

用生物信息学方法预测猪链球菌 2 型 05ZYH33 株的脂蛋白

孙理云

(河南科技大学动物科技学院, 洛阳 471003)

摘要:【目的】鉴别已公布基因组序列的 SS2 05ZYH33 株的脂蛋白 (Lpp)。【方法】首先从 Genbank 获取由 SS2 05ZYH33 基因组推测的蛋白质氨基酸序列, 然后利用 Lpp 预测软件 PrositeScan 和 DOLOP 预测 05ZYH33 株的 Lpp, 采用 SignalP 3.0 HMM、PrediSi、Phobius、LipoP-HMM 和 TMHMM version2.0 分析预测的 Lpp 的信号肽, 然后经“多数票决法”确定 Lpp, 最后利用 InterProScan 和 BlastP 服务器对鉴定的 Lpp 进行功能分析。【结果】鉴定出 34 种 Lpp, 占 SS2 致病株 05ZYH33 蛋白质组的 1.555%。结构与功能分析结果表明, 最大的一类 Lpp 是 ABC 运输蛋白的底物结合蛋白, 共有 16 种, 其中 YP_001197586、YP_001197918、YP_001198449、YP_001199435、YP_001197482、YP_001199452 和 YP_001198914 等 7 种基因可与其他成分组成完整的 ABC 运输蛋白操纵子, 这些 Lpp 在 SS2 获取糖、氨基酸和金属离子等营养物质方面具有重要作用。YP_001197698 与无乳链球菌的黏附素 Lmb 和化脓链球菌的黏附素 Lbp 及 YP_001198710 与肺炎链球菌的 SlrA 高度同源, 推测它们与黏附功能有关, 为 SS2 新的毒力因子; 其他 Lpp 包括酶、参与蛋白质折叠过程的 Lpp 及功能未知的 Lpp。【结论】这些资料提示 Lpp 在 SS2 的生理和致病性方面具有重要作用。

关键词: 脂蛋白; 猪链球菌 2 型; 生物信息学; 毒力因子

中图分类号: Q816 文献标识码: A 文章编号: 0001-6209 (2008) 08-1104-06

脂蛋白 (Lpp) 是一类广泛存在于细菌表面的蛋白质, 对细菌来讲不但具有多种生理作用, 而且作为病原相关分子模式 (pathogen-associated molecular patterns) 也可以被宿主天然免疫中 TLR 家族成员所识别, 诱导机体的天然免疫^[1], 更有一些细菌的 Lpp 作为毒力因子在致宿主发病方面起着重要作用^[2-6], 因此细菌 Lpp 的鉴别具有重要意义。

Lpp 信号肽由 N 区、H 区和 C 区 3 部分组成。N 区一般由 5~7 个氨基酸组成, 其特征是存在正电荷的氨基酸 (K 和/或 R), H 区由疏水性氨基酸组成, C 区具有 Lpp 特征性保守基序-脂盒, 该基序包括成熟 Lpp +1 位置的半胱氨酸 (C)。N 区和 H 区为 Sec 蛋白转运途径成分识别, 而 C 区为 Lpp 的脂化酶、信号肽酶所识别。基于 Lpp 信号肽的这些特征, 人们已

经建立了多种预测 Lpp 的方法^[7-11]。

猪链球菌 2 型 (*Streptococcus suis* type 2, SS2) 是革兰氏阳性、溶血性兼性厌氧的球菌, 主要可引起猪和人的脑膜炎、关节炎、败血症及猝死, 是一种重要的人兽共患病细菌。我国分别在 1998 年江苏省和 2005 年四川省在猪群和人群中暴发流行猪链球菌病并引起死亡, 引起了社会的广泛注意。引起这两次暴发疫情的病原体均为 SS2 强毒株^[12]。但其致病机制及毒力因子尚不甚清楚。虽然 Greeff 等^[13]报道 SS2 具有 Lpp 信号肽酶, 但 SS2 有哪些 Lpp, 其具有何种作用尚未见报道。为此, 本文采用生物信息学法对已公布的国内 SS2 致病株 05ZYH33 基因组进行分析, 推测出该菌株具有 34 种可能的 Lpp 并对其可能的作用进行了分析, 为进一步的研究提供了基础。

