

兽疫链球菌中与透明质酸合成相关新基因的高通量筛选及克隆

毕亚丽 王震 王璐 刘浩*

(工业发酵微生物教育部重点实验室 天津科技大学生物工程学院 天津 300457)

摘要:【目的】兽疫链球菌(*Streptococcus equi* subsp. *zooepidemicus*)中透明质酸主要的生物合成途径和相关基因已经被研究得比较透彻, 探究一种挖掘与透明质酸合成相关新基因的策略。【方法】利用自杀质粒 pSET4s::*sacB* 在宿主基因组中的随机整合作用, 筛选具有表型差异突变菌株构建突变体库, 进一步利用连接介导 PCR (Ligation mediated PCR, LM-PCR)方法和全基因组重测序, 检测质粒整合位点, 通过基因无痕敲除和回补实验验证插入位点。【结果】构建了包含 150 株具有表型差异突变株的突变体库; 以荚膜合成能力缺失的 1 号突变株(M1)作为基础研究对象, 检测到自杀质粒整合到基因组 458 960 位点上, 破坏了编码塔格糖-6-磷酸激酶的 *lacC* 基因; 无痕敲除 *lacC* 基因得到 $\Delta lacC$, 表型分析发现 $\Delta lacC$ 表现为粘性荚膜特性; 进一步全基因组重测序发现, 除了 *lacC* 基因位点存在插入突变, 206 613 位点存在碱基 G 缺失, 导致编码透明质酸合成酶的 *hasA* 基因发生移码突变, 且回补 *hasA* 基因后, M1 恢复粘性荚膜合成能力。【结论】M1 突变株粘性荚膜合成能力的缺失由 *hasA* 基因功能缺失引起, 与 *lacC* 基因功能缺失无关。初步建立了兽疫链球菌中高通量筛选与透明质酸合成相关新基因的策略, 为今后挖掘新基因奠定了基础。

关键词: 透明质酸, 兽疫链球菌, 突变体库, 整合位点, 基因敲除

High throughput screening and cloning of new genes related to synthesis of hyaluronic acid in *Streptococcus zooepidemicus*

BI Ya-Li WANG Zhen WANG Lu LIU Hao*

(MOE Key Laboratory of Industrial Fermentation Microbiology, College of Biotechnology, Tianjin University of Science & Technology, Tianjin 300457, China)

Abstract: [Objective] Since hyaluronic acid (HA) biosynthesis pathway and the related genes in *Streptococcus equi* subsp. *zooepidemicus* have been thoroughly studied, it is of great significance of exploring a kind of strategy to identify new genes related to the synthesis of HA. [Methods] Strains phenotype defect were screened using the randomly incorporation of the suicide vector pSET4s::*sacB* in host genomes and a mutant library was constructed. Integration loci were detected through ligation

基金项目: 科技部国际合作项目(No. 2011DFA31280); 天津市国际科技合作项目(No. 13RCGFSY19400)

*通讯作者: Tel: 86-22-60600127; 信箱: liuhao@tust.edu.cn

收稿日期: 2015-03-13; 接受日期: 2015-05-11; 优先数字出版日期(www.cnki.net): 2015-06-02

mediated PCR (LM-PCR) and verified by gene deletion and complementation, as well as complete genomic sequencing. **[Results]** A mutant library containing 150 mutants strains phenotype defect was constructed. We detected that the temperature sensitive vector was incorporating at 458 960 loci in *Streptococcus zooepidemicus* genomes and disrupting gene *lacC* which encodes tagatose-6-phosphate kinase. $\Delta lacC$ was obtained through markerless gene deletion system, and the phenotype analysis showed that $\Delta lacC$ still displayed the mucoid phenotype that was not the same as M1. The further complete genomic sequencing showed that base G in the locus of 206 613 in M1 mutant genome was lost, which was located in *hasA* gene encoding hyaluronate synthase and leading to frameshift mutation of *hasA*. What's more, when complementation of *hasA* in M1 mutant, interestingly, M1 mutant rescued the capsule synthesis ability and displayed mucoid phenotype. **[Conclusion]** The result indicated that the loss of capsule synthesis ability of M1 mutant was the result of the loss of *hasA* gene function, which had nothing to do with the lack of *lacC* gene function. This study preliminarily established the strategies to high throughput screening new genes related to synthesis of HA in *Streptococcus zooepidemicus*, which laid a foundation for mining new genes in future.

