

肺炎链球菌体内诱导基因的筛选及初步鉴定

孟江萍, 尹一兵*, 张雪梅, 黄远帅, 蓝 锴

(重庆医科大学医学检验系 临床检验诊断学省部共建教育部重点实验室 重庆 400016)

摘 要: 为筛选鉴定肺炎链球菌宿主体内诱导的基因, 寻找潜在的抗生素作用靶点和疫苗候选者, 应用体内表达技术, 以肺炎链球菌荚膜合成的关键基因 *galU* 作为体内报告基因, 利用其缺陷体不能合成荚膜多糖, 从而不能在宿主体内存活的特点, 筛选鉴定肺炎链球菌体内诱导基因。首先, 把肺炎链球菌基因组 DNA 的随机酶切片段(200~500bp)克隆到含有体内、体外双重报告基因(*galU-lacZ*)的报告载体 pEVP3-*galU* 的 *Bgl* II 位点, 将获得的质粒库转化肺炎链球菌 *galU* 缺陷菌株, 得到肺炎链球菌体内启动子诱捕文库, 将此文库去感染 BALB/c 小鼠, 经过两轮体内筛选, 在涂布有 X-gal 的 TSA 血清平板上得到了 165 个白色菌落, 对插入的随机片段进行测序及生物信息学分析, 共证实 15 个不同的体内诱导基因片段, 8 个为单独的 ORF, 7 个为含有多个 ORFs 的操纵子结构, 它们分别参与细菌在宿主体内的定植与粘附、能量代谢、物质转运、转录调节、DNA 复制与重组、细胞壁合成等, 另外还包括功能不明的假想蛋白。其中部分 ORFs 可能与细菌毒力相关, 可以作为候选疫苗和药物的靶标。

关键词: 肺炎链球菌; 体内诱导基因; 体内表达技术; 荚膜多糖

中图分类号: R378.14 **文献标识码:** A **文章编号:** 1001-6209(2006)04-0537-05

肺炎链球菌(*Streptococcus pneumoniae*, *S. pn*)是上呼吸道感染和严重侵袭性感染(脓毒症、脑膜炎)的最重要的病原体之一。目前全球每年死于 *S. pn* 感染的总人数大致与结核相仿, 高达 300 万至 500 万人, 其中 60 岁以上患者的病死率可以达到 20%^[1]。近年来全世界范围内 *S. pn* 耐药菌株的感染发病率显著增加, 使得 *S. pn* 感染的治疗面临越来越大的困难。而目前常用的 *S. pn* 23 价多糖荚膜疫苗只对某些型别 *S. pn* 的感染有效, 而且对两岁以下的小孩以及老年人的保护性都较弱^[2]。因此要彻底解决 *S. pn* 感染的防治, 需要全面理解肺炎链球菌分子的致病机理, 为新的药物和疫苗的研发提供可靠的理论依据。

病原菌感染宿主的每一个阶段都需要特异的基因表达, 这些基因的表达通常是受宿主体内不同环境所调节的。这些在宿主体内被诱导表达的基因简称为体内诱导基因(*ivi gene*), 在致病过程中起重要作用, 是良好的候选疫苗和药物靶标^[3]。那么, 如何才能确定哪些基因只在宿主体内表达, 研究细菌在宿主体内的真实表现呢? 目前用来筛选病原菌 *ivi* 基因的方法主要包括: 体内表达技术(*in vivo*

expression technology, *IVET*)^[4]、笔迹标识诱变(*signature-tagged mutagenesis*, *STM*)^[5]、DNA 芯片(*DNA chips*)^[6]和体内诱导抗原技术(*in vivo-induced antigen technology*, *IVIAT*)^[7], 其中基于启动子诱捕策略的 *IVET* 是一种鉴定 *ivi* 基因的有力工具, 近年来这种技术经不断地改良应用于鉴定多种革兰阳性、阴性致病菌、放线菌和真菌等的 *ivi* 基因。

