

doi:10.3969/j.issn.1672-5565.2016.02.01

# 鮰爱德华菌外膜蛋白 *OmpLC* 基因的生物信息学分析

黄艳青<sup>1</sup>, 刘港彪<sup>2</sup>, 王利<sup>2\*</sup>

1. 农业部东海与远洋渔业资源开发利用重点实验室(中国水产科学研究院东海水产研究所), 上海 200090;
2. 国家民委-教育部重点实验室(西南民族大学生命科学与技术学院), 成都 610041)

**摘要:**采用 PCR 方法从鮰爱德华菌基因组中扩增出外膜蛋白 *OmpLC* 基因, 利用生物信息学相关软件和网络数据库, 预测该基因编码产物的基本理化性质、亲疏水性、信号肽、跨膜性、二级结构、结构域及基序, 以及三级结构, 同时构建 *OmpLC* 同源基因的系统发育进化树。结果表明: 该蛋白由 360 个氨基酸组成, 分子量为 39.407 kD, 理论等电点为 4.98, 不稳定系数为 18.26, 是一种稳定的强亲水性蛋白, 有信号肽, 成熟蛋白无跨膜螺旋区。其二级结构中  $\alpha$  螺旋、 $\beta$  折叠和无规则卷曲分别占 6.67%、45.28% 和 48.06%, 其空间结构为  $\beta$  桶状, 属于 OM\_channels superfamily 的 gram\_neg\_porins 成员。蛋白质多重序列比对和聚类分析显示, 该蛋白序列与 *OmpLC* 蛋白(AEQ59632, AEQ59639) 具有高度同源性, 且在系统发育树上与二者聚为一簇。鮰爱德华菌 *OmpLC* 的生物信息学分析不仅为进一步探索该蛋白的功能提供参考资料, 也为研究鮰爱德华菌的感染和致病机理, 研制相关疫苗提供理论依据。

**关键词:** 鮰爱德华菌; 外膜蛋白; 结构预测; 生物信息学

中图分类号: S968.25 文献标志码: A 文章编号: 1672-5565(2016)02-061-10

## Bioinformatics analysis of the *Edwardsiella ictaluri* *OmpLC* gene

HUANG Yanqing<sup>1</sup>, LIU Gangbiao<sup>2</sup>, WANG Li<sup>2\*</sup>

1. Key Laboratory of East China Sea and Oceanic Fishery Resources Exploitation, Ministry of Agriculture (East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences), Shanghai 200090, China;
2. Key Laboratory of Animal Genetics and Breeding of State Ethnic Affairs Commission & Ministry of Education (College of Life Science and Technology, Southwest University for Nationalities), Chengdu 610041, China)

**Abstract:** The purpose of this study is to predict and analyze the structure and function of out membrane protein LC (*OmpLC*) of *Edwardsiella ictaluri*. The *OmpLC* gene was amplified by PCR method. The characteristics of physico-chemical parameters, hydropathy profile, signal peptide, transmembrane helices, secondary structure and tertiary structure were analyzed by bioinformatics softwares and web servers. The results showed that *OmpLC* protein of *Edwardsiella ictaluri* composed of 360 amino acids, the calculated molecular mass was 39.407 kD and the theoretical isoelectric point was 4.98, and the instability index was 18.26. *OmpLC* protein had signal peptide and strong hydrophilicity but without transmembrane helices in mature protein. In the second structure  $\alpha$ -helix,  $\beta$ -sheet and random coil made up 6.67%, 45.28%, 48.06%, respectively. The three-dimensional structure was  $\beta$ -barrel, its main domain was classified to gram\_neg\_porins which belong to OM\_channels superfamily. The multiple alignment and clustering analysis showed that *OmpLC* protein of *Edwardsiella ictaluri* has high similarity to other *OmpLC* proteins (AEQ59632, AEQ59639). The properties of *OmpLC* obtained by bioinformatic analysis can provide reference for the prediction of novel functional domain and the research on related pathogenesis and new vaccines.

**Keywords:** *Edwardsiella ictaluri*; *OmpLC*; Structure prediction; Bioinformatics

收稿日期: 2016-03-23; 修回日期: 2016-05-02.

基金项目: 国家自然科学基金面上项目 (No. 31172421); 四川省科技支撑项目 (No. 2016NZ0044); 西南民族大学研究生学位点建设项目 (No. 2016XWD-SO71007)。

作者简介: 黄艳青, 女, 博士, 研究方向: 水产动物营养与饲料学及分子生物学; E-mail: hyqrch@163.com.

\* 通信作者: 王利, 女, 副教授, 硕士生导师, 研究方向: 水产动物疾病学; E-mail: wangli@swun.cn.

鮠爱德华菌 (*Edwardsiella ictaluri*) 属肠杆菌科 (*Enterobacteriaceae*) 爱德华菌属 (*Edwardsiella*) 的一种革兰氏阴性短杆菌, 该菌主要感染鲶形目鱼类<sup>[1]</sup>。已有报道该菌曾引发黄颡鱼的红头病、斑点叉尾鮠肠型败血症 (Enteric septicemia of catfish, ESC)<sup>[2-3]</sup>, 在其它非鲶形目鱼类也可以被感染, 如在虹鳟鱼、鳖也见报道<sup>[4-5]</sup>。鮠爱德华氏菌病具有发病急、发病率和死亡率高的特点, 随着养殖规模和养殖密度的加大, 该菌对淡水鱼类的养殖危害日益严重, 但目前控制该菌引起的相关疾病主要以抗生素为主, 存在细菌耐药性增强、水体药物污染、引发食品安全等问题。因此, 研发相关中草药制剂和高效疫苗越发重要。