基金项目: 河南科技大学人才科研基金(05-010); 河南科技大学校科学研究基金(2006ZY042)

作者简介: 孙理云(1969-), 男, 河南省封丘县人, 博士, 讲师, 研究方向为预防兽医学。Tel: +86-579-64282341; E-mail: sunliyun8891@sina.com.cn

收稿日期: 2008-02-27; 修回日期: 2008-05-09

1 材料和方法

1.1 材料

SS2 致病株 05ZYH33 基因组推测的蛋白质及氨基酸组成序列由 <http://www.ncbi.nlm.nih.gov/entrez/> 获得。该菌基因组登陆号为 NC_009442, 编码 2186 种蛋白质。

1.2 Lpp 的鉴别与验证

分别利用 Prosite 程序的 PrositeScan(<http://www.expasy.org/prosite/>)^[14] 和 DOLOP 分析程序 (<http://www.mrc-lmb.cam.ac.uk/genomes/dolop/>)^[9] 对获得的蛋白质进行分析, 预测出 Lpp。然后对预测的蛋白用 SignalP 3.0 HMM (<http://www.cbs.dtu.dk/services/SignalP/>)^[15]、PrediSi(<http://www.predisi.de/>)^[16]、Phobius (<http://phobius.cgb.ki.se/>)^[17]、LipoP-HMM(<http://www.cbs.dtu.dk/services/LipoP/>)^[8]、TMHMM version2.0(<http://www.cbs.dtu.dk/services/TMHMM-2.0/>)^[18] 生物信息学工具进行分析, 最后采用“多数票决法”确认 Lpp^[19]。

1.3 脂蛋白功能的生物信息学分析

利用 InterProScan (<http://www.ebi.ac.uk/InterProScan/>) 程序检测每一 Lpp 的功能域, 利用 BlastP (<http://beta.uniprot.org/?tab=blast>) 进行比对, 同时检测其保守域。

2 结果

2.1 Lpp 的鉴定

利用 PrositeScan 鉴别出 35 个 Lpp, 利用 DOLOP 分析程序鉴别出 32 个 Lpp, 综合二结果共鉴别出 38 个 Lpp (见表 1)。

根据“多数票决法”判定 YP_001199014、YP_001197544、YP_001198844 和 YP_001198946 四种 Lpp 属于假阳性, 故预测数目为 34 (见表 1), 占 SS2 推测蛋白组的 1.555%。

2.2 Lpp 信号肽分析

约 88% 预测的 Lpp 脂盒序列符合一致序列 [L][ASTGV][AG]C。信号肽长度在 17~36 个 aa 之间, 平均约 22.62±4.75aa, 与 Sutcliffe 和 Harrington 等^[7] 报告的 G⁺细菌 Lpp 的长度相一致。H 区长度在 7~14aa 之间, 多数为 11 (8 种) 和 13 个 aa (16 种); N 区长度在 3~23aa 之间, 平均 7.76±5.16aa; 四个预测的 Lpp (YP_001199452、YP_001198743、YP_001199119 和 YP_001199210) 信号肽之所以长是由于 N 区长度所致。

2.3 Lpp 的功能分析

34 种 Lpp 按照功能可分为 4 类, 16 种 ABC 运输蛋白的底物结合蛋白 (SBP), 8 种功能未知的假定蛋白, 7 种酶和 3 种其他功能的 Lpp (见表 2)。

2.3.1 SBP: 在鉴别的 16 种 SBP 中, 其中 YP_001198914 和 YP_001198622 在原注释为假定蛋白, YP_001198623 原注释为 ABC 运输蛋白的通透酶。检测 YP_001198914, 发现其具有保守域 COG0683 (ABC 型支链氨基酸运输系统周质成分), PR00337 (Leu/Ile/Val 结合蛋白域), 与无乳链球菌 2603V/R 的支链氨基酸 ABC 运输蛋白氨基酸结合蛋白 62% (239/385) 一致 (identities), 结合其基因组的上下背景, 推测 YP_001198914 为 SBP。同样检测 YP_001198622 和 YP_001198623, 它们具有保守域 pfam04392 (ABC 运输蛋白底物结合蛋白域), 二者 53% (157/291) 一致; 与多种链球菌的 ABC 运输蛋白结合蛋白同源, 结合其基因组的上下背景, 推测它们为 SBP。