Keywords: Hyaluronic acid, *Streptococcus zooepidemicus*, Mutant library, Integration loci, Gene deletion

透明质酸(Hyaluronic acid, HA), 又称玻璃酸, 因其具有独特理化性质和生理功能, 在医药领域、化妆品领域和食品领域已被广泛的应用^[1-3]。目前全球对于 HA 的市场需求价值很大, 日本和美国均达 3 亿美元, 而欧洲市场在 1 亿美元左右^[4]。工业上主要通过兽疫链球菌发酵法生产透明质酸^[5]。近年来, 几株兽疫链球菌的全基因组测序工作相继完成并公布^[6-7], 这极大地丰富了我们对其生物学和遗传学信息的认识。随着研究的深入, 兽疫链球菌中 HA 的主要合成途径已被研究得比较透彻^[8], 探究一种挖掘与透明质酸合成相关的新基因的方法和策略, 挖掘新基因, 完善 HA 代谢网络, 并通过改造新基因构建 HA 高产菌株具有重要的意义。

突变体库的构建是功能基因研究的重要内容, 构建突变体库有多种方法^[9], 插入突变具有更方便地进行正向和反向遗传学研究的优点, 是突变体库和功能基因组学研究的首选方法^[10]。其中质粒介导的插入突变已被广泛应用于研究微生物功能基因组学, 其原理是利用自杀型质粒不能在宿主中自主复制, 因而通过抗生素能筛选到插入整合的突变菌株。Ogasawara 等利用质粒 pMUTIN 随机整合到宿主基因组, 构建 2 000 多个枯草芽孢杆菌突变体, 在研究枯草芽孢杆菌基因组方面取得较好结果^[11]。

通过正向遗传学克隆表型缺陷突变体中发生突变的基因并揭示基因功能, 是研究基因组学最简单有效的途径。迅速检测外源基因的整合位点, 是研究插入突变最重要的环节。目前检测外源基因整合位点的方法主要包括 DNA 杂交、PCR 法、Southern 印迹法等, 其中 PCR 扩增法最为简单有效, 为此也建立多种 PCR 技术, 如反向 PCR (Inverse PCR, I-PCR)和连接介导 PCR (Ligation mediated PCR, LM-PCR)^[12]。

基因敲除技术在基因功能研究方面发挥着重要的作用, 我们之前的研究利用二次同源重组和 *sacB* 编码蔗糖果聚糖酶使兽疫链球菌在含有 5%蔗糖的 THY 培养基中的致死原理作为反向筛选标记, 建立了一种简便而又高效的基因无痕打靶系统^[13]。利用该系统可以快速打靶外源基因整合位点的侧翼序列, 缺失目的基因或使基因功能屏蔽, 这是验证基因突变和表型因果关系的最有效途径。同时, 在基因功能缺失的突变株中进行基因回补实验, 通过基因回补后的表型恢复或基因功能恢复验证该基因的功能和表型相关性的有效途径。

微生物基因组重测序也是研究基因组遗传变异的有效方法。它通过对个体或群体的全基因组测序, 然后与已有物种的参考基因组序列进行比对、

拼装,再利用高性能计算平台和生物信息学手段分析单核苷酸多态位点(SNP)、结构变异(SV)和插入缺失(InDel)等多态信息,从而获得该生物的遗传变异信息。

本研究第一次提出利用温敏型质粒随机插入突变构建兽疫链球菌突变体库的研究策略,利用质粒介导的 LM-PCR 法克隆温敏型质粒整合位点,通过 *lacC* 基因无痕敲除和 *hasA* 基因回补实验验证基因功能与表型的相关性,建立了探究与透明质酸合成相关新基因的方法和策略,为今后挖掘新基因奠定了基础。

1 材料与amp;方法

1.1 材料

1.1.1 菌株和质粒:所用的质粒和菌株如表 1 所示,大肠杆菌 JM109 在 Luria-Bertani (LB)培养基中 37 °C、200 r/min 培养^[13]。兽疫链球菌在 THY 培养基中培养^[13]。使用的抗生素浓度如下:对于大肠杆菌,壮观霉素(50 mg/L)和氨苄霉素(100 mg/L);对于兽疫链球菌,壮观霉素(100 mg/L)和氯霉素(5 mg/L)。

1.1.2 工具酶、试剂盒和其他试剂:限制性内切酶、T4-DNA 连接酶、*Taq* 酶,购自 TaKaRa 公司;PCR 纯化试剂盒、琼脂糖凝胶回收试剂盒、质粒小提试剂盒,购自北京天根生化公司;壮观霉素和氨苄霉素购自上海生工生物工程有限公司。

1.2 方法

1.2.1 突变体库的构建:在 37 °C 条件下,温度敏感型质粒 pSET4s::*sacB* 不能在兽疫链球菌中正常复制,存在抗生素抗性压力时,质粒随机插入到宿主基因组中。按照标准步骤^[14]将温度敏感型质粒 pSET4s::*sacB* 电转化到兽疫链球菌 ATCC39920 感受态细胞中,得到 *S. zooepidemicus*/pSET4s::*sacB* 菌株,将此菌株接种于含壮观霉素的 THY 液体培养基中,37 °C、200 r/min 培养 12 h,稀释至适宜浓度后涂布于含壮观霉素抗性的 THY 固体平板上,37 °C 静置培养 24 h,筛选与亲代菌株表型不一致的突变菌株。

