

猪链球菌荚膜多糖合成相关基因簇保守区的克隆与分析

王楷成^{1,2}, 陆承平², 范伟兴^{1*}

¹中国动物卫生与流行病学中心, 青岛 266032

²南京农业大学动物医学院, 南京 210095

摘要: 【目的】为了解猪链球菌各血清型荚膜多糖合成相关基因保守区的功能与基因进化关系, 【方法】在分析已知的猪链球菌 1、2、7、9 型荚膜多糖合成相关基因簇序列, 及其各 orf 与猪链球菌 33 个血清型基因组 DNA 杂交结果的基础上, 提出猪链球菌荚膜多糖合成相关基因簇具有与肺炎链球菌相似的盒样结构的假设。并采用 PCR、测序和 Southern 印迹杂交等方法验证这些假设。 【结果】结果显示, 猪链球菌的荚膜多糖合成相关基因簇确存在与肺炎链球菌相似的盒样结构, 5' 端的前 4 个调节相关基因同源性极高, 基因簇两端都有保守的侧翼基因, 且在 3' 端的侧翼序列中找到了适于扩增荚膜多糖合成相关基因簇中血清型特异性区域的下游引物所在基因 (*aroA*)。分析发现, 各血清型的 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC*、*cpsD* 和 *aroA* 的亲缘关系较近。

关键词: 猪链球菌, 荚膜多糖合成相关基因簇, 保守基因, 侧翼基因

中图分类号: S852 文献标识码: A 文章编号: 0001-6209 (2012)05-0573-08

猪链球菌 (*Streptococcus suis*) 是一种人畜共患病原, 北美洲 (美国、加拿大) 到南美洲 (智利), 欧洲 (英国、荷兰、法国、丹麦、挪威、西班牙和德国), 亚洲 (中国、泰国、越南、韩国和日本), 大洋洲 (澳大利亚和新西兰) 都有猪链球菌流行^[1]。荚膜多糖是猪链球菌的重要毒力因子^[2], 也是猪链球菌血清型的区分依据^[3]。荚膜合成相关基因簇 (CPS locus) 是猪链球菌各血清型鉴别的分子基础, 但国内外对此研究均较少, 仅 Smith 等通过基因文库筛选的方法得到 2 型荚膜多糖合成相关基因簇以及 1、7、9 型荚膜多糖合成相关基因簇中的一部分序列^[4-5]。2 型 CPS locus 中各开放阅读框 (open reading frame, orf)

与 33 个血清型猪链球菌的基因组杂交结果^[4,6]显示, *cpsA-D* 在与多数血清型的基因组杂交呈阳性, 且侧翼序列 *orfZ*、*orfY* 和 *orfX* 都较保守, 而猪链球菌 2 型的 *cpsE* (*cps2E*) 仅与 1、14、1/2 型 DNA 发生杂交, 而猪链球菌 7 型的 *cpsE* (*cps7E*) 能与 3、4、5、7、9、10、11、12、17、18、19、21、23、24、29、30、32 型的基因组 DNA 发生杂交。

本文在分析其他链球菌 CPS locus 特点基础上, 用 PCR 方法扩增猪链球菌各血清型可能的保守基因, 并通过比对分析了解各血清型间这些保守基因的关系及进化, 同时寻找 CPS locus 的保守侧翼序列以用于下一步完整基因簇的扩增。

基金项目: 公益性行业 (农业) 科研专项 (20080316); 农业部“重大动物疫病防治关键技术引进”项目 (2006-G57)

* 通信作者。E-mail: fwxsjl@126.com

作者简介: 王楷成 (1981-), 女, 山东人, 助理研究员, 博士, 主要从事兽医微生物研究。Tel: +86-532-85630682; E-mail: kaichengwang@yahoo.cn

收稿日期: 2011-12-02; 修回日期: 2012-01-25

1 材料和方法

1.1 材料

1.1.1 细菌与培养:猪链球菌 33 个血清型的参考菌株由加拿大蒙特利尔大学 Gottschalk 教授提供。猪链球菌在 THB 培养基 (BD 公司, 美国) 或含 10% (v/v) 绵羊血的哥伦比亚琼脂培养基上生长。大肠杆菌在 LB 液体 (BD 公司, 美国) 或 LB 琼脂平板上生长, 需要时添加氨苄青霉素 (Amp) 至终浓度为 50 $\mu\text{g}/\text{mL}$ 。

1.1.2 主要试剂和仪器:PCR 仪为 Perkin Elmer 9600 型, 小型高速离心机购自德国 eppendorf 公司。Premix LA Taq 和 Premix Taq 购自 Takara 公司, UNIQ-10 柱式 DNA 胶回收试剂盒和质粒提取试剂盒购自上海生工生物有限公司。

1.2 模板 DNA 的提取

挑取含 10% (v/v) 绵羊血的哥伦比亚琼脂培养基上生长过夜的猪链球菌 33 个血清型的参考菌株, 接种 THB 液体培养基, 37 $^{\circ}\text{C}$ 静置培养过夜, 取 4mL 菌液用 Wizard Genomic DNA Purification Kit (Promega) 按说明书提取细菌的基因组 DNA, 并溶于灭菌超纯水, -20 $^{\circ}\text{C}$ 保存。