IVET 的原理: 首先将细菌基因组随机片段 *X'* 与无启动子的报告基因融合, 构建启动子诱捕文库, 然后将此文库转化入宿主菌, 由于随机片段 *X'* 与宿主菌染色体上的相应片段 *X* 具有同源性, 可发生同源重组, 通过插入复制的方式整合入宿主菌染色体, 构建可随机报告体内基因表达的细菌库; 用构建好的细菌库去感染动物, 如果随机片段 *X'* 中含有体内诱导基因的启动子, 那么它将会激活这些基因及报告基因, 用相应的方法筛选, 就可得到细菌在宿主体内被激活的基因, 最后剔除那些在体外也表达的基因, 就得到真正只在宿主体内表达的基因。

本研究利用无荚膜 *S. pn* 在宿主体内毒力下降, 不能存活这一特点, 选用参与荚膜多糖合成的关键基因 *galU* 作为体内报告基因, 从 *S. pn* 启动子诱

基金项目: 国家自然科学基金资助(30371275, 30400376)

* 通讯作者。Tel: 86-23-68485658; E-mail: yibingyin@21.cn.com

作者简介: 孟江萍(1973-), 女, 四川省都江堰市人, 博士研究生, 主要致力于细菌致病分子机理的研究。E-mail: mjjily2000@163.com

收稿日期: 2005-10-08; 接受日期: 2005-10-19; 修回日期: 2006-03-29

捕文库中筛选出在宿主体内诱导表达的基因,并用 *lacZ* 作为体外报告基因,剔除那些在体外也表达的基因,最终鉴定出肺炎链球菌只在宿主体内诱导表达的基因。

1 材料和方法

1.1 材料

1.1.1 菌株、质粒和培养基 肺炎链球菌血清 3 型标准菌株 CMC(B)1203 购自中国药品生物制品检定所医学微生物菌种保存中心; *S. pn* 31203 *galU*⁻ 缺陷菌株由本课题组构建;大肠杆菌(*Escherichia coli*, *E. coli*) DH5 α 由本室保存;穿梭质粒 pEVP3(含无启动子的 *lacZ* 基因,氯霉素抗性)由美国 Morrison DA 教授惠赠。报告质粒 pEVP3-*galU*(将无启动子的 *galU* 基因融合到 *lacZ* 上游)由本课题组构建。C+Y 半合成培养基,用于 *S. pn* 的培养,含 0.5% Yeast 提取物(Lacks and Hotchkiss, 1960); TSA 血平板,含 5% 新生小牛血清和 2.5 μ g/mL 氯霉素; LB 培养基用于 *E. coli* DH5 α 的培养,氯霉素终浓度为 25 μ g/mL。

1.1.2 BALB/c 小鼠 雌性, 18 ~ 22g, 购自第三军医大学动物中心。

1.1.3 主要试剂 :DNA 限制性内切酶、CIAP、DNA 连接试剂盒、*Taq* DNA 聚合酶、细菌 DNA 提取试剂盒、胶回收及质粒提取试剂盒均购自 TaKaRa 公司。质粒 pEVP3 多克隆位点两侧测序引物 P1: 5'-CCCACCTTTATCCAATTTTCG-3'; P2: 5-AGGCGATTAA GTTGGGTA-3' 由上海生物工程有限公司合成。感受肽刺激肽(CSP), 由美国 Morrison DA 教授惠赠。氯霉素、X-gal、二甲基甲酰胺等购自上海生物工程有限公司。

1.2 肺炎链球菌启动子诱捕文库的构建

用 *Sau*3AI 部分酶切 *S. pn* 31203 基因组 DNA, 回收 200 ~ 500bp 的片段。质粒 pEVP3-*galU* 用 *Bgl* II 酶切后, 用 CAIP 酶去磷酸化。将两部分 DNA 连接起来, 转化 DH5 α , 收集转化菌落, 提取质粒, 得到含 *S. pn* 基因组 DNA 随机片段的质粒库。将此质粒库转化 *S. pn* 31203 *galU*⁻ 缺陷菌株, 得到可用 *galU* 报告 *S. pn* 体内基因表达的菌株库。

