02和St ASCC 1275的CRISPR4-Cas系统。

嗜热链球菌CNRZ1066仅含1个II-A型CRISPR1-Cas系统；其他8株嗜热链球菌的CRISPR1-Cas系统亦为II-A型，CRISPR2-Cas系统均为III-A型；嗜热链球菌LMD-9、嗜热链球菌MN-ZLW-002、嗜热链球菌ND03和嗜热链球菌SMQ-301的CRISPR3-Cas系统为II-A型；嗜热链球菌MN-BM-A02和嗜热链球菌ASCC 1275均含4个CRISPR-Cas系统，其CRISPR3-Cas系统为II-A型、CRISPR4-Cas系统为I-E型。

虽然嗜热链球菌CNRZ1066仅有1个CRISPR1-Cas系统，然而在CRISPR1下游240 kb处预测到*Cas1*、*Cas2*、*Csm3*、*Csm4*、*Csm5*、*Csm6*等III-A型CRISPR相关基因。比较分析其他嗜热链球菌的相似基因座发现嗜热链球菌CNRZ1066可能发生交换重组，从而导致在该染色体位点出现一个不含CRISPR序列仅含部分*Cas*基因的区域。对9株嗜热链球菌的不同基因座结构进行详细的序列分析发现其可能通过同源重组事件发生了重复序列的删除或者退化^[14]。

2.2 不同类型CRISPR-Cas系统*Cas*基因的同源性

对标志*Cas*基因进行本地BLAST，结果表明相同类型CRISPR-Cas系统标志*Cas*基因的同源性较高(表1)。其中，CRISPR1-Cas系统和CRISPR3-Cas系统均为II-A型，但前者与后2者的*Cas9*基因大小不同，相似性较低，甚至在蛋白水平的相似性也很低，因此需设计不同的引物。本研究以*Cas9*表示CRISPR1-Cas系统的标志基因，以*Cas9**表示CRISPR3-Cas系统的标志基因。

2.3 *Cas*基因扩增与测序

根据*Cas*基因的保守性，利用Primer Premier 6.0软件在*Cas*基因内部设计引物(表2)。

PCR结果如图2，其中图A、B、C、D分别为

*Cas9*基因、*Cas10*基因、*Cas9**基因和*Cas3*基因电泳图。图2-A中6株嗜热链球菌在500 bp处均有清晰条带；图2-B中S4在300 bp处有条带，但其扩增效率较其它泳道低，且特异性较差，初步判断为干扰带，其它5株均在600 bp处有清晰条带，且扩增效率高、特异性强；图2-C中S0、79、883、KLDS3.0206和KLDS3.0207均在600 bp处有清晰条带；图2-D中仅79和KLDS3.0207在900 bp处有清晰条带。

对目的条带进行测序，测序结果的系统发生学分析如图3。由*Cas9*基因的系统发育树可以看出，供试菌KLDS3.0206、KLDS3.0207、79、883和S0与标准菌株St MN-BM-A02、St LMD-9、St SMQ-301、St ND03及St ASCC 1275位于同一独立遗传分支上，存在密切的同源进化关系。S4与标准菌株St CNRZ1066处于同一独立遗传分支上，说明S4与St CNRZ1066同源进化关系较近。*Cas10*基因测序结果BLAST表明S4在300 bp处的条带为干扰带，未比对上目的基因片段。将其他5株供试菌与标准菌株的*Cas10*基因构建系统发育树，该发育树包括2个分支，其中S0、883、KLDS3.0206和标准菌株St LMG 18311、St ND03、St ASCC 1275、St MN-BM-A02及St MN-ZLW-002位于同一分支上，进化关系较密切，KLDS3.0207和79位于另一分支上，说明KLDS3.0207和79进化关系较近。

*Cas9**基因的进化树主要分为2个类群，供试菌KLDS3.0206、KLDS3.0207、79、883、S0与标准菌株St ASCC 1275、St ND03、St MN-ZLW-002、St MN-BM-A02位于同一分支上，进化关系密切，与另一分支上的St SMQ-301和St LMD-9进化关系较远。利用DNAMAN对*Cas3*基因序列进行多序列比对，结果表明2株供试菌KLDS3.0207、79与2株标准菌株St ASCC 1275、St MN-BM-A02的同源性为97.70%。

6株供试菌中，S4仅含II-A型CRISPR1-Cas系

表1. CRISPR-Cas系统标志Cas基因的同源性
Table 1. Identities of marker *Cas* genes in CRISPR-Cas system