检查在基因组的上下背景, 可发现 YP_001197586、YP_001197918、YP_001198449、YP_001199435、YP_001197482、YP_001199452 和 YP_001198914 等 7 个 SBP 基因与其相应通透酶和 ATPase 基因组成完整 ABC 运输蛋白的操纵子。

YP_001197698 和 YP_001198383 均为“孤儿 SBP”(没有通透酶和 ATP 结合成分与之组成 ABC 运输蛋白, 仅有 SBP, 故名。), YP_001197698 注释为 ABC 型运输系统的周质成分/表面黏附素, 有 FPrintScan 数据库的 PR00690 和 PR00691 黏附素结构域, 与无乳链球菌的层连蛋白结合蛋白(Lmb)54% (167/306) 一致, 化脓性链球菌的层连蛋白结合蛋白(Lbp) 55.7% (167/306) 一致; YP_001198383 注释为 ABC 型氨基酸转运/信号转导系统的周质成分/域, 与嗜热链球菌的氨基酸 ABC 运输蛋白的 SBP52% 一致, 因而可能用于感知环境的氨基酸, 与信号的传导有关。

YP_001198472 是假基因 (编码的蛋白质功能结构域不完整, 不能执行正常的基因功能), YP_001198622、YP_001198623 基因邻近的通透酶基因为假基因, YP_001198704、YP_001198924、YP_001199280 和 YP_001199499 等 4 种 SBP 基因邻近没有 ATPase 成分基因, 均构不成完整的 ABC 运输蛋白操纵子。

表 1 Lpp 信号肽的确认

Table 1 Validation of Lpp signal peptides using various predictors of signal peptide features and a 'majority vote' approach to data interpretation