1.3 小鼠感染模型筛选肺炎链球菌体内表达基因

用构建好的启动子诱捕文库鼻腔感染 3 只 BALB/c 小鼠, 同时用 *galU* 缺陷菌株以同样方式感染 3 只 BALB/c 小鼠作对照; 感染 24h 后, 分别处死小鼠取肺组织, PBS 洗涤, 匀浆, 系列稀释后, 铺涂布

有 X-gal 的 TSA 血清平板, 37 $^{\circ}$ C 孵育 24 ~ 48h, 挑取所有白色菌落。再以同样方法进行第二轮筛选, 分别挑取最后一轮 X-gal 平板上的白色菌落于 C+Y 培养基中增菌, 50% 甘油保存。

1.4 体内诱导基因片段的克隆和测序

应用酶切自连法, 使整合到染色体上重组自杀质粒脱落下来并重新环化, 首先提取白色菌落的染色体 DNA, 用 *Bam*H I 酶切, 回收酶切片段后, 用 DNA 连接酶使酶切片段自身环化, 在含 25 μ g/mL 氯霉素的 LB 平板上筛选含重组质粒的菌落, 增菌后送上海生物工程有限公司, 用 pEVP3 多克隆位点上引物 P1 对 *galU* 基因前所插入序列进行单向测序。

1.5 生物信息学分析

基因序列同源性分析采用 NCBI 的 BLAST 软件 (blastn 和 blastx): <http://www.ncbi.nlm.gov/BLAST/>; 开放读码框搜索采用 NCBI 的 ORF Finder 软件: <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>; 开放读码框的分类与描述通过检索肺炎链球菌 TIGR4 基因组数据库获得: <http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi>; *Streptococcus pneumoniae* TIGR4, complete genome, 登陆号 AE005672。

1.6 确证所筛选出的体内诱导基因只在体内表达

体外测定含体内诱导基因的肺炎链球菌菌株的 β -半乳糖苷酶活性, 证实这些基因在体外不表达; 将所筛选出的白色菌落体外增菌后, 以同样方式分别感染小鼠, 证实这些基因在体内表达。

2 结果

2.1 启动子诱捕文库的构建

胶回收 200 ~ 500bp 的基因组 DNA 酶切片段, 与报告质粒 pEVP3-*galU* 按 3 ~ 10 : 1 的比例连接, 转化 DH5 α , 共获得约 75000 个转化子, 即为含 *S. pn* 基因组 DNA 随机片段的质粒库 (pEVP3-X-*galU*)。

由于文库质量取决于文库的容量与插入率, 考虑到片段插入方向和片段平均大小 (300bp), *S. pn* 的基因组全长约为 2.2Mb, 我们的文库大约覆盖基因组全长的 5 倍。文库的插入率 随机挑取 30 个独立克隆, 用质粒 pEVP3 多克隆位点两侧测序引物 P1, P2 进行 PCR 鉴定, pEVP3-*galU* 为对照。30 个克隆中, 有两个克隆 PCR 产物与对照在同一泳动位置, 表明没有随机片段插入, 其余 28 个克隆含有插入片段, 文库插入率达到 90% 以上, 表明文库质量较好。将此质粒库转化 *S. pn* 31203 *galU*⁻ 缺陷菌

株,得到约 450000 个 *S. pn* 转化菌落,收集这些菌落,即为可用 *galU* 随机报告 *S. pn* 体内诱导基因的菌株库。

2.2 体内诱导基因的筛选

经过小鼠腹腔感染模型的两轮筛选,我们得到了 165 个白色菌落,即包含只在体内表达基因的融合重组子,并用酶切自连法获得所有含体内诱导基因的重组质粒的克隆。

2.3 体内诱导基因的生物信息学分析

应用 BLAST 对这些基因片段进行了同源性分析,共发现 15 个不同的体内诱导基因片段,其中 9 个为已知基因,2 个为编码假想蛋白的基因,4 个为未知功能基因。对这 15 个体内诱导基因片段进行开放读码框(ORF)的搜索,总共发现 36 个开放阅读框,其中 8 个为单独的 ORF,7 个为含有多个 ORFs 的操纵元结构(表 1)。