1.2.2 插入位点的检测:所用引物如表 2 所示,酶切位点用下划线标注。选取突变株 M1 基因组为试验材料,分别用 5 种限制性内切酶(*Sal* I, *Pst* I, *Sph* I, *Eco*R I, *Hind* III) 37 °C 过夜酶切。采用苯酚-氯仿法纯化回收酶切的基因组片段,将基因组片段与对应限制性内切酶酶切,经过去磷酸化处理的 pUC19 质粒片段 16 °C 过夜连接。连接产物转化大肠杆菌 JM109 感受态细胞,涂布于含有氨苄霉素抗生素的 LB 固体平板上,37 °C 培养 12 h,挑取克隆用质粒上的引物 P1/P3, P1/P4, P2/P3, P2/P4 (P1、P2 为温敏型, P3、P4 为 pUC19)进行菌落 PCR。PCR 扩增反应体系(50 μL):ddH₂O 41 μL, 10×*Taq* reaction buffer 5 μL, 正、反向引物(10 μmol/L)各 1 μL, dNTPs (10 mmol/L) 1 μL, *Taq* 酶(5 U/μL) 1 μL,模板为菌体,下同。反应条件:94 °C 5 min; 94 °C 30 s, 55 °C 30 s, 72 °C 3 min, 30 个循环; 72 °C 10 min; 4 °C 保存。PCR 反应结束后,取 5 μL 产物进行琼脂糖凝胶电泳检测。将扩增出特异性单一条带的单克隆提取质粒,送交华大基因公司测序。将测得的结果与温度敏感型载体的已知序列比对,找出整合位点处基因组序列,并通过 BLAST 方法将其与 NCBI 中提交的兽疫链球菌基因组序列比对,找到插入位点。

1.2.3 *lacC* 基因的敲除:包括 *lacC* 基因敲除载体的构建和 *lacC* 基因的无痕敲除。(1) *lacC* 基因敲除载体的构建: *lacC* 基因敲除载体 pSET4s::*sacB*::*lacCLR* 的构建是通过重叠延伸 PCR 技术^[15] (Gene splicing by overlap extension PCR, SOE PCR)。以兽疫链球菌 ATCC39920 基因组 DNA 为模板,以 *lacCL01/lacCL02* 和 *lacCR01/lacCR02* 为引物,分别扩增 *lacC* 基因的上游片段和下游片段。两个片段通过重叠 PCR 的方法连接在一起形成大小为 1 488 bp 的 *lacCLR* 片段, *lacCLR* 经限制性内切酶酶切回收后,连接到 pSET4s::*sacB* 质粒的 *Pst* I/*Eco*R I 酶切位点,构建 *lacC* 基因敲除载体 pSET4s::*sacB*::*lacCLR*。(2) *lacC* 基因的无痕敲除:

表 1 研究所用的菌株和质粒
Table 1 Strains and plasmids used in this study

菌株和质粒 Strains or plasmids	特征 Characteristics	来源 Sources
菌株 Strains		
<i>E. coli</i> JM109	Host for cloning	Gibco BRL
<i>S. zooepidemicus</i> ATCC39920	Wild-type strain	ATCC
M1	<i>S. zooepidemicus</i> mutant	This work
$\Delta lacC$	<i>lacC</i> -deficient strain	This work
M1/pLH200::P _{pgk} :: <i>hasA</i>	M1 mutant harboring pLH200::P _{pgk} :: <i>hasA</i>	This work
质粒 Plasmids		
pUC19	Vector for routine cloning	This work
pSET4s:: <i>sacB</i>	pSET4s harboring <i>sacB</i> driven by P _{cm} promoter, spc ^r	[13]
pLH200::P _{pgk} :: <i>hasA</i>	pLH200 harboring <i>hasA</i> driven by P _{pgk} promoter, spc ^r	[13]
pSET4s:: <i>sacB</i> :: <i>lacCLR</i>	pSET4s:: <i>sacB</i> harboring the 5' and 3' flanking sequences of <i>lacC</i> , spc ^r	This work

表 2 研究所用的引物
Table 2 Primers used in this study

引物 Primers	序列 Sequences (5'→3')	下划线 Underlined
用于检测 pSET4s:: <i>sacB</i> 质粒插入位点 Primers used for detection of insertion site		
P1	AAGGCGATTAAGTTGGGTAA	
P3	ATACCGCACAGATGCGTAAG	
P4	CGTTGGCCGATTCATTAATG	
用于构建 <i>lacC</i> 基因敲除质粒 Primers used for construction of the vectors for <i>lacC</i> gene deletion		
lacCL01	A <u>ACTGCAGG</u> TGGCCACAATAATGCTCGT	<i>Pst</i> I
lacCL02	TGTTTCTTGTTAATGTCATCATGCTCTCCTAATCCTGGTA	Fusion overlap
lacCR01	TACCAGGATTAGGAGAGCATGATGACATTAACAAGAAACA	Fusion overlap
lacCR02	CG <u>GAAATTC</u> AACCAAGGTCTCTTGAACA	<i>EcoR</i> I
用于验证 <i>lacC</i> 基因敲除 Primers used for examining double crossover events of <i>lacC</i> gene		
lacC-P1	GCTTTGATGTGACAGACGTG	
lacC-P4	TCGCTCAAGCACCTGATTGA	
用于构建 <i>hasA</i> 基因回补质粒 Primers used for construction of vectors of <i>hasA</i> gene complementation		
Ppgk01	AC <u>ATGCATG</u> CATTGCGTGATGTCATGACT	<i>Sph</i> I
Ppgk02	AGGTTTTTTAATGTTCTCATAGTCAATTTAGCCATTTCAT	Fusion overlap
hasA01	ATGAAATGGCTAAATTGACTATGAGAACATTAACAAAAACCT	Fusion overlap
hasA02	ACGCGT <u>TCGACT</u> TATAATAATTTTTTACGTGTTT	<i>Sal</i> I