Accession No.	UniProt	DOLOP	SignalP analysis				Predisci	Phobius		LipoP analysis			Verdict	
	Y/N	Y/N	Cys Position	End of H-region	SP Prob	Cleavage site prob	Score	End of H-region	End of C-region	Best Prediction	Score	Margin		2nd Choice
YP_001197586	Y	Y	26	S27	0.999	0.560	0.726	23	28	SpII	12.91	9.73	TMH	Lpp
YP_001197693	Y	Y	28	L36	0.989	0.829	0.508	TMH		SpI	4.61	0.12	SpII	Lpp
YP_001197698	Y	Y	21	C21	0.999	0.776	0.614	18	23	SpII	18.49	11.32	SpI	Lpp
YP_001197780	Y	Y	29	L30	0.999	0.978	0.521	TMH		SpI	7.01	3.74	TMH	Lpp
YP_001197891	Y	Y	19	F15	0.756	0.448	0.43	20	25	SpII	16.33	13.61	SpI	Lpp
YP_001197918	Y	Y	24	C24	1.000	0.525	0.754	17	22	SpII	23.77	7.05	SpI	Lpp
YP_001198066	Y	Y	23	T19	0.976	0.797	0.396	18	23	SpII	14.96	2.73	SpI	Lpp
YP_001198111	Y	Y	21	C21	0.995	0.678	0.718	18	23	SpII	28.80	14.61	SpI	Lpp
YP_001198222	Y	Y	21	C21	1.000	0.898	0.6	16	21	SpII	24.36	7.98	SpI	Lpp
YP_001198261	Y	Y	19	S20	1.000	0.477	0.707	19	24	SpII	27.18	11.74	SpI	Lpp
YP_001198383	Y	Y	20	C20	1.000	0.730	1.000	15	20	SpII	25.18	8.47	SpI	Lpp
YP_001198449	Y	Y	21	C21	1.000	0.253	0.823	16	21	SpII	27.10	12.17	SpI	Lpp
YP_001198457	Y	Y	21	G23	0.997	0.714	0.609	18	23	SpII	19.19	15.21	SpI	Lpp
YP_001198472	Y	Y	23	C23	1.000	0.437	0.554	15	19	SpII	24.31	11.11	SpI	Lpp
YP_001198622	Y	Y	24	S26	1.000	0.751	1	20	25	SpII	24.28	13.71	SpI	Lpp
YP_001198623	Y	Y	22	C22	1.000	0.658	0.845	17	22	SpII	23.42	14.28	SpI	Lpp
YP_001198704	Y	Y	24	G26	0.999	0.588	0.852	22	26	SpII	20.97	13.56	SpI	Lpp
YP_001198710	Y	Y	18	T19	1.000	0.505	0.955	15	33	SpII	33.21	12.11	SpI	Lpp
YP_001198847	Y	Y	19	V17	0.996	0.487	0.52	19	24	SpII	10.27	3.00	SpI	Lpp
YP_001198914	Y	Y	22	G23	1.000	0.764	1	18	22	SpII	24.70	7.84	SpI	Lpp
YP_001198924	Y	Y	22	C22	1.000	0.417	0.929	20	25	SpII	19.22	5.33	SpI	Lpp
YP_001199014	Y	Y	26	V29	0.008	0.004	0	TMH		SpII	3.55	0.64	TMH	false +ve
YP_001199120	Y	Y	23	G24	1.000	0.467	0.601	20	24	SpII	25.02	16.28	SpI	Lpp
YP_001199209	Y	Y	24	S28	1.000	0.862	1	18	22	SpII	33.32	10.77	SpI	Lpp
YP_001199280	Y	Y	23	T27	1.000	0.924	0.976	18	23	SpII	18.09	6.00	SpI	Lpp
YP_001199380	Y	Y	21	C21	1.000	0.667	0.813	21	28	TMH	12.30	0.69	SpII	Lpp
YP_001199435	Y	Y	20	A19	0.998	0.503	0.734	13	18	SpII	22.13	7.98	SpII	Lpp
YP_001199499	Y	Y	22	C22	1.000	0.465	0.914	15	20	SpII	30.91	15.68	SpI	Lpp
YP_001199520	Y	Y	26	C26	1.000	0.407	0.563	23	27	SpII	24.06	15.20	SpI	Lpp
YP_001197482	Y	N	19	A18	1.000	0.459	0.725	19	24	SpII	23.20	6.23	SpI	Lpp
YP_001198234	Y	N	21	C22	1.000	0.708	0.738	19	24	SpII	25.40	10.79	SpI	Lpp
YP_001198844	Y	N	28	Q59	0.847	0.464	0.745	TMH		THM	13.29	4.81	SpI	false +ve
YP_001198946	Y	N	28	N28?	0.018	0.013	0.194	Non cyto-plasmic		CYT	-0.20		none	false +ve
YP_001199452	Y	N	30	R31	1.000	0.893	0.595	25	30	SpII	14.88	1.80	SpI	Lpp
YP_001198743	N	Y	37	Y41	1.000	0.973	0.52	34	40	SpII	17.45	3.25	SpI	Lpp
YP_001199119	N	Y	34	V32	0.914	0.455	0.473	31	35	SpII	25.02	16.28	SpI	Lpp
YP_001199210	N	Y	36	C36	0.154	0.106	0.656	35	43	SpII	4.64	4.84	CYT	Lpp

Note: Results considered inconsistent with a Lpp signal peptide are highlighted in bold and the final verdict based on the 'majority vote' principle is indicated in the final column.