1.3 CPS locus 的 5'端保守基因及侧翼序列的克隆测序

1.3.1 引物设计与合成:按 GenBank 中猪链球菌 2 型 P1/7 株的 CPS locus (GenBank 登录号: AM946016.1, position: 549473-575754), 利用 Primer 软件设计引物 orfZ-cps2E-1 和 orfZ-cps2E-2 扩增 1、6、8、14、15、16、25、27 和 1/2 型的 *orfZ-cpsE* 段基因; 参考 GenBank 中猪链球菌 7 型的一部分 CPS locus 序列 (GenBank 登录号: AAF18953.1) 设计引物 orfZ-cps7E-1 和 orfZ-cps7E-2, 扩增其他血清型的 *orfZ-cpsE* 段基因。引物由上海生工生物工程有限公司合成, 引物序列见表 1。

表 1 CPS locus 的 5'端保守基因及侧翼序列的扩增引物
Table 1 The primers to amplify the conserved genes and flanking genes at 5'-end of *S. suis* serotype 2 CPS locus

Primer	Sequence(5'→3')
orfZ-cps2E-1	ATTACAGGTGGGCTATCGGGTAGTTT
orfZ-cps2E-2	AATTCCAGAACGATTTAAATCAGCA
orfZ-cps7E-1	ATTACAGGTGGGCTATCGGGTAGTTTT
orfZ-cps7E-2	AGAATGTGGCTGAGCGGCTAAACTG

1.3.2 PCR 反应:按 Premix LA Taq 25 μL , 灭菌超纯水 22 μL , orfZ-cps2E-1 或 orfZ-cps7E-1 (10 pmol/L) 1 μL , orfZ-cps2E-2 或 orfZ-cps7E-2 (10 pmol/L) 1 μL , DNA 1 μL 配置为 50 μL 的 PCR 反应体系。反应条件为: 94 $^{\circ}\text{C}$ 5 min; 94 $^{\circ}\text{C}$ 30 sec, 60 $^{\circ}\text{C}$ 30 sec, 72 $^{\circ}\text{C}$ 10 min, 30 个循环; 72 $^{\circ}\text{C}$ 10 min。

1.3.3 目的片段的回收:PCR 产物经琼脂糖凝胶电泳后, 切下目的条带, 用 UNIQ-10 柱式 DNA 胶回收试剂盒回收凝胶中的目的片段, 具体步骤按试剂盒说明书操作。回收产物 -20 $^{\circ}\text{C}$ 保存。

1.3.4 目的片段的克隆:回收的各血清型 *orfZ-cpsE* 片段克隆至 pCR-XL-TOPO 载体 (TOPO XL PCR Cloning Kit, invitrogen), 具体操作参照说明书进行。短暂离心连接反应体系, 取 5 μL 连接反应产物加入 200 μL 感受态细胞 TOP10 Chemically Competent *E. coli* (invitrogen) 中, 混匀, 冰上放置 30 min。42 $^{\circ}\text{C}$ 热激 90 s, 迅速置于冰上 2 min, 加入 800 μL LB 液体培养基, 37 $^{\circ}\text{C}$ 150 rpm 振荡培养 1 h。3000 r/min 离心 5 min, 弃上清液 800 μL , 用剩余的 200 μL 重悬细胞, 涂布含 Amp 的 LB 平板, 37 $^{\circ}\text{C}$ 培养 16 h。

1.3.5 克隆鉴定与测序:挑取平板上长出的重组菌, 接种 LB 液体培养基 (含 Amp 50 $\mu\text{g}/\text{mL}$), 37 $^{\circ}\text{C}$ 振荡培养过夜, 按质粒提取试剂盒说明书提取质粒, 以扩增该片段的引物进行 PCR, 鉴定质粒中是否已克隆目的片段。每个片段的阳性克隆各挑取一个送上海生工生物工程有限公司进行双向 Walking 测序, 测序结果经 Vector NTI 拼接并进行 orf 分析。

1.4 CPS locus 的 3'端保守侧翼序列的筛选

按 GenBank 中猪链球菌 2 型 P1/7 株的 CPS locus (GenBank 登录号: AM946016.1, position: 549473-575754) 序列, 利用 Primer 软件设计引物分别扩增其终止子下游的 IS30、*ssu0554*、*aroA* 和 *aroK* 4 个基因, 引物序列和反应条件见表 2。其中, 16S rDNA 的扩增引物同文献^[7]报道。反应体系均为: Premix Taq (TaKaRa) 25 μL , 上下游引物 (10 pmol/L) 各 1 μL , 2 型参考菌株 735 的基因组 DNA 1 μL , 超纯水补足至 50 μL 。PCR 产物经琼脂糖凝胶电泳后, 切下目的条带, 用 UNIQ-10 柱式 DNA 胶回收试剂盒回收目的片段。回收的产物用 DIG High Prime DNA Labeling and Detection Starter Kit I (Roche) 标记为探针, 与猪链球菌 33 个血清型参考菌株的基因组进行 Southern 印迹杂交实验^[8]。杂交温度为 65 $^{\circ}\text{C}$,

同时以 16S rDNA 探针与猪链球菌 33 个血清型参考菌株基因组 DNA 进行杂交,验证核酸绑定效果。

表 2 猪链球菌 2 型 CPS locus 3'端保守侧翼基因扩增引物与反应条件

Table 2 The sequences of the primers and the PCR programs to amplify the flanking genes at 3'-end of *S. suis* serotype 2 CPS locus