外膜蛋白 (Outer membrane proteins, OMPs) 是革兰氏阴性细菌外膜的主要结构成分, 是外膜 A 蛋白、微孔蛋白、脂蛋白和微量蛋白等蛋白的总称<sup>[6]</sup>, OMPs 除作为细菌的结构组成之外, 对细胞物质的运输和有关物质的合成和致病力也有重要作用。研究表明外膜蛋白具有良好的免疫原性, 不仅可诱导体内产生较强的细胞和体液免疫, 而且不同血清型分离菌株具有交叉保护作用, 因而成为疫苗研发的热点。随着基因组学、蛋白质组学、生物信息学等相关学科的发展, 越来越多病原菌的外膜蛋白的基因序列得以测出, 这为从病原菌中扩增目标基因, 进行重组表达, 快速筛选疫苗抗原提供了可能。目前, 已有学者对 OmpK、OmpU、OmpS2、OmpW、OmpC 等外膜蛋白的结构和免疫原性进行了研究<sup>[7-11]</sup>。但是 OmpLC (Outer membrane porin protein LC) 作为一种特定外膜蛋白, 在国内外鲜见报道, 其主要特性和功能尚不是很清楚。本文以鮠爱德华菌 *OmpLC* 基因序列为基础, 对该蛋白进行一系列生物信息学分析, 以期今后进一步研究该蛋白的功能及鮠爱德华菌相关发病机制以及疫苗的研制提供理论依据。

## 1 材料与方法

### 1.1 实验材料

#### 1.1.1 菌种

鮠爱德华菌, 由西南民族大学生命科学与技术学院水产养殖实验室从患病黄颡鱼肠道采样, 经回感实验确认为此次黄颡鱼患病的主要病原菌。鮠爱德华菌经分子鉴定后, 保存-20℃冰箱中备用。

#### 1.1.2 主要试剂和仪器

营养琼脂 (Nutrient Agar; 北京奥博星生物技术有限责任公司)。TaqMix、DNA Marker DL2000 购自

TIANGEN BIOTECH (BEIJING) CO., LTD。Yeast Extract (OXOID)、Tryptone (OXOID), PCR 仪 (Eppendorf Germany), 电泳仪 (BIORAD) 凝胶成像系统 (BioRAD Laboratories Segrate Italy), 超低温冰箱, 离心机 (Eppendorf), 恒温振荡培养箱 (HZQ-F100), 恒温培养箱 (上海齐欣科学仪器科技有限公司)。高压灭菌锅 (SANYO Electric CO., LTD Japan)。

### 1.2 *OmpLC* 基因序列扩增

参考 GENBANK 已有的 *OmpLC* 基因 CDS 序列, 利用软件 Primer Premier 5 设计上游引物 (*ompLCF*): 5'CCCCGATGGAATCAAAT3'、下游引物 (*ompLCR*): 5'CCTGCCTG ATCGG A TA 3', 委托上海生工生物工程有限公司合成。以反复低温和煮沸法粗提细菌总 DNA 模板。

以提取的鮠爱德华菌总 DNA 为模板, 用引物 *OmpLCR*、*OmpLCF*, 对 *OmpLC* 基因进行 PCR 扩增。PCR 反应体系: 上下游引物各 1 μL (引物浓度均为 10 μmol/L), DNA 模板 3 μL, ddH<sub>2</sub>O 20 μL, Mix 25 μL, 总体系为 50 μL。PCR 反应条件: 94℃ 预变性 4 min; 94℃、30 s, 51.5℃、30 s, 72℃、1 min, 30 个循环; 72℃ 延伸 4 min。PCR 扩增产物用 1% 琼脂糖凝胶电泳检测, 将阳性样品胶回收后送交上海生工生物工程有限公司进行测序。

### 1.3 *OmpLC* 基因生物信息学分析

#### 1.3.1 核酸序列分析

测序完成后获得上下序列, 利用 DNASTAR 软件包的 SeqMan 程序进行序列拼接, 并根据测序峰图进行调整。获得完整序列后利用 NCBI 网站提供的 Blastn 和 Blastx 程序进行序列比对, 查找相似序列, 初步确定其功能。用 BioEdit 软件对基因序列进行 ORF 查找, 确定 CDS 区, 并将 CDS 区翻译成氨基酸序列。用 CodonW1.4 对编码区进行密码子偏好性分析。最后将该序列提交至 GenBank。

#### 1.3.2 蛋白质序列分析

用 ProtParam 分析蛋白的一级结构和及其理化参数 (分子量、等电点、吸光系数等); 用 ProtScal 和 DNASTAR 软件包 Protean 程序预测其亲疏水性; 用 TMHMM Server 2.0 进行跨膜区分析; 用 SignalP 4.0 Server 软件进行信号肽预测; 用 PredictProtein 预测蛋白质二级结构; 用 InterProScan、MotifScan、CDD 预测其结构功能域。用 Phyre2 进行三级结构预测; 用 DNASTAR 软件包中的 Megalign 进行分子进化分析。所用网络服务资源见表 1。

表 1 生物信息学分析在线网络资源

Table 1 Online database and software of bioinformatic analysis

软件	网址
ProtParam	http://expasy.org/tools/protparam.html
ProtScal	http://web.expasy.org/protscal/
TMHMM Server 2.0	http://www.cbs.dtu.dk/services/TMHMM-2.0
SignalP 4.0 Server	http://www.cbs.dtu.dk/services/SignalP/
PredictProtein	https://www.predictprotein.org/
InterProScan	http://www.ebi.ac.uk/Tools/pfa/iprscan/
MotifScan	http://hits.isb-sib.ch/cgi-bin/PFSCAN
CDD	http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
Phyre2	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index

## 2 实验结果

### 2.1 *OmpLC* 基因序列分析

#### 2.1.1 *OmpLC* 基因 CDS 区分析

PCR 产物经琼脂糖凝胶电泳后, 结果见图 1。经测序分析 *OmpLC* 基因序列片段大小为 1 133 bp, 序列提交至 GenBank, 收录号为 JX064520。鮰爱德华菌 (*Edwardsiella ictaluri*) *OmpLC* 基因的 cDNA 从 32 到 1 111 全长 1 080 bp, 编码 360 个氨基酸。

#### 2.1.2 *OmpLC* 基因密码子偏好性分析

不同物种不同基因对于密码子的使用具有明显的偏好性, 密码子使用的偏性与基因的时空表达、表达产物的结构及功能有密切关系。对鮰爱德华菌 *OmpLC* 基因进行偏好性分析表明: 编码 Phe 的密码子

UUC、编码 Leu 的密码子 CUG、编码 Ile 的密码子 AUC 等 20 种密码子为该基因的偏好性密码子(见图 2)。

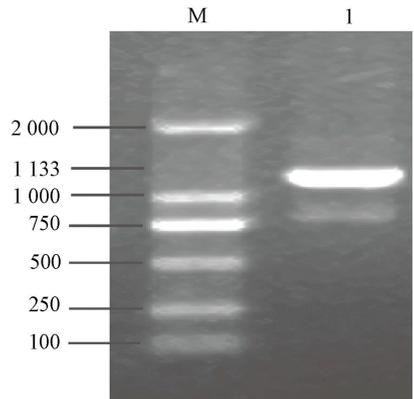


图 1 *OmpLC* PCR 扩增结果

Fig. 1 The PCR amplification results of *OmpLC*  
注: M; DNA Marker 2 000 bp; 1; *OmpLC* 基因 PCR 扩增产物。  
Notes: M; DNA Marker 2 000 bp; 1; The PCR amplification product.