Strains	Identities	Gaps/%	Expect
<i>Cas9</i>			
<i>St.</i> MN-BM-A02	3366/3366 (100%)	0	0.0
<i>St.</i> ASCC 1275	3366/3366 (100%)	0	0.0
<i>St.</i> LMD-9	3361/3366 (99%)	0	0.0
<i>St.</i> JIM 8232	3357/3366 (99%)	0	0.0
<i>St.</i> ND03	3351/3366 (99%)	0	0.0
<i>St.</i> MN-ZLW-002	3347/3366 (99%)	0	0.0
<i>St.</i> CNRZ1066	3322/3388 (98%)	1	0.0
<i>St.</i> LMG 18311	3280/3373 (97%)	0	0.0
<i>Cas9*</i>			
<i>St.</i> SMQ-301	4167/4167 (100%)	0	0.0
<i>St.</i> LMD-9	4167/4167 (100%)	0	0.0
<i>St.</i> MN-ZLW-002	4159/4167 (99%)	0	0.0
<i>St.</i> MN-BM-A02	4157/4167 (99%)	0	0.0
<i>St.</i> ASCC 1275	4157/4167 (99%)	0	0.0
<i>St.</i> ND03	4158/4167 (99%)	0	0.0
<i>Cas10</i>			
<i>St.</i> MN-BM-A02	2277/2277 (100%)	0	0.0
<i>St.</i> ASCC 1275	2277/2277 (100%)	0	0.0
<i>St.</i> ND03	2268/2277 (99%)	0	0.0
<i>St.</i> MN-ZLW-002	2268/2277 (99%)	0	0.0
<i>St.</i> LMG 18311	2263/2277 (99%)	0	0.0
<i>St.</i> LMD-9	2258/2277 (99%)	0	0.0
<i>St.</i> JIM 8232	2176/2287 (95%)	1	0.0
<i>St.</i> SMQ-301	2258/2277 (99%)	0	0.0
<i>Cas3</i>			
<i>St.</i> ASCC 1275	2781/2781 (100%)	0	0.0
<i>St.</i> MN-BM-A02	2780/2781 (99%)	0	0.0

表2. 嗜热链球菌Cas基因的扩增引物
Table 2. *Cas* amplification primers of *Streptococcus thermophilus* strains

Primer name	Primer sequence(5'→3')	Annealing temperature/°C	Size/bp
Cas3	F: CGCTAAGTCGCTTATCAGTACA/R: GAGGATTGAACAGTGGACAGAG	56	904
Cas9	F: CGGACTGATTATGGTCGTTACA/R: CGTCAACTTGCTTCTGGCTAA	54	520
Cas9*	F: TTA CTGTTCCGTACCGCTCTATT/R: TTGTTGAGATGGCTCGTGAA	58	600
Cas10	F: ACCGACATAGAGCATATTAGCC/R: GAGGAGAGTGACGAGGATACA	55	596