Gene	Primer	Sequence(5'→3')	PCR programs	Size/bp ^a
IS30	IS30-1	AAGAAAATGAAAGTGGGT	94℃ 5 min; 94℃ 30 sec, 45℃ 30 sec, 72℃ 30 sec, 30 个循环; 72℃ 10 min	448
	IS30-2	ATGAAGAATAGGGATGCG		
ssu0554-1	ssu0554-1	TCAAACCATCCAAAATAATAAC	94℃ 5 min; 94℃ 30 sec, 44℃ 30 sec, 72℃ 2 min, 30 个循环; 72℃ 10 min	2040
	ssu0554-2	AGCACATACAGATCAAAAAGTCA		
aroA	aroA-1	CCATGCAGGTCTTTCCG	94℃ 5 min; 94℃ 30 sec, 40℃ 30 sec, 72℃ 1 min, 30 个循环; 72℃ 10 min	781
	aroA-2	ATCAACCGAGGAATCAA		
aroK	aroK-1	TTACTTGGATTATGGGA	94℃ 5 min; 94℃ 30 sec, 46℃ 30 sec, 72℃ 30 sec, 30 个循环; 72℃ 10 min	462
	aroK-2	CCTTGCTACTTGTCTGG		
16S rDNA	16SrDNA-1	AGAGTTTGATCCTGGCTNAG	94℃ 5 min; 94℃ 30 sec, 55℃ 30 sec, 72℃ 1.5 min, 30 个循环; 72℃ 10 min	1457
	16SrDNA-2	TGACGGGCGGTGTGTACAA		

1.5 CPS locus 的 3'端保守侧翼序列的测序

通过探针杂交试验筛选出的较靠近 CPS locus 终止子的保守 orf(*aroA*) 后,扩增 33 个血清型此基因的 PCR 产物,经胶回收后送上海生工生物工程有限公司进行双向测序。

1.6 序列分析

所有血清型的 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC*、*cpsD* 和 CPS locus 5'端保守侧翼序列(*aroA*) 序列,分别用 ClustalX^[9] 进行多重序列比对,并用 MEGA (version 4) 以 neighbour-joining 算法进行序列比较并绘制进化树^[10]。

2 结果

2.1 CPS locus 的 5'端保守基因及侧翼序列

各血清型的 PCR 扩增结果见图 1。引物 orfZ-cps7E-1 和 orfZ-cps7E-2 扩增产物约为 8000 bp,引物 orfZ-cps2E-1 和 orfZ-cps2E-2 扩增产物约为 7000 bp。这两对引物扩增 26 和 33 型 DNA 时,均未得到目的条带。各血清型的此部分基因已被 GenBank 收录,登陆号为:JF791152-JF791167,以及 JF273644-JF273656 和 HQ694980 中的 orfZ-cpsE 区域的序列。这些血清型均具有 *orfY* 与 *orfX* 这两个 CPS locus 5'端的侧翼基因,以及 *cpsA*、*cpsB*、*cpsC* 和 *cpsD* 这 4 个保守基因。分析 GenBank 中 9 个猪链球菌的全基因序列(登陆号为 AM946016, AAFA00000000, AARD00000000, FM252031, FM252032, CP000407, CP000408, CP002465.1, CP000837.1),可知该细菌的核酸 GC 含量为 41%。而得到的所有 31 个血清型的 *orfY*、*orfX*、*cpsA* 和 *cpsD* 的 GC 含量均约为

38%,猜测猪链球菌的这部分基因是通过侧向基因转移(lateral gene transfer, LGT) 机制^[11,12],从其他细菌或其它来源获得。多数血清型 *cpsB* 的 GC 含量约为 41%,仅 6 型此基因的 GC 含量为 38%,推测 6 型的 *cpsB* 来源于其他细菌或物种。同样的情况也存在于 *cpsC* 基因,除 6 型和 22 型(38%)外,其余血清型的此基因 GC 含量均约为 41%。

2.2 CPS locus 的 3'端保守侧翼序列的筛选与测序

Southern 印迹杂交结果见表 3, *aroA*、*aroK* 和阳性对照(16S rDNA) 与 33 个血清型的基因组都产生杂交, *aroA* 与 *aroK* 是 CPS locus 3'端的保守侧翼序列。为下一步 CPS locus 型特异性部分的扩增,扩增所有血清型的 *aroA* 基因,除 26 型未得到目的片段外,其余各血清型均有约 750bp 的目的条带(图 2)。各血清型的 *aroA* 已被 GenBank 收录,登陆号为: JF791121-JF791151。各血清型的 *aroA* 基因 GC 含量平均约为 42.6%,与猪链球菌基因组的 GC 含量相近。

2.3 进化树分析

测得的猪链球菌各血清型的 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC*、*cpsD* 和 *aroA* 的基因比对结果见图 3。结果显示,不同血清型中的这 7 个基因亲缘关系均较近。各血清型 CPS locus 5'端的侧翼序列(*orfY* 和 *orfX*) 和第 1 个调节基因(*cpsA*) 的进化关系极为相似,尤其是 group A(3、4、5、8、23 型)、group B(1、2、14、29、1/2 型) 和 group C(17、18、19、22 型) 这 3 组的 *orfY*、*orfX* 和 *cpsA* 在进化树中始终各处于一个分支,猜测这 3 组血清型的 CPS locus 5'端分别来源于各自的共同祖先。*cpsB* 和 *cpsC* 进化树形式相似,多