将 *lacC* 基因敲除载体 pSET4s::*sacB*::*lacCLR* 电转化到兽疫链球菌 ATCC39920 感受态细胞中, 涂布于含有壮观霉素(100 mg/L)的 THY 培养基上, 30 °C 培养至长出单菌落。将壮观霉素抗性转化子在 THY 液体培养基中 30 °C、200 r/min 培养 12 h, 然后将其移至 37 °C 培养 4 h。培养物适当稀释后涂布于含有 5% (质量体积比) 蔗糖的 THY 固体培养基上, 37 °C 培养 24 h。筛选蔗糖抗性和壮观霉素敏感的克隆, 用引物 lacC-P1/lacC-P4 验证 *lacC* 基因敲除。

1.2.4 全基因组重测序: 提取 M1 突变株基因组送交华大基因有限公司进行全基因组重测序, 重测序完成后获取原始数据(Raw reads), 经过滤去污染, 去接头, 然后与 NCBI 公布的兽疫链球菌基因组序列比对, 去除因 PCR 扩展过多导致的重复 Paired-end Reads。根据序列比对结果计算出相对于参考基因组的覆盖度和测序深度、全基因变异中的单核苷酸多态性位点(SNP), 插入/缺失位点(Insertion/Deletion, InDels)等。

1.2.5 *hasA* 基因回补: 包括 *hasA* 基因回补质粒的构建及 *hasA* 基因回补。(1) *hasA* 基因回补质粒的构建: 以兽疫链球菌 ATCC39920 基因组 DNA 为模版, 以 Ppgk01/Ppgk02 为引物, 扩增包含 *pgk* 基因启动子序列的片段; 以 hasA01/hasA02 为引物, 扩增包含 *hasA* 基因开放阅读框的片段。两个片段通过重

叠 PCR 的方法连接在一起形成大小为 1 469 bp 的 P_{pgk} -*hasA* 片段, 经限制性内切酶酶切回收后, 连接到 pSET1 质粒的 *Sph* I/*Sal* I 酶切位点, 构建 *hasA* 基因回补质粒 pSET1:: P_{pgk} ::*hasA*。(2) *hasA* 基因回补: 将 pSET1:: P_{pgk} ::*hasA* 质粒电转化到 M1 突变株中, 涂布于含有氯霉素(5 mg/L)的 THY 平板上, 筛选出氯霉素抗性克隆进一步分析。

2 结果与分析

2.1 插入突变构建突变体库

兽疫链球菌野生型菌株菌落形态较大、湿润且光亮。温度敏感型质粒随机整合到兽疫链球菌基因组中, 插入破坏透明质酸合成的关键基因会导致荚膜缺陷, 表现为菌落形态小、干燥且没有光泽(图 1)。通过高通量筛选, 获得由 150 株与兽疫链球菌亲代菌株表型差异明显的突变株构成的突变体库, 突变体分别编号为 M1-M150, 这些突变株表型差异是由插入突变还是随机自发突变引起的, 还需进一步研究。我们首先选取荚膜合成能力缺失的 1 号突变株(M1)作为基础研究对象, 探究引起兽疫链球菌野生型菌株荚膜合成能力缺失的原因, 从而探究一种挖掘新基因的策略。

2.2 pSET4s::*sacB* 质粒插入位点分析

2.2.1 插入位点 PCR 扩增检测结果: M1 突变株的

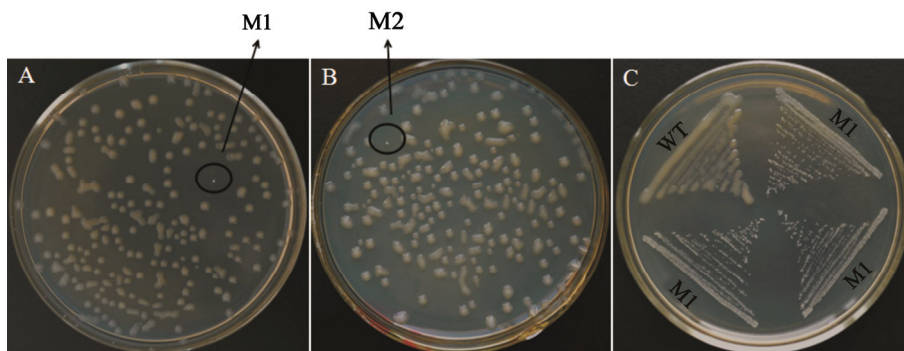


图 1 突变体的筛选及 M1 突变体与野生型菌株表型分析

Figure 1 Selection of mutant strains and assay of the phenotypes between M1 and wild type strains

注: A、B: 突变体的筛选图, A 图○内为 M1 突变体; B 图○内为 M2 突变体。C: M1 突变体与野生型菌株表型分析图, WT: 野生型菌株; M1: 1 号突变株。

Note: A, B: Screening mutants on THY plate, A: ○ indicate M1 mutant; B: ○ indicate M2 Mutant. C: Assay of the phenotypes between M1 and wild type strains, WT: Wild type strains; M1: No.1 mutant.