Phe UUU	5 0.63	Ser UCU	5 1.76	Tyr UAU	6 0.43	Cys UGU	0 0.00
UUC	11 1.38	UCC	6 2.12	UAC	22 1.57	UGC	0 0.00
Leu UUA	0 0.00	UCA	0 0.00	TER UAA	0 0.00	TER UGA	0 0.00
UUG	1 0.25	UCG	0 0.00	UAG	0 0.00	Trp UGG	5 1.00
GUU	2 0.50	Pro CCU	0 0.00	His CAU	0 0.00	Arg CGU	8 4.00
CUC	2 0.50	CCG	1 0.67	CAC	1 2.00	CGC	4 2.00
CUA	0 0.00	CCA	1 0.67	Gln GAA	0 0.00	CGA	0 0.00
CUG	19 4.75	CCG	4 2.67	CAG	11 2.00	CGG	0 0.00
Ile AUU	1 0.30	Thr ACU	2 0.40	Asn AAU	3 0.19	Ser AGU	0 0.00
AUC	9 2.70	AGC	18 3.60	AAC	29 1.81	AGC	6 2.12
AUA	0 0.00	ACA	0 0.00	Lys AAA	17 1.55	Arg AGA	0 0.00
Met AUG	8 1.00	ACG	0 0.00	AAG	5 0.45	AGG	0 0.00
Val GUU	9 1.38	Ala GCU	7 0.72	Asp GAU	12 0.89	Gly GGU	19 1.81
GUC	6 0.92	GCC	21 2.15	GAC	15 1.11	GCC	23 2.19
GUA	5 0.77	GCA	4 0.41	Glu GAA	11 1.57	GGA	0 0.00
GUG	6 0.92	GCG	7 0.72	GAG	3 0.43	GGG	0 0.00

图 2 密码子偏好性分析

Fig. 2 Analysis of codon preferences

### 2.2 *OmpLC* 编码产物的基本性质分析

#### 2.2.1 *OmpLC* 编码产物的基本理化性质分析

蛋白质的基本理化性质包括其相对分子量、氨基酸组成、等电点等。用 ProtParam tool 软件预测鮰爱德华菌 *OmpLC* 蛋白的理化性质。*OmpLC* 蛋白分子量为 39.407 kDa, 理论等电点 PI 为 4.98, 带负电荷的氨基酸残基 (Asp+Glu) 总数为 41, 带正电荷的氨基酸残基 (Arg+Lys) 总数为 34, 不稳定系数为 18.26。根据不稳定参数的数值低于 40 才是稳定的蛋白<sup>[12]</sup>, 鮰爱德华菌 *OmpLC* 蛋白是一种稳定性蛋白, 软件预测该蛋白在体外环境下的半衰期为 30 h。

酸为 His(0.28%) (结果见表 2)。推测 *OmpLC* 蛋白分子量为 39.407 kDa, 理论等电点 PI 为 4.98, 带负电荷的氨基酸残基 (Asp+Glu) 总数为 41, 带正电荷的氨基酸残基 (Arg+Lys) 总数为 34, 不稳定系数为 18.26。根据不稳定参数的数值低于 40 才是稳定的蛋白<sup>[12]</sup>, 鮰爱德华菌 *OmpLC* 蛋白是一种稳定性蛋白, 软件预测该蛋白在体外环境下的半衰期为 30 h。

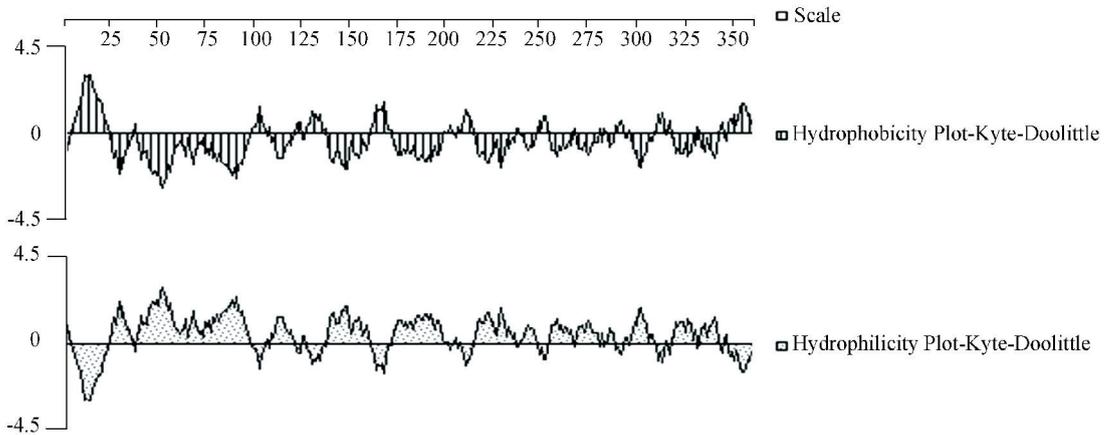
表2 氨基酸基本组成

Table 2 Basic composition of amino acids of *OmpLC*

Amino Acid	Number	Mol%	Amino Acid	Number	Mol%
Ala	39	10.83	Met	8	2.22
Cys	0	0.00	Asn	32	8.89
Asp	27	7.50	Pro	6	1.67
Glu	14	3.89	Gln	11	3.06
Phe	16	4.44	Arg	12	3.33
Gly	42	11.67	Ser	17	4.72
His	1	0.28	Thr	20	5.56
Ile	10	2.78	Val	26	7.22
Lys	22	6.11	Trp	5	1.39
Leu	24	6.67	Tyr	28	7.78