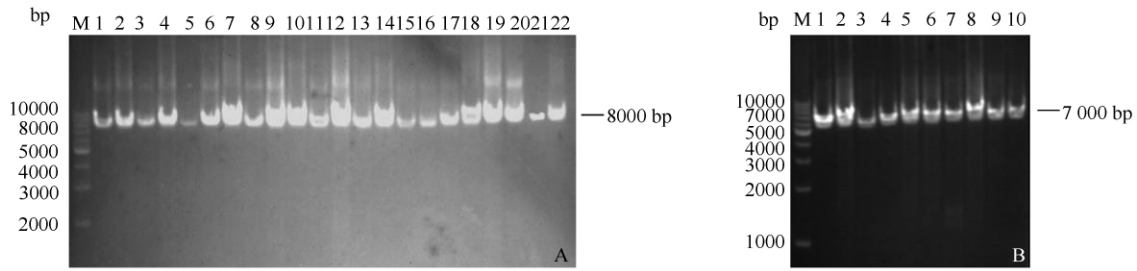


图1 猪链球菌各血清型 *orfZ-cpsE* 段扩增产物电泳图

Fig. 1 Ethidium bromide-stained agarose gel showing *orfZ-cpsE* fragments of *S. suis* strains. (A) *orfZ-cpsE* fragments amplified by *orfZ-cps7E-1* and *orfZ-cps7E-2*. M: molecular size marker (1 Kbp DNA Ladder Marker, Takara); 1: *S. suis* serotype 3; 2: *S. suis* serotype 4; 3: *S. suis* serotype 5; 4: *S. suis* serotype 7; 5: *S. suis* serotype 9; 6: *S. suis* serotype 10; 7: *S. suis* serotype 11; 8: *S. suis* serotype 12; 9: *S. suis* serotype 13; 10: *S. suis* serotype 17; 11: *S. suis* serotype 18; 12: *S. suis* serotype 19; 13: *S. suis* serotype 20; 14: *S. suis* serotype 21; 15: *S. suis* serotype 22; 16: *S. suis* serotype 23; 17: *S. suis* serotype 24; 18: *S. suis* serotype 28; 19: *S. suis* serotype 29; 20: *S. suis* serotype 30; 21: *S. suis* serotype 31. (B) *orfZ-cpsE* fragments amplified by *orfZ-cps2E-1* and *orfZ-cps2E-2*. M: molecular size marker (1 Kbp DNA Ladder Marker, Takara); 1: *S. suis* serotype 1; 2: *S. suis* serotype 2; 3: *S. suis* serotype 6; 4: *S. suis* serotype 8; 5: *S. suis* serotype 14; 6: *S. suis* serotype 15; 7: *S. suis* serotype 16; 8: *S. suis* serotype 25; 9: *S. suis* serotype 27; 10: *S. suis* serotype 1/2.

表3 猪链球菌2型 CPS locus 的3'端保守侧翼序列与33个血清型基因组DNA的杂交结果

Table 3 Hybridization of the flanking genes at 3'-end of CPS locus with chromosomal DNA of *S. suis* serotypes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	33	1/2	
IS30	+	-	-	-	+	-	+	+	+	+	-	-	-	+	+	+	+	+	+	+	-	±	-	-	-	+	-	+	+	+	-	-	+	
ssu0554	+	+	+	+	+	+	+	+	+	+	±	±	±	+	+	+	+	+	+	+	±	±	±	±	±	±	+	±	+	+	+	±	+	+
aroA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	±	+	+	+	±	+	+	+	+	+	+	+	+
aroK	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16S rDNA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

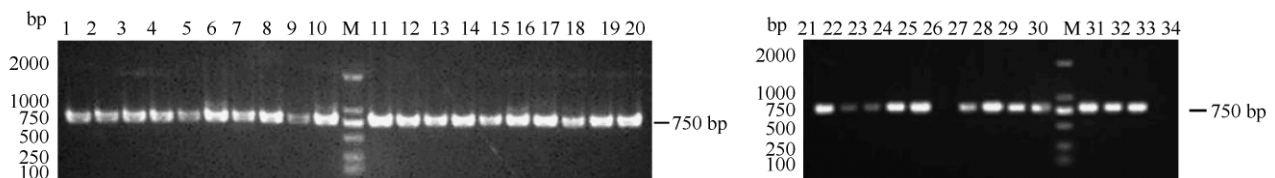


图2 猪链球菌各血清型 *aroA* 段扩增产物电泳图

Fig. 2 Ethidium bromide-stained agarose gel showing *aroA* fragments of *S. suis* strains. M: molecular size marker; 1-31: *S. suis* serotype 1-31; 32: *S. suis* serotype 33; 33: *S. suis* serotype 1/2; 34: negative control.

数血清型的该基因属同一分支,亲缘关系极近,仅 *cps6B*、*cps6C* 和 *cps22C* 与其余血清型的同一基因亲缘关系较远,结合 2.1 中 GC 含量的分析结果可知, *cps6B*、*cps6C* 和 *cps22C* 从外源 DNA 获取的几率较大。

3 讨论

猪链球菌 2 型 CPS locus 中各 *orf* 与 33 个血清型猪链球菌的基因组杂交结果显示, *cpsA-D* 在与多数血清型的基因组发生杂交,且 2 型的 *cpsE* 与 1、

14、1/2 的基因组杂交呈阳性,而与其它血清型的基因组杂交呈阴性^[4,6]。猪链球菌 7 型与 9 型的 *cpsE* 同源性极高,且与 1、14、1/2 之外的其它多数血清型的基因组杂交呈阳性^[4,5]。根据这一现象,本实验分别设计了 2 对引物扩增 33 个血清型的 CPS locus 的 5'端保守基因及侧翼序列,经测序知此部分序列均包含 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC* 和 *cpsD*,与 2 型一致。猪链球菌 31 个血清型(不含 26 型和 33 型)荚膜多糖合成相关基因簇的前 4 个 *orf*(*cpsABCD*)同源性极高。*cpsA* 翻译的蛋白被证实猪链球菌的荚膜多糖合成过程中起调节作用^[13]。而 *cpsB* 翻译