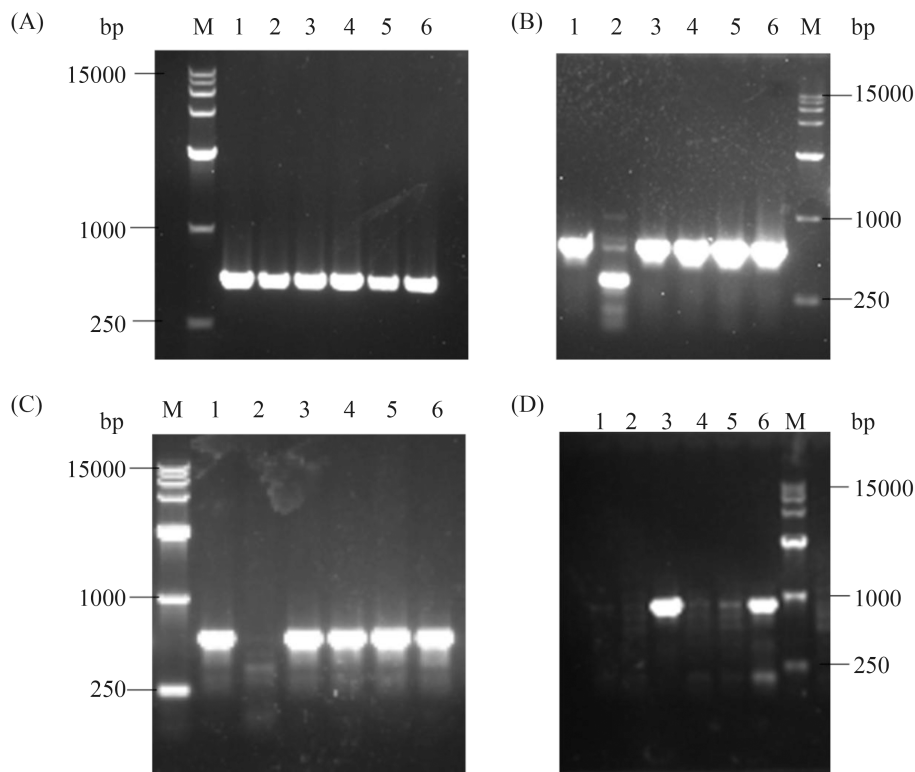


图 2. *Cas*基因扩增结果

Figure 2. Amplification results of *Cas* genes. A: *Cas9* gene; B: *Cas10* gene; C: *Cas9*^{*} gene; D: *Cas3* gene. M: DNA Marker; lane 1: S0; lane 2: S4; lane 3: 79; lane 4: 883; lane 5: KLDS3.0206; lane 6: KLDS3.0207.

统标志基因*Cas9*基因; S0、883及KLDS3.0206均含有 II -A型CRISPR1-Cas系统标志基因*Cas9*基因、III -A型CRISPR2-Cas系统标志基因*Cas10*基因和 II -A型CRISPR3-Cas系统标志基因*Cas9*^{*}基因; 79和KLDS3.0207含有 II -A型CRISPR1-Cas系统标志基因*Cas9*基因、III -A型CRISPR2-Cas系统标志基因*Cas10*基因、II -A型CRISPR3-Cas系统标志基因*Cas9*^{*}基因和 I -E型CRISPR4-Cas系统标志基因*Cas3*基因。该结果与邓凯波等^[15]的研究结果S4仅含1个CRISPR序列、S0含有3个CRISPR序列一致。

3 讨论

S4与标准菌株 CNRZ1066, S0、883、KLDS3.0206与标准菌株MN-ZLW-002、SMQ-301、ND03、LMD-9, KLDS3.0207、79与标准菌

株ASCC 1275、MN-BM-A02均含相同数目和类型的CRISPR-Cas系统。考虑CRISPR-Cas系统类型和数目相同的菌株间可能具有较近的进化关系, 为探究供试菌株的性能提供视角。

CNRZ1066、JIM8232、LMG18311与其他几株嗜热链球菌的全基因组高度保守, 但前三者却不含 II -A型CRISPR3-Cas系统, 在前三者中检测到退化的CRISPR3重复序列。在其他几株菌的*Cas9*与其下游的磷酸丝氨酸磷酸酶之间也检测到一个退化的重复序列。有趣的是, 基因组序列在重复序列的上游, 却位于不完全重复序列的下游, 这表明在末端重复序列与不完全重复序列之间可能发生了重组事件, 从而导致整个CRISPR3相关*Cas*基因片段的缺失或插入。在 I -E型和 II -A型CRISPR-Cas系统中, CRISPR序列通常位于*Cas*基因的上游或下游, 而在 III -A型CRISPR-