2.2.2 *OmpLC* 编码产物的疏水性分析

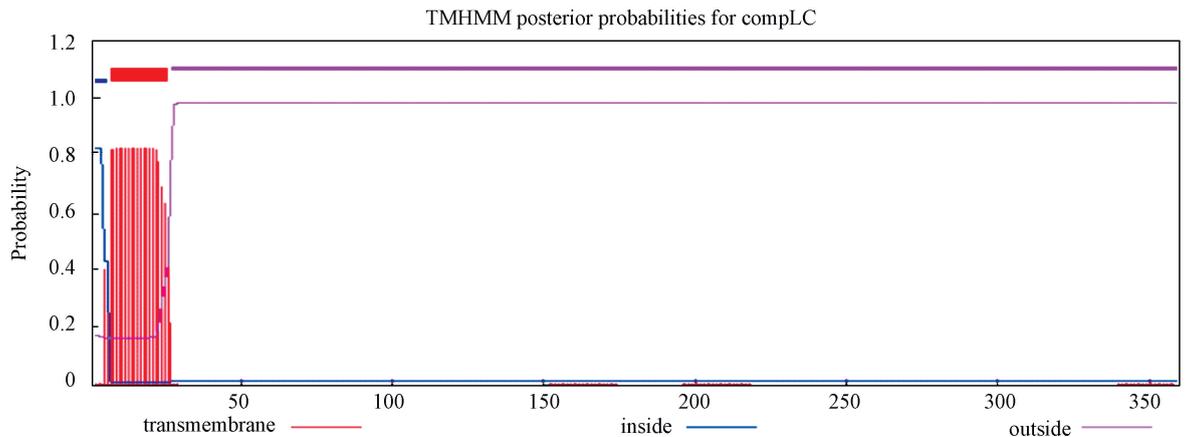
运用 Protean 程序 Kyte-Doolittle 算法对 *OmpLC* 编码产物进行亲水性/疏水性预测疏水性图谱:正值越大说明越疏水,负值越大说明越亲水;亲水性图谱反之。由亲疏水性图谱可知在前 3-22 氨基酸残基区域有较强的疏水性,参考 ProtScal 的分析结果,对于疏水性第 10 位的 Val 具有最大值为 3.000,第 51 位的 Asp 具有最小值为 -2.867,平均疏水性值为 -0.449,整个蛋白表现为亲水性,由此预测该蛋白是一种亲水性蛋白(见图 3)。

图3 *OmpLC* 编码产物亲疏水性图谱Fig. 3 Hydrophobicity analysis of *OmpLC* protein

## 2.2.3 编码产物的跨膜区分析

TMHMM,是一种基于隐马尔可夫模型的跨膜螺旋预测算法。利用 TMHMM server 2.0 在线软件对 *OmpLC* 编码产物的跨膜区域分析(见图 4),结果

显示,1-6 区域位于膜内,7-26 区域预存在一个跨膜螺旋,27-360 区域则位于膜外,N 端在细胞膜内预测的概率为 0.83。这与上述疏水性分析相吻合。

图4 *OmpLC* 编码产物跨膜区预测\*Fig. 4 Transmembrane region prediction of *OmpLC* protein

\* 彩图见电子版(<http://swxxx.alljournals.cn/ch/login.aspx>) (2016 年第 2 期 doi:10.3969/j.issn.1672-5565.2016.02.01)。

2.2.4 *OmpLC* 编码产物信号肽预测

信号肽是分泌蛋白质前体 N 端 15-30 个氨基酸组成的多肽,引导新生蛋白的定位。本文采用的 SignalP4.0 是个强大的信号肽及其剪切位点检测工

具,该算法基于神经网络算法,利用已知信号序列的革兰氏阴性原核生物、革兰氏阳性原核生物及真核生物的序列作为训练集,对分泌型蛋白的信号肽进行预测,并且给出 C、S、Y-score 计算结果<sup>[13]</sup>。其中

C 值是信号肽酶切位点分值, S 值是信号肽分值, Y 值是由 C 值和 S 值综合得出的剪切位点分值, 用于更精确的确定信号肽酶切位点<sup>[12]</sup>。一个典型的信号肽, 其 S 值在剪切位点之前高在剪切位点之后变低, Y 值和 C 值在剪切位点达到最大值并趋近于 1。

本实验预测结果(见图 5)表明: C 值和 Y 值在第 23 位分别获得最大值 0.884、0.898, S 值在第 3 位最大为 0.949, 1-22 位平均 S 值为 0.923, D 值是 S-mean 和 Y-max 的平均值, 为 0.910, 从而可以判断 *OmpLC* 编码产物存在信号肽, 剪切位点位于 22~23。

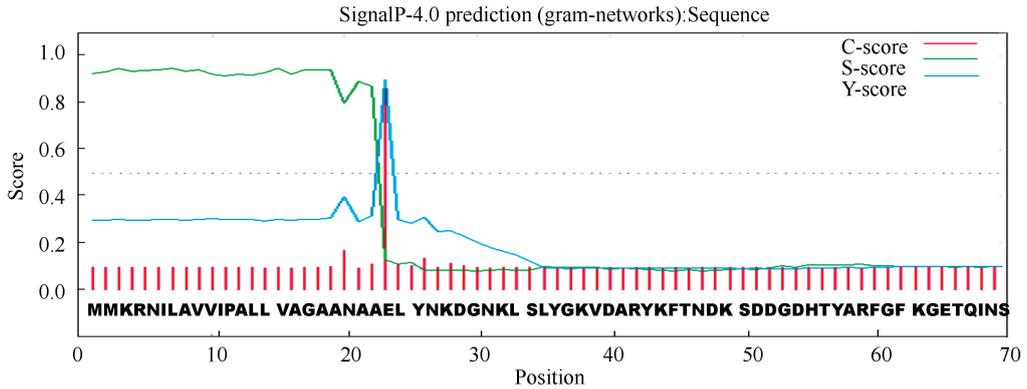


图 5 *OmpLC* 编码产物信号肽预测\*

Fig. 5 Signal peptide prediction of *OmpLC* protein

\* 彩图见电子版 (<http://swxx.alljournals.cn/ch/login.aspx>) (2016 年第 2 期 doi:10.3969/j.issn.1672-5565.2016.02.01)。