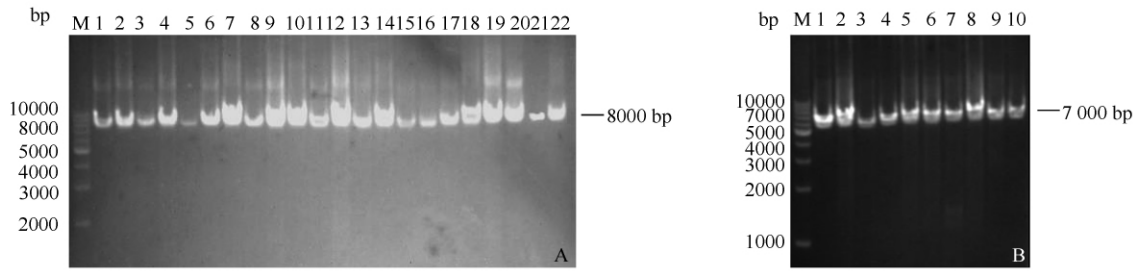


图1 猪链球菌各血清型 *orfZ-cpsE* 段扩增产物电泳图

Fig. 1 Ethidium bromide-stained agarose gel showing *orfZ-cpsE* fragments of *S. suis* strains. (A) *orfZ-cpsE* fragments amplified by *orfZ-cps7E-1* and *orfZ-cps7E-2*. M: molecular size marker (1 Kbp DNA Ladder Marker, Takara); 1: *S. suis* serotype 3; 2: *S. suis* serotype 4; 3: *S. suis* serotype 5; 4: *S. suis* serotype 7; 5: *S. suis* serotype 9; 6: *S. suis* serotype 10; 7: *S. suis* serotype 11; 8: *S. suis* serotype 12; 9: *S. suis* serotype 13; 10: *S. suis* serotype 17; 11: *S. suis* serotype 18; 12: *S. suis* serotype 19; 13: *S. suis* serotype 20; 14: *S. suis* serotype 21; 15: *S. suis* serotype 22; 16: *S. suis* serotype 23; 17: *S. suis* serotype 24; 18: *S. suis* serotype 28; 19: *S. suis* serotype 29; 20: *S. suis* serotype 30; 21: *S. suis* serotype 31. (B) *orfZ-cpsE* fragments amplified by *orfZ-cps2E-1* and *orfZ-cps2E-2*. M: molecular size marker (1 Kbp DNA Ladder Marker, Takara); 1: *S. suis* serotype 1; 2: *S. suis* serotype 2; 3: *S. suis* serotype 6; 4: *S. suis* serotype 8; 5: *S. suis* serotype 14; 6: *S. suis* serotype 15; 7: *S. suis* serotype 16; 8: *S. suis* serotype 25; 9: *S. suis* serotype 27; 10: *S. suis* serotype 1/2.

表3 猪链球菌2型 CPS locus 的3'端保守侧翼序列与33个血清型基因组DNA的杂交结果

Table 3 Hybridization of the flanking genes at 3'-end of CPS locus with chromosomal DNA of *S. suis* serotypes

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	33	1/2	
IS30	+	-	-	-	+	-	+	+	+	+	-	-	-	+	+	+	+	+	+	+	-	±	-	-	-	+	-	+	+	+	-	-	+	
ssu0554	+	+	+	+	+	+	+	+	+	+	±	±	±	+	+	+	+	+	+	+	±	±	±	±	±	±	+	±	+	+	+	±	+	+
aroA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	±	+	+	+	±	+	+	+	+	+	+	+	+	+
aroK	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16S rDNA	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

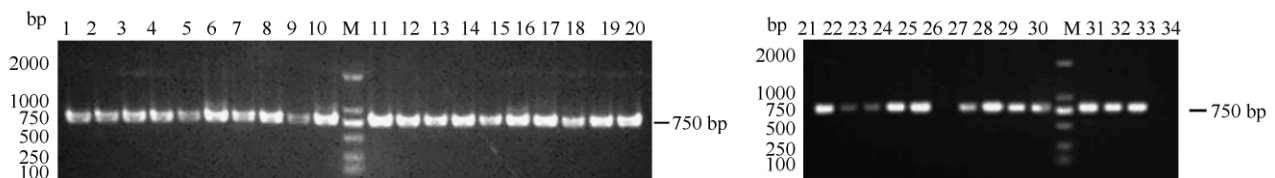


图2 猪链球菌各血清型 *aroA* 段扩增产物电泳图

Fig. 2 Ethidium bromide-stained agarose gel showing *aroA* fragments of *S. suis* strains. M: molecular size marker; 1-31: *S. suis* serotype 1-31; 32: *S. suis* serotype 33; 33: *S. suis* serotype 1/2; 34: negative control.