表 2 鉴定的 Lpp 功能分类
Table 2 Functional categorization of putative Lpp identified in the SS2 genome

category	Accession No.	Putative function
SBPs	YP_001197698	ABC-type metal ion transport system, periplasmic component/surface adhesin
	YP_001197482	ABC-type metal ion transport system, periplasmic component/surface adhesin
	YP_001199452	High-affinity zinc uptake system protein znuA precursor
	YP_001197918	amino acid ABC transporter, amino acid-binding protein
	YP_001198383	ABC-type amino acid transport/signal transduction system, periplasmic component/domain
	YP_001199435	ABC-type amino acid transport/signal transduction system, periplasmic component/domain
	YP_001198914	ABC-type branched-chain amino acid transport systems, periplasmic component
	YP_001198472	ABC-type phosphate transport system, periplasmic component
	YP_001198704	ABC-type sugar transport system, periplasmic component
	YP_001198924	ABC-type sugar transport system, periplasmic component
	YP_001199280	ABC-type sugar transport system, periplasmic component
	YP_001199499	ABC transporter substrate-binding protein - maltose/maltodextrin
	YP_001198622	ABC transporter substrate binding protein
	YP_001198623	ABC transporter substrate binding protein
	YP_001198449	Uncharacterized ABC-type transport system, periplasmic component/surface lipoprotein
YP_001197586	ABC-type nitrate/sulfonate/bicarbonate transport system, periplasmic component	
Enzymes	YP_001197693	lipase
	YP_001197780	penicillin-binding protein
	YP_001198111	D-alanyl-D-alanine carboxypeptidase
	YP_001198710	PPIase - cyclophilin family
	YP_001198066	ATPase (PilT family)
	YP_001199520	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit
	YP_001198457	Uncharacterized protein involved in cytokinesis, contains TGc (transglutaminase/protease-like) domain
Hypotheticals	YP_001198222	hypothetical protein
	YP_001198743	hypothetical protein
	YP_001197891	hypothetical protein
	YP_001198261	hypothetical protein
	YP_001198847	hypothetical protein
	YP_001198234	hypothetical protein
	YP_001199210	hypothetical protein
	YP_001199380	hypothetical protein
Other	YP_001199119	Membrane-associated lipoprotein involved in thiamine biosynthesis
	YP_001199120	Major membrane immunogen, membrane-anchored lipoprotein
	YP_001199209	Clostridium cellulosome enzyme, dockerin type

2.3.2 起酶作用的 Lpp: 在鉴别出 Lpp 中, 7 种预测具有酶活性。其中 YP_001197693 与一种未培养细菌的脂酶 39% (44/111) 一致^[20]。脂酶大多存在微生物细胞内, 因为需催化胞内脂肪的合成及分解, 或者分泌胞外酶利用胞外的脂质。而在 SS2 致病株 05ZYH33 中预测的脂酶位于细菌细胞膜上, 其有何种生理方面或致病作用, 有待试验证明。YP_001197780 与多种链球菌的青霉素结合蛋白 1A 有较高的一致性, N 端具有转糖基酶活性的 PF00912 结构域, C 端具有结合青霉素和转肽酶活性的 PF00905 结构域。YP_001198111 预测为 D-丙氨酰-D-丙氨酸-羧肽酶, 具有 pfam02557 (D, D-羧肽酶 VanY 家族)。在 InterProScan 的多个数据库中均发现 YP_001198710 有亲环素家族的肽基脯氨酸顺反异构酶(PPIase)域, PPIase 是一种伴侣分子, 与蛋白质的折叠有关; 该蛋白与肺炎链球菌的 PPIase58% (156/268) 一致。YP_001198066 与 *Streptococcus sanguinis* SK36 的 ATPase (PilT family)73% (147/200) 一致。YP_001199520 具有保守域 COG1053 (SdhA,

琥珀酸脱氢酶/延胡索酸还原酶的黄素蛋白亚单位), 与 Pfam 数据库的 HI0933-样蛋白家族相匹配, HI0933-样蛋白家族包括具有二核苷结合基序的氧化还原酶和脱氢酶。YP_001198457 有保守域 smart00460 (转谷氨酰胺酶/蛋白酶样同源物), 与猪链球菌 89/1591 的转谷氨酰胺酶样蛋白 97% (398/409) 一致。