### 2.3 *OmpLC* 编码产物二级结构预测

PredictProtein 是欧洲分子生物学实验室提供的蛋白质序列和结构预测服务, 其分析方法包括: 用 PROFsec 分析序列的二级结构, 用 PROFacc 分析残基溶剂可及性, PHDhtm 分析潜在的跨膜拓扑结构,

通过 PROSITE 搜索模体, 借助 ProDom 预测功能结构域。发现该编码产物二级结构中螺旋占 6.67%, 折叠占 45.28%, 无规则卷曲占 48.06%, 其预测图如图 6 所示, 可以看出, 该产物二级结构主要由折叠和环构成。

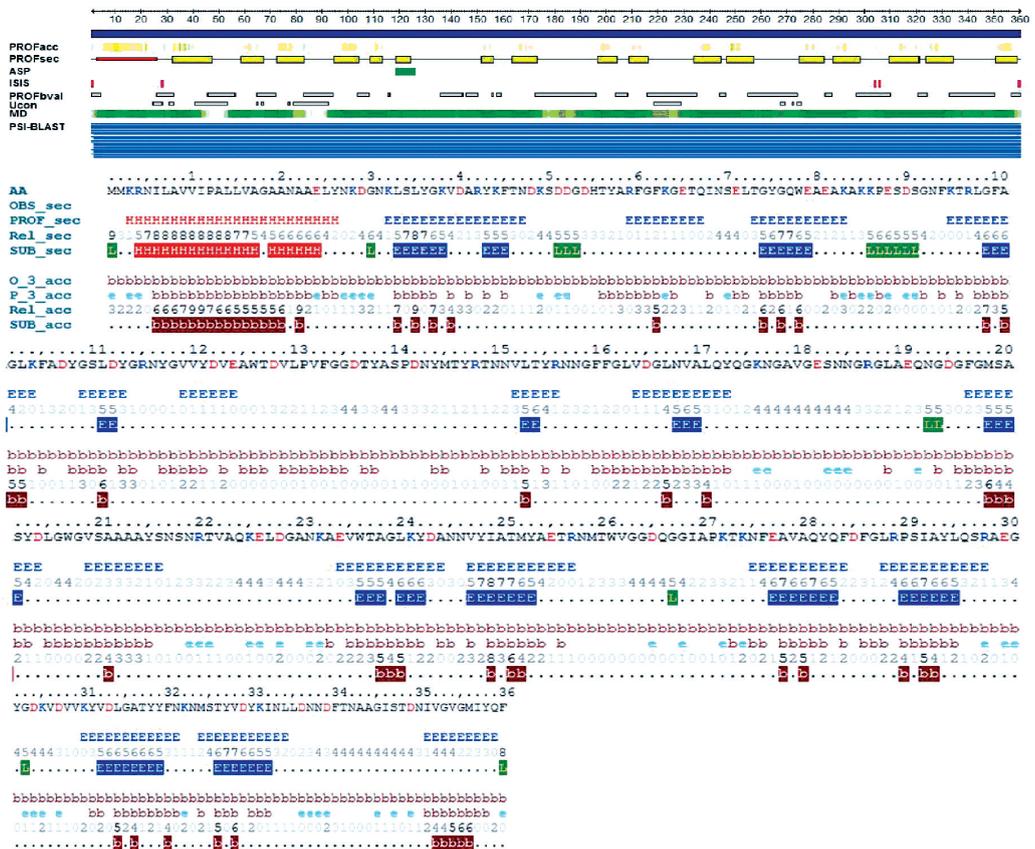


图 6 *OmpLC* 编码产物二级结构预测\*

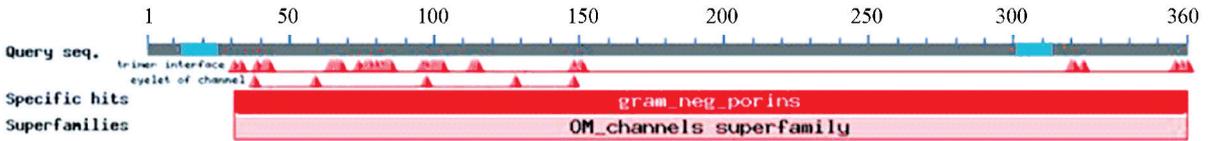
Fig. 6 Secondary structure prediction of the *OmpLC* protein

\* 彩图见电子版 (<http://swxx.alljournals.cn/ch/login.aspx>) (2016 年第 2 期 doi:10.3969/j.issn.1672-5565.2016.02.01)。

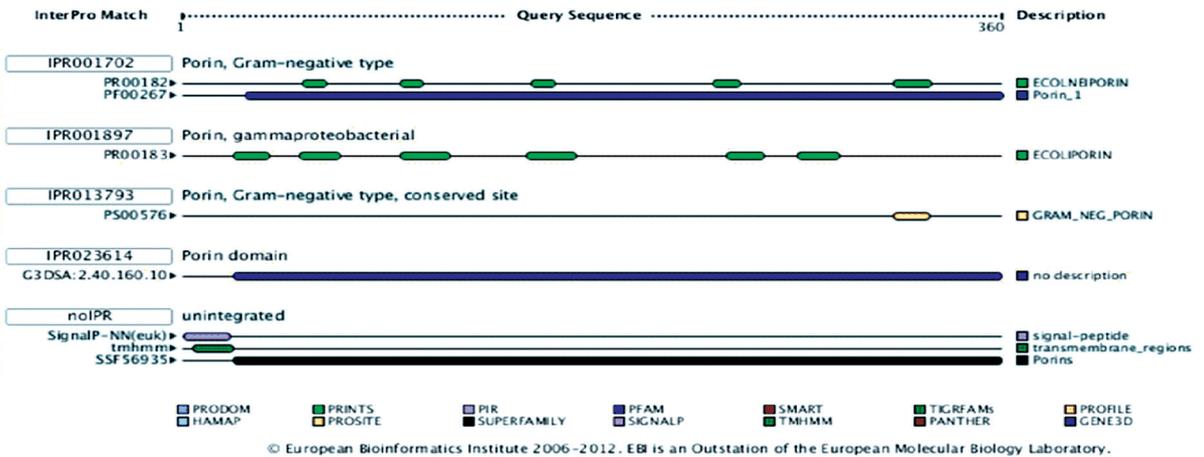
### 2.4 OmpLC 编码产物结构域和功能位点预测

InterPro 数据库由 EBI 开发,整合蛋白质家族、结构域和功能位点等资源。整合 UniProt、PROSITE、Pfam 等 12 个成员数据库,检索结果准确。用 CDD、InterProScan、MotifScan 对 *OmpLC* 的编码产物的基序及结构域进行分析,结果均指向 cl01155 (OM\_channels superfamily) 的成员 gram\_neg\_porins [cd00342] (见图 7a,7b)。*OmpLC* 编码蛋白由 360