表 1 IVET 鉴定的 15 个肺炎链球菌小鼠体内诱导基因

Table 1 15 <i>S. pn</i> i vi genes identified by IVET					
Function	Strain	Accession No.	Gene	Production	ORFs
energy metabolism	spn090	AE007344	<i>PTS-EII</i>	PTS system , sugar-specific IIC component	7
	spn137	AE007491	<i>ptcC</i>	PTS system , cellobiose-specific IIC component	4
	spn059	AE007479	<i>dexS</i>	putative dextran glucosidase	1
surface proteins	spn105	AE007351	<i>cbpF</i>	choline binding protein F	2
	spn151	AE007376	<i>zmpB</i>	putative zinc metalloprotease ZmpB	2
peptidoglycan synthesis in cell wall	spn005	AE007443	<i>pgdA</i>	peptidoglycan GlcNAc deacetylase A	1
	spn130	AE007349	<i>pbpA</i>	penicillin-binding protein 1A	2
DNA synthesis	spn011	AE007340	<i>polC</i>	DNA polymerase III , alpha subunit	1
material transport	spn145	AE007496	ABC-NP	ABC transporter , ATP-binding/membrane spanning	1
putative functions	spn026	AE007402		hypothetical protein 52% homology with endolysin of <i>Streptococcus thermophilus</i>	1
	spn114	AE007360		hypothetical protein 51% homology with Histone acetyltransferase HPA2 of <i>Streptococcus suis</i>	6
unknown functions	spn001	AE007482		hypothetical protein	5
	spn041	AE007356		hypothetical protein	1
	spn108	AE007457		hypothetical protein	1
	spn139	AE007480		hypothetical protein	1

2.4 体内、体外实验证实所筛选出的体内诱导基因只在体内表达

体外 β-半乳糖苷酶活性分析结果显示,15 个含体内诱导基因的肺炎链球菌菌株 OD₄₂₀ 读数均低于线性范围下限,说明这些菌株基本没有 β-半乳糖苷酶活性,与之在 X-gal 平板上呈现白色菌落一致,证实这些基因在体外不表达,将所筛选出的白色菌落体外增菌后,以同样浓度(5 × 10⁶)分别感染小鼠,从小鼠血中分离到 *S. pn*,而且在 X-gal + TSA 平板上为均己的白色菌落,证实这些基因在体内表达。

3 讨论

对于 IVET 而言,报告基因的选择是一个关键问题,它不仅会影响到筛选的可信度,还会影响到筛选的可行性。目前 IVET 中应用的报告基因主要有:管家基因、抗生素耐药基因、γδ 解离酶基因和绿色荧光蛋白基因。这几种报告基因各有优缺点:采用病原菌所必需的管家基因作为报告基因的 IVET 系统又可称为营养缺陷型,是由 Mahan 等^[4]最早建立,它

筛选严格,假阳性率低,但是由于并非所有的病原菌都存在减毒的营养缺陷型,限制了它的应用。采用宿主耐受的抗生素基因作为报告基因,可以对不存在营养缺陷型的病原菌进行筛选,并且调整抗生素的剂量可以分离到不同活性水平的体内诱导启动子,这一策略的缺点在于必须给足够浓度的抗生素给宿主动物并保证能到达感染部位,会干扰病原菌体内的感染过程。在严格的筛选压力下,管家基因和抗生素筛选所鉴定的体内诱导基因都是在感染过程中高水平表达的,这样就会漏掉一些弱表达或一过性表达的启动子。而采用 γδ 解离酶基因作为报告系统可解决这个问题,这一体系比前两种 IVET 系统敏感,并且可以检测到感染不同时刻,不同组织中细菌特异表达的基因,但是需要将解离酶的底物引入相应菌株的染色体。用绿色荧光蛋白作为报告基因的方法,是一种新的鉴定体内诱导基因的方法,它在分析弱启动子和瞬时表达的启动子方面优于其它 IVET,还能对单个细菌内基因的表达水平进行相对定量,但需要昂贵的分选细菌的流式细胞小颗粒分