数血清型的该基因属同一分支,亲缘关系极近,仅 *cps6B*、*cps6C* 和 *cps22C* 与其余血清型的同一基因亲缘关系较远,结合 2.1 中 GC 含量的分析结果可知, *cps6B*、*cps6C* 和 *cps22C* 从外源 DNA 获取的几率较大。

3 讨论

猪链球菌 2 型 CPS locus 中各 *orf* 与 33 个血清型猪链球菌的基因组杂交结果显示, *cpsA-D* 在与多数血清型的基因组发生杂交,且 2 型的 *cpsE* 与 1、

14、1/2 的基因组杂交呈阳性,而与其它血清型的基因组杂交呈阴性^[4,6]。猪链球菌 7 型与 9 型的 *cpsE* 同源性极高,且与 1、14、1/2 之外的其它多数血清型的基因组杂交呈阳性^[4,5]。根据这一现象,本实验分别设计了 2 对引物扩增 33 个血清型的 CPS locus 的 5'端保守基因及侧翼序列,经测序知此部分序列均包含 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC* 和 *cpsD*,与 2 型一致。猪链球菌 31 个血清型(不含 26 型和 33 型)荚膜多糖合成相关基因簇的前 4 个 *orf*(*cpsABCD*)同源性极高。*cpsA* 翻译的蛋白被证实猪链球菌的荚膜多糖合成过程中起调节作用^[13]。而 *cpsB* 翻译

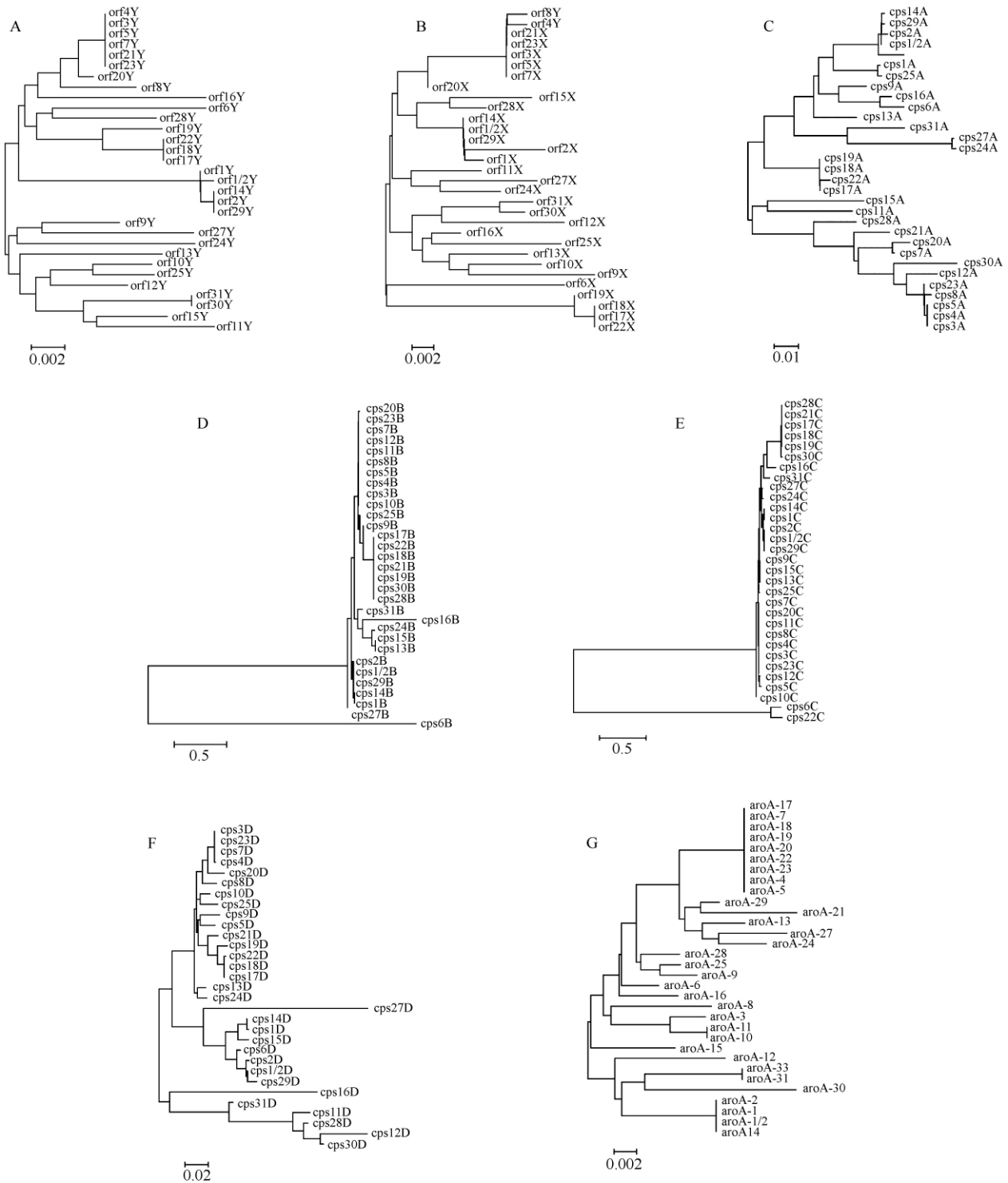


图 3 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC*、*cpsD* 和 *aroA* 的基因进化树

Fig. 3 Phylogenetic trees (minimum-evolution, neighbor-joining phylogenetic trees) of *orfY*, *orfX*, *cpsA*, *cpsB*, *cpsC*, *cpsD* and *aroA*. A: *orfY*; B: *orfX*; C: *cpsA*; D: *cpsB*; E: *cpsC*; F: *cpsD*; G: *aroA*. The scale bar at the bottom of each phylogenetic tree was used to measure the branch length and showed the expected number of substitutions per site.