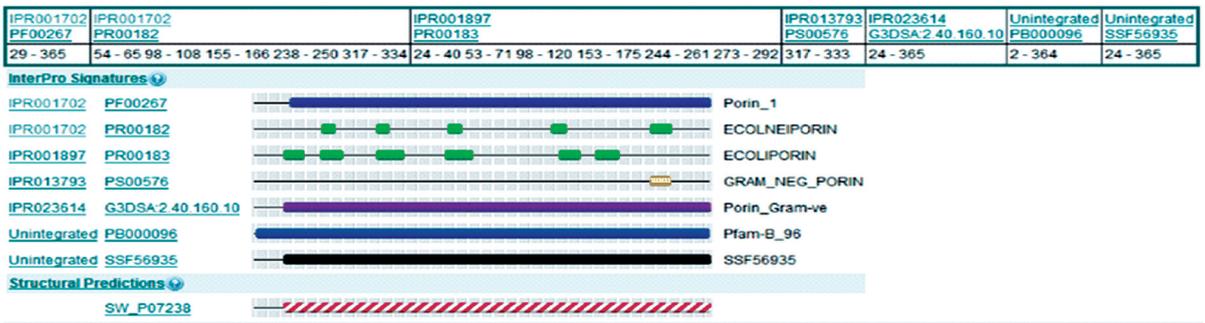
个氨基酸组成,其中 28-360 个氨基酸残基为保守结构域。该结构域由多个基序(53-64、96-106、153-164、233-245、312-348 等)组成(见图 7c),其中最典型的模体是第 312 - 348 位的表达式为 [LIVMFY].{2}G.{2}Y.F.K.{2}[SN][STAV][LIVMFYW]V,目标蛋白的模体为 VDLGATYYFNKNMSTYV,该家族蛋白在主要存在于细菌中,在病毒和和真核生物偶有出现(见图 7d)。



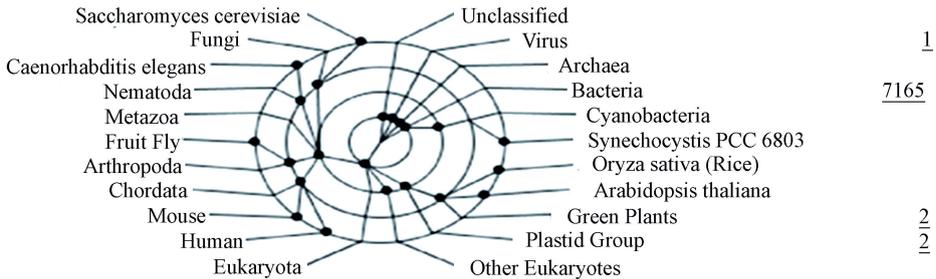
(a)通过 CDD 预测的保守结构域



(b)通过 InterProScan 预测的保守结构域



(c)通过 InterProScan 预测的保守基序特征



(d)革兰氏阴性型 Porin 家族在各种生物中的分布情况

图 7 *OmpLC* 编码产物结构域预测结果\*

Fig. 7 Functional domain of *OmpLC* protein

\* 彩图见电子版 (<http://swxx.alljournals.cn/ch/login.aspx>) (2016 年第 2 期 doi:10.3969/j.issn.1672-5565.2016.02.01)。

### 2.5 *OmpLC* 编码蛋白三级结构分析

CPHmodels 是采用同源建模来预测蛋白质三级结构的一个网络服务器,同时也采用了以预测距离为基础的穿线(threading)算法。同源建模是在蛋白质结构数据库中搜索目标蛋白的同源蛋白,然后将目标蛋白和同源蛋白进行对比,根据模板结构建模,并进一步优化模型作为预测的结果<sup>[14]</sup>。对于远源同源蛋白,结构虽然有一定的相似性,但其一致性小于 30%,此时很难确定序列和结构之间的关系,此时选择线串法进行蛋白结构预测效果较好。该算法是将待测序列“穿”入已知蛋白的基本骨架内,通过计算各种折叠的概率,来预测蛋白的核心结构。Phyre 是一种基于 profile-profile 比对和线串法,对预测蛋白和模板进行

1D-3D 的序列谱的比对预测。因此本文采用这两种方法进行预测,CPHmodels 结果显示与 1osm\_A 的相似性达 64.5%,其 3D 结构图如图 8a 所示。Phyre2 结果显示,同源蛋白为 d1osma,该蛋白属于 porins 超家族。目标蛋白中有 335 个氨基酸残基(93%)以 100% 置信率参与建模(见图 8b)。将二者所得建模参数文件导入 Swiss-Pdb Viewer 3.7,将其结构进行叠加比对(见图 8c),分析结构类似性,并利用 Ramachandran 参数进行检验。结果如图 7 所示。黄色区域是理想的  $\Phi$  角和  $\Psi$  角分布区域,而蓝色区域为不合理区域。从图 8d 中可以看出,模拟得到的 *OmpLC* 蛋白的残基的二面角绝大部分在合理区域内,表明该预测结构与自然结构趋势相同,具有稳定的空间结构。