2.3.3 功能未知的假定蛋白: YP_001198222、YP_001198743、YP_001197891、YP_001198261、YP_001198847 和 YP_001198234 均未发现保守域有相匹配的蛋白结构域, 功能未知。YP_001198222 原注释为翻译起始因子 1, YP_001198743 为葡糖 6 磷酸异构酶。YP_001198222、YP_001198743、YP_001197891、YP_001198261 和 YP_001198847 与猪链球菌 89/1591 的假定蛋白 SsuiDRAFT_1709、SsuiDRAFT_1306、SsuiDRAFT_0365、SsuiDRAFT_1731 和 SsuiDRAFT_1191 分别 99% (248/249)、97% (254/260)、84% (84/99)、84% (105/124) 和 100% (134/134) 一致; YP_001198234 与猪链球菌 98HAH33 的假定蛋白 SSU98_0871 100% (279/279) 一致。

YP_001199210 与猪链球菌 89/1591 的大肠菌素裂解蛋白 97% (217/223) 一致, 在 N 端均具有与 Pfam 数据库的 PF02402 大肠菌素裂解蛋白家族相匹配的域, 脂蛋白酶的裂解位置位于其中。大肠菌素裂解蛋白是一类小分子脂蛋白, 成熟的蛋白由 28~33 个氨基酸组成, 为大肠杆菌释放大肠菌素所需^[21]。

YP_001199380 具有保守域 PRK02944 (OxaA 样蛋白前体) 和与 Pfam 数据库的 60 kDa 内膜蛋白家族相匹配的域, 与 *Streptococcus gordonii* str. Challis substr. CH1 的膜蛋白 OxaA1 前体 68% (184/270) 一致, 60 kDa 内膜蛋白是伴侣分子, 参与新合成的内在蛋白插入细胞膜过程^[22]。

2.3.4 其他: YP_001199119 在 05ZYH33 基因组注释为与硫胺素合成有关的膜连 Lpp, 有 PF02424 结构域, 与猪链球菌 89/1591 的 ApbE 样蛋白 99% (347/349) 一致, ApbE 是鼠伤寒沙门氏菌的一种 Lpp, 与硫胺素的合成有关^[23]。

另外一个 Lpp YP_001199120 具有保守域 COG4939 为膜主要免疫原/膜锚定 Lpp。该蛋白与 Pfam 数据库的黄素单核苷酸 (FMN) 结合蛋白家族匹配。

YP_001199209 没有相匹配的结构域, 但与热纤梭菌 ATCC27405 的纤维素小体酶的型锚定素有 37% 的一致性。

3 讨论

Lpp 是一类共价结合于细菌表面的重要蛋白, 不但具有多种重要的生理作用, 还可以刺激活化宿主的多种免疫细胞产生促炎因子和抗体, 可用于开发疫苗和佐剂, 因而备受人们关注^[24-25]。已有试验证实, 一些链球菌的 Lpp 具有免疫保护作用^[26-27]。随着越来越多细菌基因组序列测定的完成, 预测鉴别出其 Lpp 具有重要的意义。

目前用于预测 Lpp 的程序有 G+LPP^[6]、Lipo P^[7]、DOLOP^[8]、Lipped^[9]和 PrositeScan^[13], 它们都是在分析试验确证的 Lpp 序列的基础上得出的预测规则, 由于分析的数据不同, 得出的预测规则稍有差异, 加之采用的算法不同, 对同一基因组的预测结果不同。这由我们的结果可得到验证。这些方法因为未能区分 N 端的疏水区是分泌蛋白信号肽的 H 区或是跨膜蛋白的跨膜区, 因而易造成结果假阳性出现, 而本试验在综合采用 DOLOP 和 UniProt 预测 Lpp 的基础上, 结合 SignalP 3.0 HMM、PrediSi、Phobius、LipoP 和 TMHMM, 对预测的 N 端的疏水区进行综合判定, 不