蛋白与大肠杆菌脂多糖的合成调节蛋白同源性高，在大肠杆菌中与 CpsB 序列相似的蛋白能决定脂多

糖 O-抗原的糖链长度，能够增加长链的糖产物，减少短糖链的产生^[14]。CpsB 在猪链球菌荚膜多糖合

成过程中的重要作用也已被 Smith 等^[13]证实,人工缺失 *cps2B* 基因的猪链球菌 2 型菌株不产生荚膜。*cpsC* 和 *cpsD* 基因在肺炎链球菌、唾液链球菌 (*Streptococcus salivarius*)、嗜热链球菌 (*Streptococcus thermophilus*)、无乳链球菌 (*Streptococcus agalactiae*) 和金黄色葡萄球菌 (*Staphylococcus aureus*) 的荚膜合成相关基因簇中都存在^[15]。Paulsen 等对这些细菌中的 *cpsC* 和 *cpsD* 的同源性进行了分析^[16],这些基因的翻译蛋白与革兰氏阴性菌的多糖合成相关基因簇中某些基因翻译蛋白的 N 端和 C 端都很相似,这类蛋白可归类为 MPA1^[16],这类蛋白被证明与多糖的输出有关。在肺炎链球菌的荚膜多糖合成相关基因簇中,与 *cpsD* 相似基因的翻译蛋白被证明是一种胰蛋白激酶,而 *cpsC* 的翻译蛋白是其底物^[17]。*CpsC* 和 *CpsD* 在荚膜多糖的链长度控制和输出过程中都起到重要作用。

aroA 是猪链球菌的一个看家基因^[18],猪链球菌各血清型均具有此基因且保守。肠炎沙门菌的 *aroA* 与荚膜多糖和生物被膜的合成有关^[19],人工缺失 *aroA* 的菌株产生荚膜多糖和生物被膜的能力下降。本文扩增了 32 个血清型 (不含 26 型) 的 *aroA*,测序结果显示其同源性极高,同时考虑到其较 *aroK* 距 CPS locus 更近,且为猪链球菌的看家基因,故欲选择 *aroA* 基因作为下一步扩增 CPS locus 中型特异性区域下游引物设计的靶基因。

侧向基因转移 (lateral gene transfer, LGT), 又称水平基因转移 (horizontal gene transfer), 是指在差异物种之间所进行的遗传物质的交流,是推动物种进化的重要动力,对微生物基因组进化表现出深刻的影响。细菌可以通过 3 种方式进行 LGT,即接合 (conjugation)、转化 (transformation) 和转导 (transduction)。许多细菌都是通过侧向基因转移机制获得新的 CPS locus 或其中的部分基因^[20-22],各种细菌的基因转移频度不同。鉴定 LGT 的常用方法有两种:一种是 DNA 碱基组成分析,利用某些特异序列来区分外源 DNA 与细菌原有 DNA,一种是细菌之间的基因组序列比对。最先用来鉴定 LGT 的方法是 DNA 的碱基组成分析。它的理论基础是,不同生物体具有不同的 G + C 含量和密码子选择偏好 (codon usage patterns)。简单来说,某一开放读码

框架 (open reading frame, ORF) 的 G + C 含量如果明显高于或低于整个基因组的平均 G + C 含量,它就可能是来自 LGT。密码子选择偏好是指生物体多偏向于使用同义密码子中的某一个。如果一个 orf 中使用的同义密码子与基因组中其他地方的不同,那么这一 orf 就可能来自 LGT。通常用密码子适应指数 (codon adaptation index) 来表示密码子选择情况。细菌中高度表达的基因通常都使用宿主适应性密码子 (host adapted codon),而那些新插入的外源基因则不使用这些密码子^[23-25]。从本章对 *orfY*、*orfX*、*cpsA*、*cpsB*、*cpsC*、*cpsD* 的分析来看,猪链球菌的 CPS locus 可能大部分来源于外源物种的 DNA,同时在某些选择性压力下发生基因重组,且不同血清型 CPS locus 的来源各不相同,在进化过程中也可能产生过其他尚未被发现的血清型,其具体的机制和过程尚不清楚。

参考文献

- [1] Staats JJ, Feder I, Okwumabua O, Chengappa MM. *Streptococcus suis*: past and present. *Veterinary Research Communications*, 1997, 21(6): 381-407.
- [2] Charland N, Kobisch M, Martineau-Doize B, Jacques M, Gottschalk M. Role of capsular sialic acid in virulence and resistance to phagocytosis of *Streptococcus suis* capsular type 2. *FEMS Immunology & Medical Microbiology*, 1996, 14(4): 195-203.
- [3] Higgins R, Gottschalk M. An update on *Streptococcus suis* identification. *The Journal of Veterinary Diagnostic Investigation*, 1990, 2(3): 249-252.
- [4] Smith HE, Veenbergen V, van der Velde J, Damman M, Wisselink HJ, Smits MA. The *cps* genes of *Streptococcus suis* serotypes 1, 2, and 9: development of rapid serotype-specific PCR assays. *Journal of Clinical Microbiology*, 1999, 37(10): 3146-3152.
- [5] Smith HE, Bruijnsvoort L, Buijs H, Wisselink HJ, Smits MA. Rapid PCR test for *Streptococcus suis* serotype 7. *FEMS Microbiology Letters*, 1999, 178(2): 265-270.
- [6] Smith HE, Vries R, Slot R, Smits MA. The *cps* locus of *Streptococcus suis* serotype 2: genetic determinant for the synthesis of sialic acid. *Microbial Pathogenesis*, 2000, 29(2): 127-134.