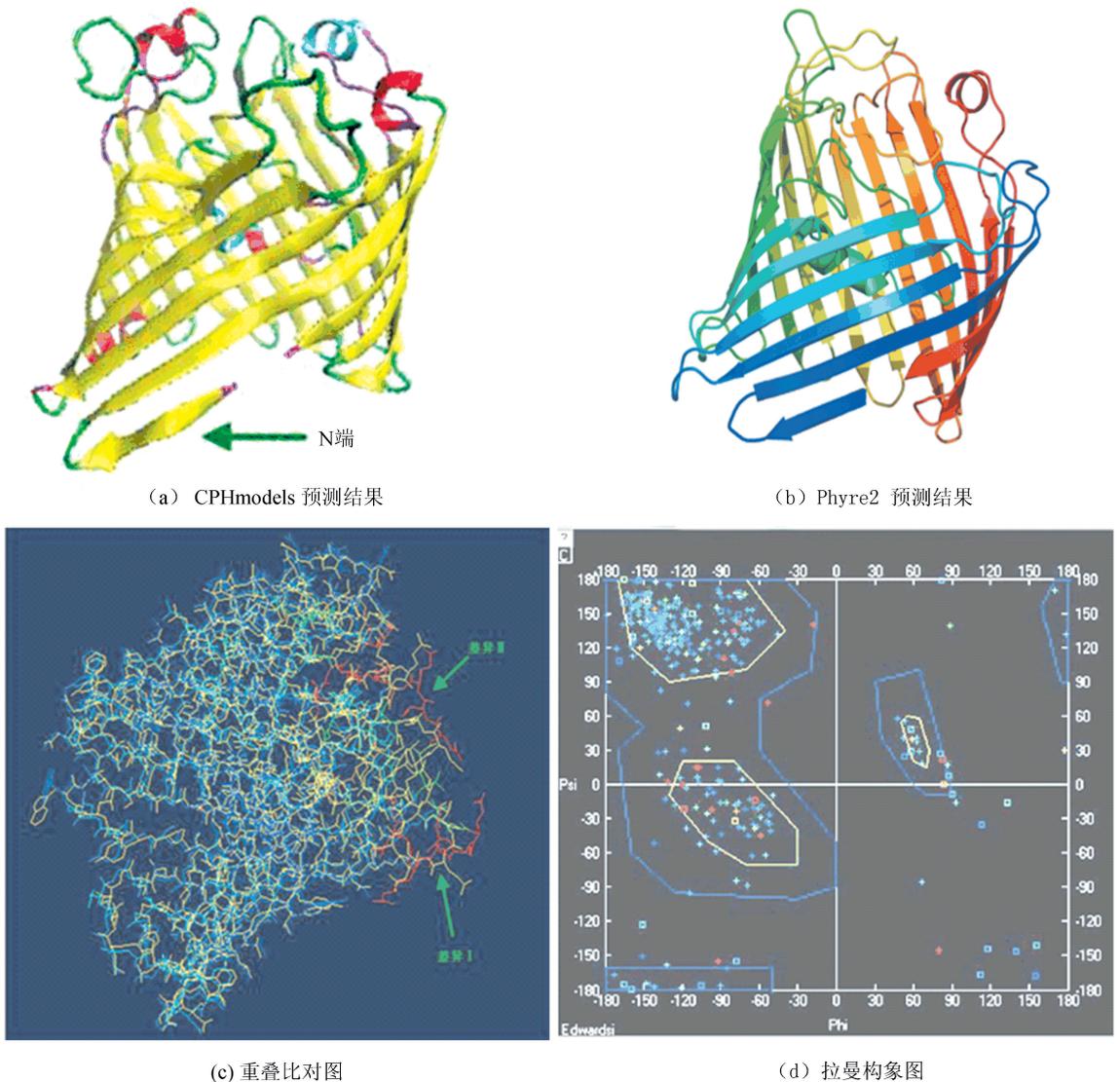


图 8 两种建模方法所得 3D 图及和重叠比较分析图\*

Fig. 8 The three-dimensional(3D) structure model and graphic superimposition analysis of the model predicted by two modeling methods.The predicted 3D structure of *OmpLC* protein

\* 彩图见电子版(<http://swxxx.alljournals.cn/ch/login.aspx>)(2016 年第 2 期 doi:10.3969/j.issn.1672-5565.2016.02.01)。

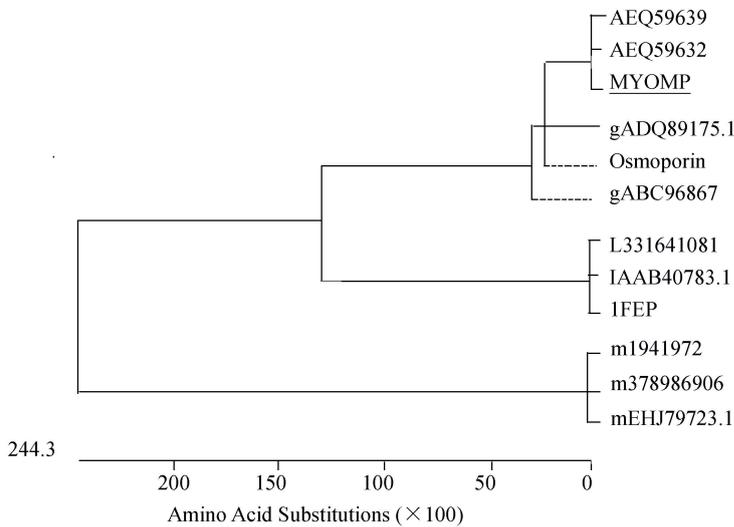
## 2.6 *OmpLC* 编码蛋白分子进化分析

本文选用 Blast p 程序,利用特异性位点迭代比对算法(Position-Specific Iterated BLAST, PSI-Blast),搜索非冗余数据库(GenBank CDS translations+PDB+SwissProt+PIR+PRF),进行同源蛋白搜索。遗传距离分析结果显示,目标蛋白序列 MYOMP(已用下划线标记)与 *OmpLC* 蛋白序列 AEQ59632、和 AEQ59639 序列相似性分别达到 100%和 99.7%(见图 9a),且从系统进化树上可以看出 MYOMP 和二者

聚为一类(见图 9b),因此可以判断目标蛋白为 *OmpLC*。12 个序列可以明显分为三类,首字母为 m 的序列 1941972、378986906、EHJ79723.1 属于类麦芽孔道蛋白类(Maltoporin-like),首字母为 l/l 的序列 331641081、AAB40783.1 和 1FEP 为配体门控通道蛋白类(ligand\_gated\_channel),其它的则属于革兰氏阴性孔道蛋白(gram\_neg\_porins)。该目标蛋白为革兰氏阴性孔道蛋白。