但可防止漏掉一些 Lpp, 也可以排除假阳性结果, 提高了预测结果的准确性。

由推测的功能看出, SS2 的 Lpp 具有重要的生理作用。它们有的负责营养物质的吸收 (YP_001197918、YP_001198449、YP_001199435、YP_001197482、YP_001199452 和 YP_001198914 等 7 种属 SBP 的 Lpp), 有的参与参与细菌细胞壁的合成 (如 YP_001197780 和 YP_001198111), 有的参与物质代谢 (如 YP_001199119)。

无乳链球菌的 Lmb、化脓性链球菌的 Lbp 和肺炎链球菌的 SlrA 均为 Lpp, 为相应细菌的毒力因子, Lmb 具有介导无乳链球菌黏附人层连蛋白^[2], Lbp 促进侵袭人脑毛细血管内皮细胞的作用^[3]及化脓性链球菌黏附上皮细胞和内化化脓性链球菌的作用^[4-5]; SlrA 具有促进肺炎链球菌黏附和侵袭宿主细胞作用, 抵抗细胞吞噬作用^[6]。SS2 的 YP_001197698 与 Lmb 和 Lbp 同源, YP_001198710 与 SlrA 同源, 因此推测二者是 SS2 的毒力因子, 具有黏附宿主的细胞外基质和细胞等作用。

参 考 文 献

- [1] Brightbill HD, Libraty H, Krutzik SR, et al. Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. *Science*, 1999, 285 (5428): 732-736.
- [2] Spellerberg B, Rozdzinski E, Martin S, et al. Lmb, a protein with similarities to the LraI adhesin family, mediates attachment of *Streptococcus agalactiae* to human laminin. *Infect Immun*, 1999, 67(2): 871-878.
- [3] Tenenbaum T, Spellerberg B, Adam R, et al. *Streptococcus agalactiae* invasion of human brain microvascular endothelial cells is promoted by the laminin-binding protein Lmb. *Microbes Infect*, 2007, 9(6): 714-720.
- [4] Terao Y, Kawabata S, Kunitomo E, et al. Novel laminin-binding protein of *Streptococcus pyogenes*, Lbp, is involved in adhesion to epithelial cells. *Infect Immun*, 2002, 70(2): 993-997.
- [5] Elsner A, Kreikemeyer B, Braun-Kiewnick A, et al. Involvement of Lsp, a member of the LraI-lipoprotein family in *Streptococcus pyogenes*, in eukaryotic cell adhesion and internalization. *Infect Immun*, 2002, 70(9): 4859-4869.
- [6] Hermans PW, Adrian PV, Albert C, et al. The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization. *J Biol Chem*, 2006, 281(2): 968-976.
- [7] Sutcliffe IC, Harrington DJ. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiol*, 2002, 148 (7): 2065-2077.
- [8] Juncker AS, Willenbrock H., von Heijne G, et al. Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci*, 2003, 12 (8): 1652-1662.
- [9] Babu MM, Priya ML, Selvan AT, et al. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol*, 2006, 188: 2761-2773.
- [10] Taylor PD, Toseland CP, Attwood TK et al. LIPPRED: A web server for accurate prediction of lipoprotein signal sequences and