选器,限制了它在国内的应用。

本研究中我们对 Mahan 等最早建立的 IVET 进行了改良,选择参与病原菌生物合成的基因作为报告基因。肺炎链球菌是有一种有荚膜的条件致病菌,1928 年著名的 Griffith 实验就证实荚膜多糖是肺炎链球菌的主要毒力因子,无荚膜的肺炎链球菌不能在小鼠体内存活。有研究表明在所有肺炎链球菌中都存在 *galU* 基因,其编码的 UDP-葡萄糖焦磷酸化酶(UDPG :PP)是荚膜多糖生物合成途径中的关键酶^[8]。利用荚膜缺陷菌株,只有报告基因 *galU* 被激活,病原菌才能在宿主体内存活的特点,我们构建了以 *galU-lacZ* 双重报告基因系统。从感染小鼠体内筛选到了 15 个体内诱导基因。

根据目前的研究资料,*S. pn* 的毒力因子主要有荚膜多糖、溶血素、表面粘附蛋白、自溶素、神经氨酸酶等。人们研究 *S. pn* 致病机制已有多,虽然鉴定了一些与毒力相关的基因,并且不断有新的毒力基因被鉴定,但是对其毒力因子的具体数量、本质和最终致病的分子机理还远不明了。根据对这些基因的生物信息学分析,为肺炎链球菌致病的分子机理带来了新的认识。定植与粘附是肺炎链球菌致病的首要步骤。与革兰阴性菌和其它链球菌不同,肺炎链球菌表面没有菌毛、纤毛等结构,因此它裸露的表面与宿主细胞相互作用的机制一直是研究者关注的焦点。*S. pn* 表面有多种蛋白质修饰,它们共价或非共价结合于细胞壁上,其中最独特的一类蛋白质是胆碱结合蛋白,它们就象 *S. pn* 表面胆碱和宿主细胞之间的“分子桥”,C 端锚定在胆碱上,N 端暴露在细菌表面,对 *S. pn* 的粘附、侵袭非常重要。在我们所鉴定的 *ivi* 基因片段中有一个胆碱结合蛋白基因(*cbpF*),虽然 *cbpF* 缺陷菌株无明显的表型改变^[9],但是它和上游 13 个 bp 处的 *cbpG* 在同一个操纵元结构中,受 *cbpG* 上游同一启动子的调控,因此这个启动子是在体内诱导表达的。*CbpG* 的这种在粘附和引起败血症中的双重作用以及它在体内诱导表达的特点提示这个蛋白可以作为肺炎链球菌新的疫苗候选者。

我们所鉴定的 *ivi* 基因 *zmpB*,编码的锌金属蛋白酶 *ZmpB*,可通过 LPXTGE 序列与细胞壁肽聚糖共价结合。有研究表明 *ZmpB* 可能参与肺炎链球菌表面蛋白的分选或靶向作用^[10],*zmpB* 缺陷的菌株表面 *CbpA*、*CbpE*、*CbpF*、*CbpJ*、*LytA* 表达缺失,*CbpA* 定位到胞浆,细菌粘附力下降,体外培养不能形成链状,失去自溶能力,自然转化能力也下降^[11]。

肽聚糖是革兰阳性菌细胞壁的主要成分,分裂形成新细胞,必然需要大量新合成的肽聚糖,其中青霉素结合蛋白(PBPs)参与了细胞壁肽聚糖的合成^[12]。现已明确肺炎链球菌基因组中有 6 个编码青霉素结合蛋白的基因,分别为 *pbpX*、*pbpA*、*pbp2b*、*pbp2a*、*pbp1bp* 和 *pbp3*^[13],其中 *pbpA* 是我们鉴定的 *ivi* 基因,表明细菌在宿主体内分裂过程中,需要 *PbpA* 参与合成细胞壁肽聚糖。