基因组经 *Sph* I 酶切, 与 pUC19 质粒连接转化后, 以 P1/P3, P1/P4, P2/P4 引物进行菌落 PCR 未扩增出条带, 用 P2/P3 引物扩增出单一的特异性条带, 条带大小在 250–500 bp (图 2)。基因组经 *Sal* I、*Pst* I、*Eco* R I、*Hind* III 酶切连接后菌落 PCR 均未扩增出特异性条带。

2.2.2 克隆测序序列比对: 将 2.2.1 扩增出特异性条带的质粒测序, 将测序结果与 pSET4s::*sacB* 质粒序列比对。结果表明: 其中 226 bp 为已知 pSET4s::*sacB* 质粒的序列, 238 bp 为未知序列(图 3)。将该未知序列与 NCBI 网站中的兽疫链球菌 ATCC35246 菌株的基因组比对后发现, 克隆到的未知序列与兽疫链球菌 ATCC35246 基因组 458 723–458 960 之间的序列一致性为 100%, 此位点位于编码塔格糖-6-磷酸激酶的 *lacC* 基因内部。通过编码的蛋白比对发现, 已经破坏了 *lacC* 基因的编码框。

2.3 *lacC* 基因敲除及敲除菌株表型分析

将在含有 5% 蔗糖的 THY 固体平板上长出的单菌落接种于 THY 液体培养基培养 12 h 后, 提取基

因组, 以 *lacC*-P1/*lacC*-P4 为引物进行 PCR 扩增, 如果 *lacC* 基因被敲除, 则能扩增出比对照组小 933 bp 的条带(图 4), 结果表明, 筛选到的 4 株菌均能扩增出比野生型小 900 bp 左右的条带, 可能为 *lacC* 敲除菌。进一步通过对扩增得到的条带进行纯化回收及测序, 分析发现扩增得到的基因序列与野生型之间的序列一致性为 100%。结果表明, 获得 4 株 *lacC* 基因敲除菌株。

将兽疫链球菌野生型菌株、M1 突变株和 $\Delta lacC$ 突变株接种于 THY 液体培养基中, 37 °C、200 r/min 培养 8 h (OD_{660} 约为 0.6), 将 3 株菌划线于 THB 固体平板上, 37 °C 静置培养 24 h 后, 观察表型(图 5), 结果表明, $\Delta lacC$ 突变株与野生型菌株表型一致, 没有缺失荚膜合成能力, M1 突变株的荚膜缺陷不是由 *lacC* 基因功能的缺失引起的, 可能是自发突变也可能有其他位点插入, 所以进一步对该菌株进行重测序。