- cleavage sites. *Bioinformatics*, 2006, 1 (5):176–179.
- [11] Setubal J C, Reis M, Matsunaga J, et al. Lipoprotein computational prediction in spirochaetal genomes. *Microbiol*, 2006, 152 (1): 113–121.
- [12] Tang JQ, Wang CJ, Feng YJ, et al. Streptococcal toxic shock syndrome caused by *Streptococcus suis* serotype 2. *PLoS Med*, 2006, 3 (5): 151.
- [13] de Greeff A, Hamilton A, Sutcliffe IC, et al. Lipoprotein signal peptidase of *Streptococcus suis* serotype 2. *Microbiol*, 2003, 149 (6): 1399–1407.
- [14] Wu CH, Apweiler R, Bairoch A, et al. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res*, 2006, 34: D187–D191.
- [15] Bendtsen JD, Nielsen H, von Heijne G, et al. Improved prediction of signal peptides: SignalP 3.0. *J Mol Bio*, 2004, 340 (4): 783–795.
- [16] Hiller K, Grote A, Scheer M, et al. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res*, 2004, 32: W375–W379.
- [17] Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol*, 2004, 338 (5): 1027–1036.
- [18] Krogh A, Larsson B, von Heijne G, et al. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol*, 2001, 305 (3): 567–580.
- [19] Tjalsma H, van Dijk JM. Proteomics-based consensus prediction of protein retention in a bacterial membrane. *Proteomics*, 2005, 5 (17): 4472–4482.
- [20] Henne A, Schmitz RA, Bömeke M, Gottschalk G, et al. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl Environ Microbiol*, 2000, 66(7): 3113–3116.
- [21] Cavard D, Baty D, Howard SP, et al. Lipoprotein nature of the colicin A lysis protein: effect of amino acid substitutions at the site of modification and processing. *J Bacteriol*, 1987, 169(5): 2187–2194.
- [22] Kuhn A, Stuart R, Henry R, et al. The Alb3/Oxa1/YidC protein family: membrane-localized chaperones facilitating membrane protein insertion? *Trends Cell Biol*. 2003, 13(10): 510–516.
- [23] Downs DM, Beck BJ. The *apbE* gene encodes a lipoprotein involved in thiamine synthesis in *Salmonella typhimurium*. *J Bacteriol*, 1998, 180(4): 885–891.
- [24] Lei B, Liu M, Chesney GL, et al. Identification of new candidate vaccine antigens made by *Streptococcus pyogenes*: purification and characterization of 16 putative extracellular lipoproteins. *J Infect Dis*, 2004, 189 (1): 79–89.
- [25] Spohn R., Buwitt-Beckmann U, Brock R, et al. Synthetic lipopeptide adjuvants and Toll-like receptor 2—structure–activity relationships. *Vaccine*, 2004, 22(19): 2494–2499.
- [26] Aranda J, Garrido ME, Cortés P, et al. Analysis of the protective capacity of three *Streptococcus suis* proteins induced under divalentation-limited conditions. *Infect Immun*, 2008, 76(4): 1590–1598.
- [27] Oliveira MLS, Arêas APM, Campos IB, et al. Induction of systemic and mucosal immune response and decrease in *Streptococcus pneumoniae* colonization by nasal inoculation of mice with recombinant lactic acid bacteria expressing pneumococcal surface antigen A. *Microbes Infect*, 2006, 8(4): 1016–1024.

Putative lipoproteins of *Streptococcus suis* type 2 identified by bioinformatic genome analysis

Liyun Sun*

(College of Livestock Science & Technology, Henan University of Science and Technology, Luoyang 473001, China)

Abstract: [Objective] To identify putative lipoproteins of *Streptococcus suis* serotype 2 (SS2). [Methods] First, I used the ScanProsite feature of Prosite and the website Database Of bacterial LipOProteins(DOLOP) to identify putative lipoproteins in the recently published genome of SS2 strains 05ZYH33, then validated the identified putative lipoproteins to exclude the false positive by a ‘majority vote’ approach following analysis with standard tools for signal peptide verification. Finally, putative functions were attributed to individual lipoproteins by reference to the identification of conserved domains of InterPro. Homologous proteins were identified by unfiltered BlastP homology searches (including conserved domain detection). [Results] Among the 38 identified putative lipoproteins, 34 were validated as lipoprotein and 4 as false positive lipoproteins. The largest functional category (16/34, 44%) of lipoproteins was that predicted to comprise of substrate binding proteins of ATP-binding cassette transport systems. Other roles included lipoproteins that participated in adhesion (YP_001197698 and YP_001198710), protein export and folding. [Conclusion] Lipoproteins contributed to the physiology of SS2 and influenced its virulence.

Keywords: lipoproteins; *Streptococcus suis* type 2; bioinformatics; virulence factors

Supported by the Scientific Research Foundation for Talents in Henan University of Science and Technology (05-010) and the Scientific Research Foundation of Henan University of Science and Technology (2006ZY042)

*Corresponding author. Tel: +86-579-64282341; E-mail: sunliyun8891@sina.com.cn

Received: 27 February 2008/ Revised: 9 May 2008