在哺乳动物体内存在大量的溶菌酶,细菌入侵宿主体内后,感染部位溶菌酶浓度急剧升高,可以破坏细菌细胞壁上的肽聚糖,而 *S. pn* 细胞壁含有高比例的己糖胺单位(>80%的葡萄糖胺和约 10%的胞壁酸)都没有被乙酰化,可以抵抗溶菌酶对肽聚糖的破坏作用^[14]。我们鉴定的 *ivi* 基因 *pgdA* 编码肽聚糖乙酰葡萄糖胺脱乙酰基酶 A,催化细胞壁肽聚糖乙酰葡萄糖胺的脱乙酰基作用,参与细胞壁肽聚糖抗溶菌酶作用。正是 *PgdA* 的脱乙酰基作用增强了细菌在宿主体内的毒力,使细菌能够抵抗宿主溶菌酶的攻击。

磷酸转移酶系统(PTS)广泛存在于各种细菌中,具有多样性,反映出细菌在生长过程中优先利用碳源的不同。在我们鉴定的 *ivi* 基因中包括 2 个含 PTS 的操纵元和葡聚糖苷酶(*dexS*),主要参与葡萄糖的磷酸化和水解。肺炎链球菌可以比其它细菌更高效地利用糖原,这与它的定植部位、生活方式有关。在体内鉴定的 2 个 PTS 系统分别为纤维二糖和蔗糖特异性,表明肺炎链球菌在小鼠体内可以高效利用这两种碳源,对它在体内存活比较关键。营养物质如氨基酸等的摄取是肺炎链球菌在宿主体内存活的关键的,我们所鉴定的 3 个 ABC 转运蛋白,具有转运细胞生长、代谢等基本生命活动所需物质的功能,但是这 3 个转运蛋白的作用底物目前还不是很清楚,它们在肺炎链球菌致病中的作用有待进一步的研究。

我们所鉴定的 15 个 *ivi* 基因中不包括一些已知的毒力因子,比如溶血素、自溶素等,分析可能的原因有:①一些毒力因子是肺炎链球菌体外生存必需的,在体外筛选中被剔除;②一些基因可能在感染的不同阶段表达具有时间特异性,一些基因的表达还具有组织特异性;③这种筛选方式严格,假阳性少,筛选到的都是高活性的诱导启动子,但是可能漏掉那些表达较弱的或一过性表达的启动子。

应用 IVET 鉴定的 *ivi* 基因不一定是病原菌的毒力基因,也就是说要对这些 *ivi* 基因进行功能分析才

能有助于阐明它们在致病过程中的重要性及更好地理解宿主-病原菌之间复杂的关系。因此,还需要通过基因缺失技术构建突变体,进一步分析其功能。我们在获得肺炎链球菌的 15 个 *ivi* 基因后,进一步的工作就是要通过插入失活方法构建这些基因的突变体,对其功能进行深入分析。