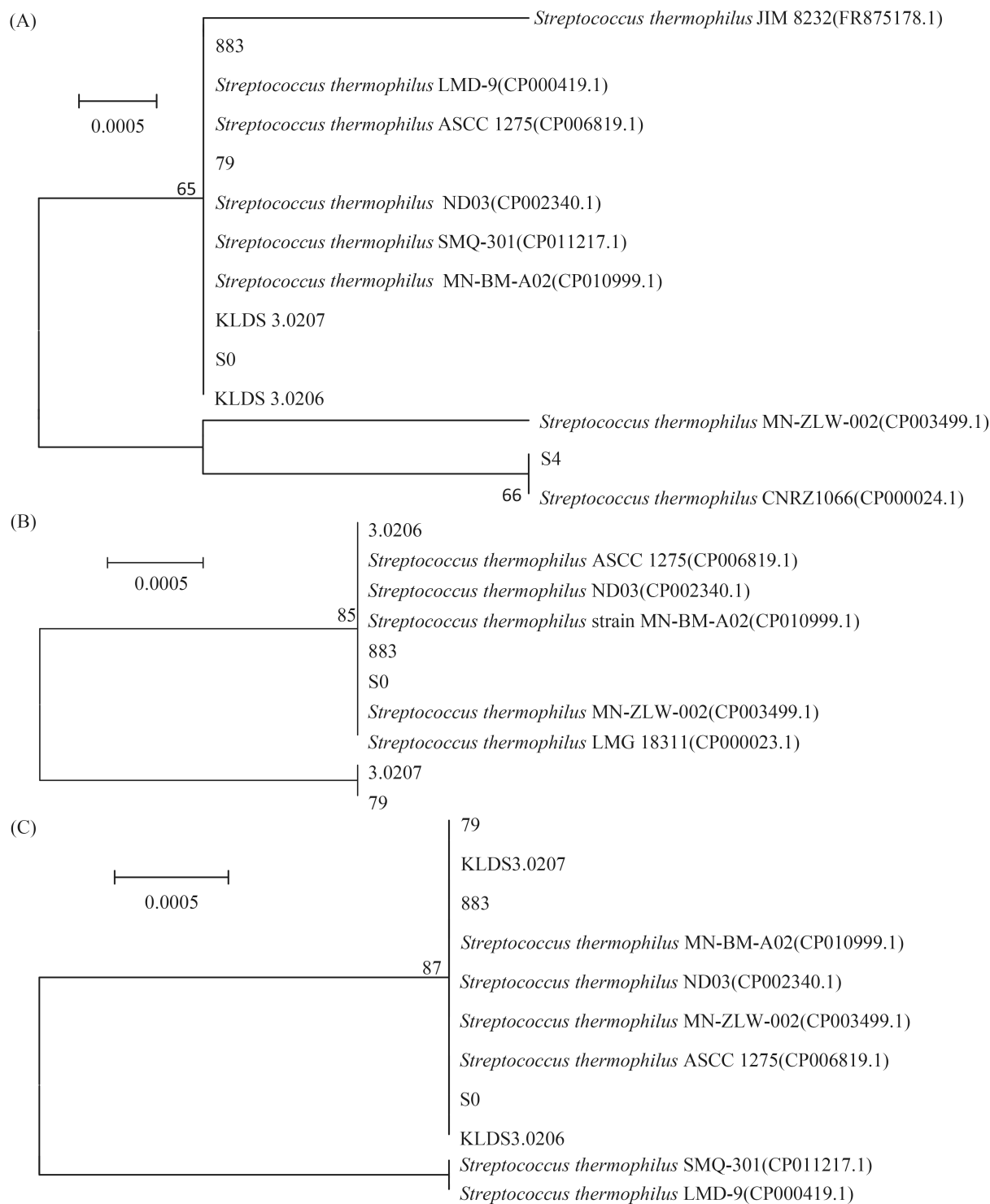


图 3. *Cas*基因系统发育树

Figure 3. Phylogenetic tree of *Cas* genes. A: Phylogenetic tree of *Cas9* gene; B: Phylogenetic tree of *Cas10* gene; C: Phylogenetic tree of *Cas9** gene. The numbers of nodes indicate bootstrap values, and represent the percentages of 1000 bootstrap replications in which the taxa to the right are placed together. Numbers in parentheses represent the sequences accession number in GenBank. The scale bar represents the horizontal branch lengths and the number of substitutions per nucleotide position.

Cas系统中, CRISPR序列位于Cas基因之间, 将Cas1、Cas2与其他Cas基因隔开, 这可能与其特定的作用机制有关。

9株嗜热链球菌的CRISPR-Cas系统的重复-间隔单元最少为1个, 最多42个。截止至2014年8月, CRISPRs database(<http://crispr.u-psud.fr/crispr/>)中所记载的重复-间隔单元最高可达374个(*Chloroflexus* sp. Y-400-fl), 但大多数基因座通常少于50个重复-间隔单元, 这在乳酸菌基因组中得到例证^[16]。9株标准菌株的重复序列在同类型CRISPR-Cas系统中高度保守, 因此重复序列被当做CRISPR-Cas系统的分型依据之一^[17]。间隔序列则是高度可变的, 不同CRISPR-Cas系统的间隔序列均不相同。将间隔序列进行BLAST, 由比对结果知间隔序列可能来自噬菌体、质粒等外源基因元件。研究表明间隔序列的排列顺序可反映噬菌体对宿主的侵染历史。

尽管直观看来CRISPR基因座不可能无限扩大, 但定义CRISPR基因座最佳尺寸及最大尺寸的参数还未可知。此外, 宿主CRISPR扩大的抗性代价应与病毒类群中的CRISPR逃逸相比较, 以确定在这场基因战争中是捕食者还是被捕食者承担了较高的进化代价。CRISPR基因座最初是通过噬菌体感染整合新的间隔区而得到进化, 其末端间隔区发生删除, 这可能是由于CRISPR重复序列之间的同源重组造成的^[2,14,18]。这或许使得宿主能够限制CRISPR基因座的扩大, 使其不会增大到不利的水平。