2.4 M1 突变株全基因组重测序

兽疫链球菌基因组大小仅为 2.17 Mb, 对 M1 进行全基因组重测序, 得到约 11 G 高质量的数据

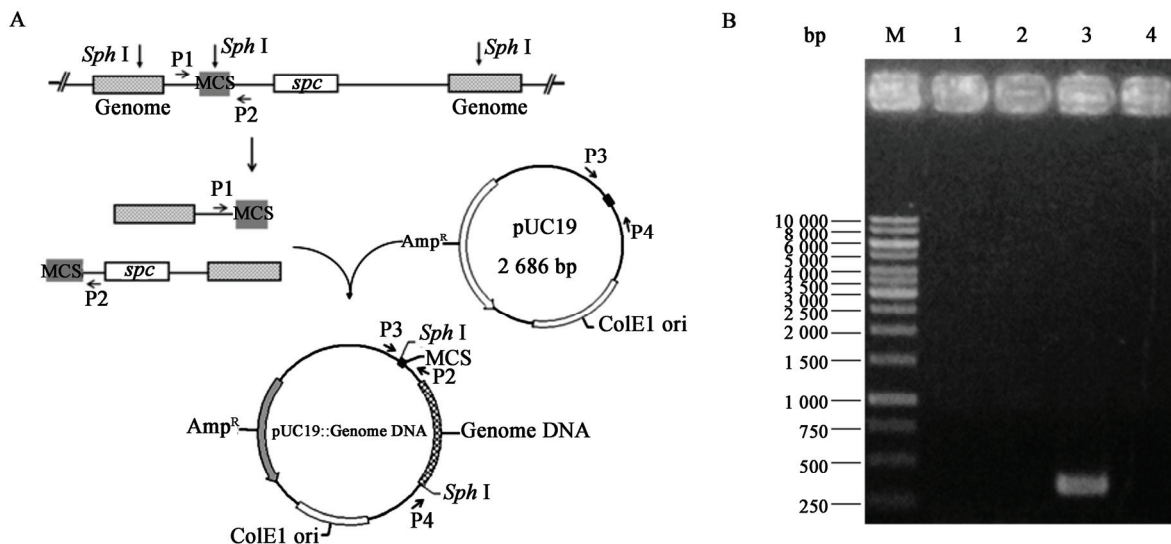


图 2 pSET4s::*sacB* 质粒插入位点检测

Figure 2 Detection of the insertion position of pSET4s::*sacB* plasmid

注: A: pSET4s::*sacB* 质粒插入位点检测方法示意图; B: pSET4s::*sacB* 质粒插入位点检测 PCR 扩增图。

Note: A: Detection method of the insertion position of pSET4s::*sacB*; B: PCR amplification of the insertion position of pSET4s::*sacB*. M: 1 kb DNA marker, 1–4: PCR amplification using the primers P1/P3, P1/P4, P2/P3, P2/P4, respectively.

```

1 GCCGTTTGTAACTCAGCCGTGCGTTTTTTATCGCTTTGCAGAAGTTTTTACTTTCTTG 60
2 GCCGTTTGTAACTCAGCCGTGCGTTTTTTATCGCTTTGCAGAAGTTTTTACTTTCTTG 60
1 ACGGAAGAATGATGTGCTTTTGGCCATAGTATGCTTTGTTAAATAAAGATTCTTCGCCTTG 120
2 ACGGAAGAATGATGTGCTTTTGGCCATAGTATGCTTTGTTAAATAAAGATTCTTCGCCTTG 120
1 GTAGCCATCTTCAGTTCAGTGTGTTGCTTCAAATACTAAGTATTTGTGGCCTTTATCTTC 180
2 GTAGCCATCTTCAGTTCAGTGTGTTGCTTCAAATACTAAGTATTTGTGGCCTTTATCTTC 180
1 TACGTAGTGAGGATCTCTCAGCGTATGGTTGTGCGCTGAGCTGTAGTTGCCTTCAAGAAC 240
2 TACGTAGTGAGGATCTCTCAGCGTATGGTTGTGCGCTGAGCTGTAG..... 226
1 TGCCCTTAAGGGGGCCCCGAACAGTCCAAGACAGTCAAGACACCCGTAGGCTTTAGCATT 300
2 ..... 226
1 TGCAATCAGCCTTTGATAGTAATCAGGCTCTAGGCCTCTTGGCAGACTTCCAGAAAGCGT 360
2 ..... 226
1 CATAACAGCACATGACTCAGCCAGCATACTTAAGTGTGCCATAAGGCCTGCTGCCTCCTC 420
2 ..... 226
1 AGAAGCAATCAAAGGACCTGCCTCCAAAAGCTCAGTCTGCATGC 464
2 ..... 226
    
```

图3 测序序列与自杀质粒序列比对结果

Figure 3 Result of sequencing sequence and pSET4s::*sacB* plasmid sequence BLAST

注: 1: 包含基因组系列的重组质粒测序序列; 2: pSET4s::*sacB* 质粒序列.

Note: 1: Sequencing sequence of the recombinant plasmid containing the genome; 2: Sequence of the vector pSET4s::*sacB*.

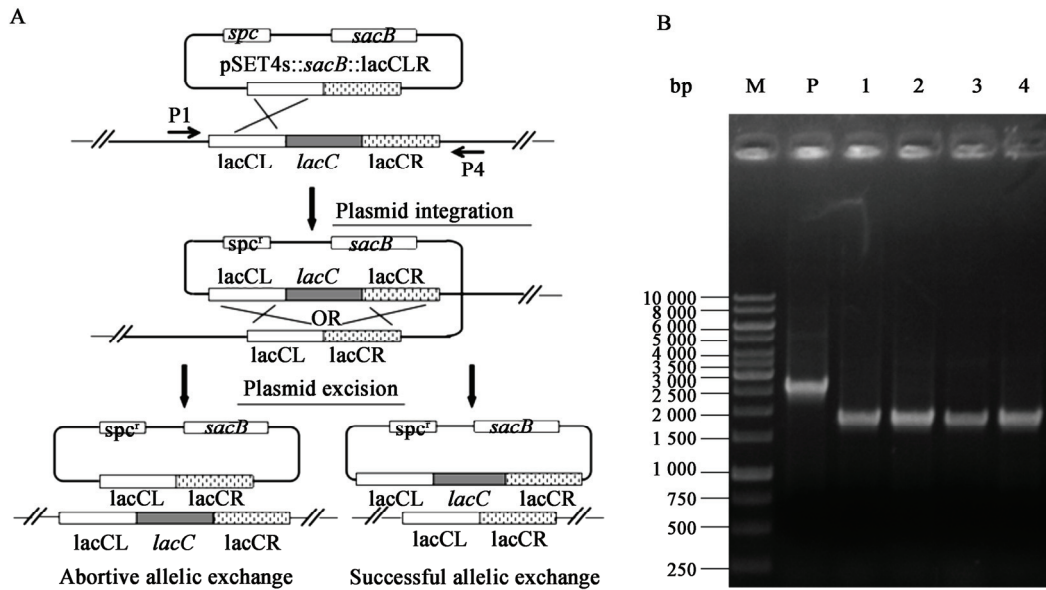


图4 *lacC* 基因敲除菌株 PCR 验证

Figure 4 PCR validation of the $\Delta lacC$ strains

注: A: *lacC* 基因敲除原理示意图; B: *lacC* 基因敲除验证图.