- [7] Rasmussen SR , Andresen LO. 16S rDNA sequence variations of some *Streptococcus suis* serotypes. *International Journal of Systematic Bacteriology* , 1998 , 48 Pt 3: 1063-1065.
- [8] Sambrook J. Molecular Cloning: A Laboratory Manul. 3nd ed. New York: Cold Spring Harbor Laboratory press , 2001: 502-509.
- [9] Thompson JD , Gibson TJ , Plewniak F , Jeanmougin F , Higgins DG. The CLUSTAL _ X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* , 1997 , 25(24) : 4876-4882.
- [10] Tamura K , Dudley J , Nei M , Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* , 2007 , 24(8) : 1596-1599.
- [11] Campbell AM. Lateral gene transfer in prokaryotes. *Theoretical Population Biology* , 2000 , 57(2) : 71-77.
- [12] Ochman H , Lawrence JG , Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature* , 2000 , 405(6784) : 299-304.
- [13] Smith HE , Damman M , Velde J , Wagenaar F , Wisselink HJ , Stockhofe-Zurwieden N , Smits MA. Identification and characterization of the *cps* locus of *Streptococcus suis* serotype 2: the capsule protects against phagocytosis and is an important virulence factor. *Infection and Immunity* , 1999 , 67(4) : 1750-1756.
- [14] Franco AV , Liu D , Reeves PR. The *wzz* (*cld*) protein in *Escherichia coli*: amino acid sequence variation determines O-antigen chain length specificity. *Journal of Bacteriology* , 1998 , 180(10) : 2670-2675.
- [15] Wessels MR. Biology of streptococcal capsular polysaccharides. *Society for Applied Bacteriology Symposium Series* , 1997 , 26: 20S-31S.
- [16] Paulsen IT , Beness AM , Saier MH. Computer-based analyses of the protein constituents of transport systems catalysing export of complex carbohydrates in bacteria. *Microbiology* , 1997 , 143 (Pt 8) : 2685-2699.
- [17] Morona JK , Morona R , Miller DC , Paton JC. Mutational analysis of the carboxy-terminal (YGX)₄ repeat domain of CpsD , an autophosphorylating tyrosine kinase required for capsule biosynthesis in *Streptococcus pneumoniae*. *Journal of Bacteriology* , 2003 , 185(10) : 3009-3019.
- [18] King SJ , Leigh JA , Heath PJ , Luque I , Tarradas C , Dowson CG , Whatmore AM. Development of a multilocus sequence typing scheme for the pig pathogen *Streptococcus suis*: identification of virulent clones and potential capsular serotype exchange. *Journal of Clinical Microbiology* , 2002 , 40(10) : 3671-3680.
- [19] Malcova M , Karasova D , Rychlik I. *aroA* and *aroD* mutations influence biofilm formation in *Salmonella* Enteritidis. *FEMS Microbiology Letters* , 2009 , 291(1) : 44-49.
- [20] Nakhamchik A , Wilde C , Chong H , Rowe-Magnus DA. Evidence for the horizontal transfer of an unusual capsular polysaccharide biosynthesis locus in marine bacteria. *Infection and Immunity* , 2010 , 78(12) : 5214-5222.
- [21] Martins ER , Melo-Cristino J , Ramirez M. Evidence for rare capsular switching in *Streptococcus agalactiae*. *Journal of Bacteriology* , 2010 , 192(5) : 1361-1369.
- [22] Hakenbeck R , Balmelle N , Weber B , Gardes C , Keck W , Saizieu A. Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infection and Immunity* , 2001 , 69(4) : 2477-2486.
- [23] Sharp PM , Li WH. The codon Adaptation Index—a measure of directional synonymous codon usage bias , and its potential applications. *Nucleic Acids Research* , 1987 , 15(3) : 1281-1295.
- [24] Wang B. Limitations of compositional approach to identifying horizontally transferred genes. *Journal of Molecular Evolution* , 2001 , 53(3) : 244-250.
- [25] Koski LB , Morton RA , Golding GB. Codon bias and base composition are poor indicators of horizontally transferred genes. *Molecular Biology and Evolution* , 2001 , 18(3) : 404-412.

Cloning and analysis of the conserved regions in the capsular polysaccharides synthesis locus of *Streptococcus suis*

Kaicheng Wang^{1 2}, Chengping Lu², Weixing Fan^{1*}

¹ China Animal Health and Epidemiology Center, Qingdao 266032, China

² Nanjing Agricultural University, College of Veterinary, Nanjing 210095, China

Abstract: [Objective] To learn about the function and gene phylogenetic relationship of the conserved regions in capsular polysaccharides synthesis (CPS) locus, [Methods] based on the CPS locus sequence of *Streptococcus suis* serotype 1, 2, 7 and 9 and the results to the cross-hybridization experiments, the CPS locus of *Streptococcus suis* was hypothesized as cassette-like structure which is similar to *Streptococcus pneumoniae*. PCR, sequencing and southern blotting was used to certify the hypothesis. [Results] The CPS locus of *Streptococcus suis* was cassette-like structure. The 5'-end of the CPS locus contained 4 regulatory genes which are highly similar in all the serotypes. The flanking genes of the CPS locus are conservative. The flanking gene *aroA*, which is at the downstream to the CPS locus, was selected as the objective gene to develop the PCR to amplify the serotype-specific regions. The *orfY*, *orfX*, *cpsA*, *cpsB*, *cpsC*, *cpsD* and *aroA* are conserved with high sequence identity in different serotypes.

Keywords: *Streptococcus suis*, capsular polysaccharides locus, conserved genes, flanking genes

(本文责编:王晋芳)

Supported by the Special Fund for Public Welfare Industry of Chinese Ministry of Agriculture (200803016) and by the Fund for Key Technique Introduction of the Important Animal Disease Prevention from MOA(2006-G57)

* Corresponding author. E-mail: fwxsjl@126.com

Received: 2 December 2011 / Revised: 25 January 2012