9株标准嗜热链球菌的CRISPR上下游序列保守性较差, 扩增CRISPR序列的实验较繁琐。已证实Cas基因是CRISPR-Cas系统抵抗噬菌体侵染所不可或缺的, 且Cas基因同源性较高, 易于设计引物。本研究针对不同类型CRISPR-Cas系统的标志Cas基因设计引物, 扩增得到实验室未测序菌株的Cas基因, 并能直接推测供试菌所含CRISPR-

Cas系统的类型, 实验较为有效、简易, 为进一步探究嗜热链球菌中CRISPR-Cas系统的功能奠定基础。

参考文献

- [1] Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goltsman E, Mazur M, Pusch GD, Fonstein M, Overbeek R, Kyprides N, Purnelle B, Prozzi D, Ngui K, Masuy D, Hancy F, Burtreau S, Boutry M, Delcour J, Goffeau A, Hols P. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nature Biotechnology*, 2004, 22(12): 1554–1558.
- [2] Jansen R, van Embden JDA, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 2002, 43(6): 1565–1575.
- [3] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 2007, 315(5819): 1709–1712.
- [4] van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences*, 2009, 34(8): 401–407.
- [5] Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science*, 2010, 327(5962): 167–170.
- [6] Deveau H, Garneau JE, Moineau S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annual Review of Microbiology*, 2010, 64: 475–493.
- [7] Young JC, Dill BD, Pan CL, Hettich RL, Banfield JF, Shah M, Fremaux C, Horvath P, Barrangou R, VerBerkmoes NC. Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS One*, 2012, 7(5): e38077.
- [8] Sun CL, Barrangou R, Thomas BC, Horvath P, Fremaux C, Banfield JF. Phage mutations in response to CRISPR diversification in a bacterial population. *Environmental Microbiology*, 2013, 15(2): 463–470.
- [9] 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.
- [10] Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology*, 2005, 1(6): e60.

- [11] 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.
- [12] Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, 2007, 35(S2): W52–W57.
- [13] Selengut JD, Haft DH, Davidsen T, Ganapathy A, Gwinn-Giglio M, Nelson WC, Richter AR, White O. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Research*, 2007, 35(S1): D260–D264.
- [14] Horvath P, Romero DA, Coûté-Monvoisin AC, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *Journal of Bacteriology*, 2008, 190(4): 1401–1412.
- [15] Deng KB, Huo GC. Detection and homology analysis of CRISPR in *Streptococcus thermophilus*. *Food Science*, 2013, 34(3): 153–157. (in Chinese)
- [16] 邓凯波, 霍贵成. 嗜热链球菌中CRISPR序列的检测与同源性分析. *食品科学*, 2013, 34(3): 153–157.
- [17] Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *International Journal of Food Microbiology*, 2009, 131(1): 62–70.
- [18] Lange SJ, Alkhnabashi OS, Rose D, Will S, Backofen R. CRISPRmap: an automated classification of repeat conservation in prokaryotic adaptive immune systems. *Nucleic Acids Research*, 2013, 41(17): 8034–8044.
- [19] Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology*, 2008, 190(4): 1390–1400.

Detection of CRSPR-Cas system in *Streptococcus thermophilus*

Wan Li, Hongzhang Liang, Danqing Zhang, Nana Wang, Yaru Tang, Bailiang Li, Guicheng Huo*

Key Laboratory of Dairy Science, Ministry of Education, Northeast Agricultural University, Harbin 150030, Heilongjiang Province, China

Abstract: [Objective] We aimed to detect the CRSPR-Cas system of six *Streptococcus thermophilus*. [Methods] Bioinformatics method was used to predict CRSPR-Cas system of nine *S. thermophilus* that published in National Center for Biotechnology Information. Four primers were designed according to the flanking sequences of standard strains and the CRISPR-Cas system of six *S. thermophilus* have been detected by PCR method. [Results] *S. thermophilus* S4 had a *Cas9* gene, others all had *Cas9* gene, *Cas10* gene and *Cas9** gene. In addition, 79 and KLDS3.0207 still had *Cas3* gene. [Conclusion] Signature genes amplification of CRSPR-Cas system could predict the type of CRSPR-Cas system in unsequenced strains, these findings will help establish the foundation for the study of CRSPR-Cas system in lactic acid bacteria.

Keywords: *Streptococcus thermophilus*, clustered regularly interspaced short palindromic repeats(CRISPR), CRSPR-Cas system, *Cas* genes, bioinformatics method

(本文责编: 张晓丽)

Supported by the National Natural Science Foundation of China (31171717) and by the National Programs for High Technology Research and Development of China (2012AA022108)

*Corresponding author. Tel: +86-451-55191807; Fax: +86-451-55190340; E-mail: gchuo@vip.0451.com

Received: 20 July 2015; Revised: 9 September 2015; Published online: 15 September 2015