Note: A: Principle of *lacC* gene deletion; B: Verification of *lacC* gene deletion. M: 1 kb DNA marker; P: PCR amplification using wild type genome as templates and lacC-P1/lacC-P4 as primers; 1-4: PCR amplification using different transformants genomes as templates and lacC-P1/lacC-P4 as primers.

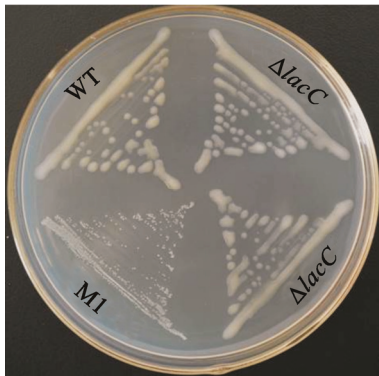


图5 *lacC* 基因敲除菌株表型分析

Figure 5 Assay of the phenotypes of $\Delta lacC$ mutants

注: WT: 野生型菌株; M1: 1号突变株; $\Delta lacC$: *lacC* 基因敲除菌株。

Note: WT: Wild type strain; M1: No.1 mutant; $\Delta lacC$: *lacC* gene deletion strain.

量, 所有测序 Reads 数量为 12 138 824, 所有比对上参考基因组的 Reads 数目为 11 650 709, 比对上参考基因组上的平均比对深度达 527, 参考基因组被 Reads 比对的覆盖度达 99.99%。进一步与兽疫链球菌野生型基因组序列比对, 发现除了已成功克隆的 *lacC* 基因位点存在插入突变, 另有 206 612 位点处碱基 G 缺失, 在导致编码透明质酸合成酶的 *hasA* 基因发生移码突变, 致使其功能域发生改变 (图 6)。

2.5 M1 菌株回补 *hasA* 基因分析

将本实验室构建的 *hasA* 回补质粒

pWV01::P_{pgk}::*hasA* 转化到 M1 突变株感受态细胞中, 得到 M1/pLH200::P_{pgk}::*hasA*, 将其与兽疫链球菌野生型菌株 M1 突变株接种于 THY 液体培养基中, 37 °C、200 r/min 培养 8 h (*OD*₆₆₀ 约为 0.6), 将 3 株菌划线于 THB 固体平板上, 37 °C 静置培养 24 h 后, 观察表型 (图 7)。结果表明, M1 突变株在回补 *hasA* 基因功能后恢复生成荚膜的能力, 即 M1 突变株表型缺陷由 *hasA* 基因移码突变引起。

通过对 M1 全基因组重测序和对 *hasA* 基因的回补实验表明, M1 突变是由随机自发突变而不是插入突变引起; M1 突变株粘性荚膜合成能力的缺失由 *hasA* 基因功能缺失引起, 与 *lacC* 基因功能缺失无关。

通过质粒介导的插入突变构建的突变体库, 出发点是希望质粒能够插入到基因组中且对菌株透明质酸合成有一定的影响, 从而从该突变体库中探究新的与透明质酸合成相关的基因。从本研究的结果中不能确定构建的突变体库中突变体表型的差异均为插入突变引起, 可能也是菌株自发突变引起的。不过, 自然界中发生随机自发突变的频率原则上是很低的, 筛选到的 150 株突变体对于探究与透明质酸合成相关新基因还是有重要的意义, 且本研究探究的策略为进一步从突变体库中探究与透明质酸合成相关新基因奠定了很好的基础。

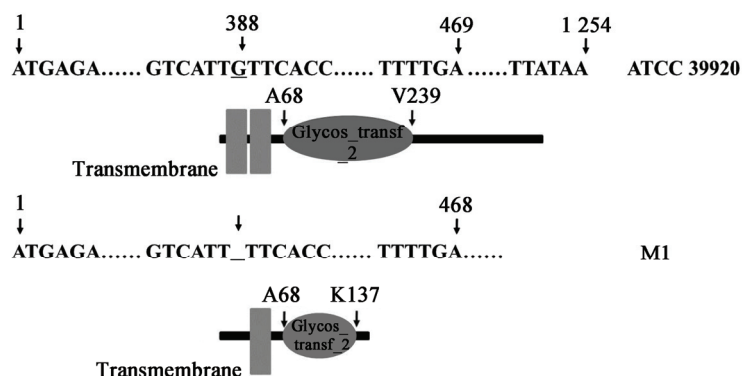


图6 野生型菌株与 M1 突变株 *hasA* 基因序列及功能域分析

Figure 6 Assay of the sequences of *hasA* gene and domain of wide type strain and M1 mutant

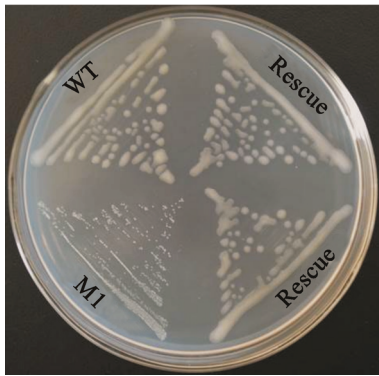


图 7 回补 *hasA* 基因的 M1 突变株表型分析

Figure 7 Assay of the phenotypes of M1 mutant rescuing *hasA*

注: WT: 野生型菌株; M1: 1 号突变株; Rescue: 回补 *hasA* 基因的 M1 突变株。

Note: WT: Wild type strain; M1: No.1 mutant; Rescue: *hasA* complementation in M1 mutant.