参 考 文 献

- [1] Balakrishnan I, Crook P, Morris R, *et al.* Early predictors of mortality in *Pneumococcal* bacteraemia. *J Infect*, 2000, **40**: 256 – 261.
- [2] Darkes MJ, Plosker GL. *Pneumococcal* conjugate vaccine (Pneumovax): a review of its use in the prevention of *Streptococcus pneumoniae* infection. *Paediatr Drugs* 2002 **4**: 609 – 630.
- [3] Su LC, David WH. In vivo genetic analysis of bacterial virulence. *Annu Rev Microbiol*, 1999 **53**: 129 – 154.
- [4] Mahan MJ, Slauch JM, Mekalanos JJ. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science*, 1993, **259**: 686 – 688.
- [5] Mei JM, Nourbakhsh F, Ford CW, *et al.* Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. *Mol Microbiol*, 1997 **26**: 399 – 407.
- [6] Orihuela CJ, Radin JN, Sublett JE, *et al.* Microarray analysis of pneumococcal gene expression during invasive disease. *Infect Immun*, 2004 **72**: 5582 – 5596.
- [7] Handfield M, Brady LJ, Progulske-Fox A, *et al.* IVIAT: a novel method to identify microbial genes expressed specifically during human infections. *Trends Microbiol*, 2000 **8**(7): 336 – 339.
- [8] Mollerach M, Lopez R, Garcia E. Characterization of the *galU* gene of *Streptococcus pneumoniae* encoding a uridine diphosphoglucose pyrophosphorylase: a gene essential for capsular polysaccharide biosynthesis. *J Exp Med*, 1998 **188**(11): 2047 – 2056.
- [9] Tuomanen EI, Measure HR, Gosink KK, *et al.* Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infect Immun*, 2000 **68**(10): 5690 – 5695.
- [10] Novak R, Charpentier E, Braun JS, *et al.* Extracellular targeting of choline-binding proteins in *Streptococcus pneumoniae* by a zinc metalloprotease. *Mol Microbiol* 2000 **36**(2): 366 – 376.
- [11] Berge M, Garcia P, Iannelli F, *et al.* The puzzle of *zmpB* and extensive chain formation, autolysis defect and non-translocation of choline-binding proteins in *Streptococcus pneumoniae*. *Mol Microbiol*, 2001 **39**(6): 1651 – 1660.
- [12] Macheboeuf P, Di Guilmi AM, Job V, *et al.* Active site restructuring regulates ligand recognition in class A penicillin-binding proteins. *Proc Natl Acad Sci USA* 2005 **102**(3): 577 – 582.
- [13] Sanbongi Y, Ida T, Ishikawa M, *et al.* Complete sequences of six penicillin-binding protein genes from 40 *Streptococcus pneumoniae* clinical isolates collected in Japan. *Antimicrob Agents Chemother*, 2004 **48**(6): 2244 – 2250.
- [14] Vollmer W, Tomasz A. The *pgdA* gene encodes for a peptidoglycan *N*-Acetylglucosamine deacetylase in *Streptococcus pneumoniae*. *J Biol Chem*, 2000 **275**(27): 20496 – 20501.

Screen and identification of *Streptococcus pneumoniae* genes specifically induced in host

MENG Jiang-ping, YIN Yi-bing*, ZHANG Xue-mei, HUANG Yuan-shuai, LAN Kai

(Key Laboratory of Laboratory Medical Diagnostics of Ministry of Education, Department of Laboratory Medicine in Chongqing University of Medical Sciences, Chongqing 400016, China)

Abstract: To identify in vivo-induced genes of *S. pneumoniae* and search new potential drug targets and vaccine candidates, a selection system was developed based on the in vivo expression technology (IVET). Promoter *galU* gene which is critical for the capsular polysaccharide biosynthesis and *lacZY* gene which encodes β -galactosidase were employed as dual reporter genes. The *galU*-deficient mutant of *S. pneumoniae* is incapable of utilizing galactose and synthesizing capsular polysaccharide, therefore can't survival in the host. Firstly, the random pieces of *S. pneumoniae* chromosomal DNA (200 ~ 500bp), obtained by partial *Sau3AI* restriction digestion, were subcloned into the *Bgl*II site of pEVP3-*galU*. Transformation by this plasmid library yielded promoter-trap library in *S. pneumoniae*. Then, the library was used to infect animals. Bacteria were harvested from lung tissue. White clones on TSA agar containing X-gal were used to reinfect animals. The sequence of infection and sorting was performed twice, 165 white clones harvested from the final round of infection were analyzed. A total of 15 unique sequences were obtained through in vivo screen. The bioinformatics analysis showed that these *ivi* genes involved in colonization and adherence, energy metabolism, nutrient substance transport, transcription regulation, DNA metabolism and cell wall synthesis. And there were some hypothetical proteins have unknown functions. Part of these genes may be related with virulence and can be used as vaccine candidates and drug targets.

Keywords: *Streptococcus pneumoniae*; in vivo expression technology; in vivo induced gene; Capsular polysaccharide

Foundation item: National Natural Sciences Foundation of China (30371275, 30400376)

* Corresponding author. Tel: 86-23-68485658; E-mail: yibingyin@21.cn.com

Received: 8 October 2005/Accepted: 19 October 2005/Revised: 29 March 2006