3 结论

随着生物信息学和分子生物学技术的发展, 阐明 HA 合成机理, 在分子水平改造兽疫链球菌获得 HA 高产菌株具有重要意义。本研究建立了一种挖掘与兽疫链球菌 HA 合成相关新基因的研究策略, 即通过自杀质粒随机插入基因组破坏与 HA 合成相关基因, 并通过直观的表型分析筛选构建兽疫链球菌突变体库, 通过质粒介导连接 PCR 技术克隆自杀质粒整合位点, 并通过敲除整合位点侧翼序列和基因回补实验, 验证表型和基因功能的相关性。本实验通过自杀载体随机插入成功构建包含 150 株突变株的突变体库。利用克隆载体 pUC19 介导的 LM-PCR 技术成功克隆到自杀载体整合位点, LM-PCR 方法虽然实验结果准确但工作效率较低, 操作比较繁琐(需要酶切、连接、转化、PCR 验证), 通过进一步探索优化实验步骤提高检测效率仍是工作重点。同时研究发现, 在基因组上除质粒插入突变外还存在另一位点的碱基缺失突变, 这种非特异性突变对通过突变体库挖掘新基因的工作有一定的影响, 也说明了通过基因敲除和基因回补验证

基因功能和表型相关性的重要性, 至于这种碱基缺失突变与自杀质粒随机插入是否具有相关性暂不清楚。

参考文献

- [1] Balazs EA, Denlinger JL. Viscosupplementation: a new concept in the treatment of osteoarthritis[J]. The Journal of Rheumatology, 1993, 39(Supplement): 3-9
- [2] Vázquez JA, Montemayor MI, Fraguas J, et al. Hyaluronic acid production by *Streptococcus zooepidemicus* in marine by-products media from mussel processing wastewaters and tuna peptone viscera[J]. Microbial Cell Factories, 2010, 9: 46
- [3] Yamamoto H. Antiaging and cosmetic effects of dietary hyaluronic acid (extracellular matrix extract)[J]. New Food Industry, 1998, 40(11): 33-41
- [4] Beasley KL, Weiss MA, Weiss RA. Hyaluronic acid fillers: a comprehensive review[J]. Facial Plastic Surgery, 2009, 25(2): 86-94
- [5] Sugahara K, Schwartz NB, Dorfman A. Biosynthesis of hyaluronic acid by *Streptococcus*[J]. Journal of Biological Chemistry, 1979, 254(14): 6252-6261
- [6] Beres SB, Sesso R, Pinto SWL, et al. Genome sequence of a lancefield group C *Streptococcus zooepidemicus* strain causing epidemic nephritis: new information about an old disease[J]. PLoS One, 2008, 3(8): e3026
- [7] Ma Z, Geng JN, Zhang H, et al. Complete genome sequence of *Streptococcus equi* subsp. *zooepidemicus* strain ATCC 35246[J]. Journal of Bacteriology, 2011, 193(19): 5583-5584
- [8] Chong BF, Blank LM, McLaughlin R, et al. Microbial hyaluronic acid production[J]. Applied Microbiology and Biotechnology, 2005, 66(4): 341-351
- [9] Zhang NZ, Lu ZX. The methods of constructing microbial mutant[J]. Biotechnology Bulletin, 2010(2): 68-71 (in Chinese) 张念章, 逯忠新. 构建微生物突变体的方法综述[J]. 生物技术通报, 2010(2): 68-71
- [10] Muskett PR, Clissold L, Marocco A, et al. A resource of mapped dissociation launch pads for targeted insertional mutagenesis in the *Arabidopsis* genome[J]. Plant Physiology, 2003, 132(2): 506-516
- [11] Ogasawara N. Systematic function analysis of *Bacillus subtilis* genes[J]. Research in Microbiology, 2000, 151(2): 129-134
- [12] Liu D, Li B, Zhang DJ. Detection of integrated exogenous sequence by using the ligation-mediated PCR method[J]. Journal of Northeast Agricultural University, 2013, 44(12): 93-97 (in Chinese) 刘娣, 李斌, 张冬杰. 利用连接介导 PCR 检测外源基因整合位点体系的建立[J]. 东北农业大学学报, 2013, 44(12): 93-97
- [13] Sun XY, Yang DD, Wang YY, et al. Development of a markerless gene deletion system for *Streptococcus zooepidemicus*: functional characterization of hyaluronan synthase gene[J]. Applied Microbiology and Biotechnology, 2013, 97(19): 8629-8636
- [14] Takamatsu D, Osaki M, Sekizaki T. Thermosensitive suicide vectors for gene replacement in *Streptococcus suis*[J]. Plasmid, 2001, 46(2): 140-148
- [15] Ho SN, Hunt HD, Horton RM, et al. Site-directed mutagenesis by overlap extension using the polymerase chain reaction[J]. Gene, 1989, 77(1): 51-59