		Percent identity													
		1	2	3	4	5	6	7	8	9	10	11	12		
Divergence	1	■	99.7	17.8	66.3	67.3	66.5	17.8	17.8	9.4	9.6	9.6	99.7	1	AEQ59632
	2	0.3	■	17.8	66.3	67.3	66.5	17.8	17.8	9.7	9.9	9.9	100.0	2	AEQ59639
	3	277.0	277.0	■	21.3	21.2	20.5	98.3	98.5	10.8	10.7	10.7	17.8	3	1FEP
	4	44.6	44.6	230.0	■	95.8	99.1	21.3	21.3	11.0	11.3	11.3	66.3	4	gABC96867
	5	42.8	42.8	231.0	4.3	■	96.5	21.2	21.2	10.7	10.9	10.9	67.3	5	gADQ89175.1
	6	44.3	44.3	240.0	0.9	3.6	■	20.5	20.5	11.0	11.0	11.0	66.5	6	Osmoporin
	7	277.0	277.0	1.7	230.0	231.0	240.0	■	99.9	10.8	10.7	10.7	17.8	7	L331841081
	8	277.0	277.0	1.5	230.0	231.0	240.0	0.1	■	10.8	10.7	10.7	17.8	8	IAAB40783.1
	9	560.0	538.0	466.0	456.0	476.0	461.0	466.0	466.0	■	100.0	99.8	9.7	9	m1941972
	10	545.0	524.0	476.0	447.0	466.0	461.0	476.0	476.0	0.0	■	99.8	9.9	10	m378986906
	11	545.0	524.0	476.0	447.0	466.0	461.0	476.0	476.0	0.2	0.2	■	9.9	11	mEHJ79723.1
	12	0.3	0.0	277.0	44.6	42.8	44.3	277.0	277.0	538.0	524.0	524.0	■	12	<u>MYOMP</u>
		1	2	3	4	5	6	7	8	9	10	11	12		

(a) 遗传距离图



(b) 进化分析图

图 9 *OmpLC* 蛋白序列系统进化分析

Fig. 9 The phylogenetic analysis of *OmpLC* protein

## 3 讨论

生物信息学是一门建立在生物科学、数学、计算机科学基础之上的交叉科学。随着生物技术的快速发展和多个物种的基因组计划的相继完成,生物信息学在基因的功能发现、疾病基因诊断、蛋白质结构预

测、分子药物设计等方面发挥着重要作用。目前有很多生物信息学分析软件和数据库可以从网上免费下载,有些机构还有专门的在线服务系统。生物信息学借助这些广泛的信息资源及计算机辅助软件,能够对核酸、蛋白质等序列进行快速、高效、全面的分析,可以节约实验成本、缩短研究周期、避免试验的盲目性。*OmpLC* 基因能够编码一种特

定的外膜孔蛋白,有研究表明 *OmpLC* 能够作为噬菌体  $\Phi$ eiAU 和  $\Phi$ eiDWF 的吸附受体<sup>[15]</sup>,而关于其它相关功能尚还未见报道,因此选取鲷爱德华氏菌 *OmpLC* 作为目标蛋白,利用生物信息学方法分析其结构和功能。

鲷爱德华氏菌 *OmpLC* 蛋白富含 Gly 和 Ala,且成熟蛋白 N 端的第一个氨基酸是 Ala,它和 Gly 都属于稳定型残基,且不稳定系数为 18.26 (<40),为一种稳定蛋白。跨膜区分析结果说明在第 7-26 个氨基酸残基区域存在一个跨膜螺旋,这和二级结构预测的结果较一致,但因为其 1-22 区域为信号肽区域,因此成熟蛋白中应该无明显跨膜螺旋。此结果与朱珊珊等<sup>[16]</sup>研究的主要外膜蛋白(MOMP)结构有所不同。这可能和外膜蛋白种类及不同菌株的特异性有关。二级结构和三级结构表明该蛋白主要由 16 个  $\beta$ -折叠和一些无规则卷曲组成,三维结构呈桶状。通过同源建模和线串法两种建模方法,得到结果较一致且理论可靠的 *OmpLC* 蛋白模型。该模型同已报道 *ompC*、*ompF* 的结构相似<sup>[11,17-18]</sup>。有研究表明,*ompC* 和 *ompF* 基因的存在对于细菌的抗酸性和相关抗生素的扩散具有重要作用<sup>[17-18]</sup>,而鲷爱德华氏菌 *OmpLC* 和 *ompC* 及 *ompF* 具有相似的结构,但其是否也具有相应的功能尚不清楚,且 *OmpLC* 和 *ompC* 及 *ompF* 之间的进化关系如何亦有待于进一步研究。

革兰氏阴性孔蛋白在细菌生命过程中主要具有分子筛特性,对亲水性复合物进行过滤<sup>[19]</sup>。某些选择性孔蛋白(Porins)还具有溶质结合位点,允许特定的溶质跨膜。作为外膜蛋白的重要组成部分,孔蛋白(Porins)还可以作为噬菌体和细菌素的受体<sup>[20]</sup>。鲷爱德华氏菌 *OmpLC* 蛋白的结构域和同源性分析结果显示:鲷爱德华氏菌 *OmpLC* 蛋白属于外膜蛋白通道超家族(OM\_channels superfamily)的革兰氏阴性孔蛋白(gram\_neg\_porins)成员,因此,鲷爱德华氏菌 *OmpLC* 蛋白未来可作为噬菌体  $\Phi$ eiAU 和  $\Phi$ eiDWF 的吸附受体,而 *OmpLC* 基因的表达水平高低也可成为噬菌体感染敏感性重要分子因素<sup>[20]</sup>。

由于细菌外膜蛋白(OMP)的种类多样及其在生物体中位置的特异性,其纯化较为困难,因而限制了对其结构和功能的研究。本文对鲷爱德华氏菌的外膜蛋白——*OmpLC* 蛋白的序列进行一系列生物信息学分析,为进一步研究鲷爱德华氏菌致病的相关发病机制、鲷爱德华氏菌 *OmpLC* 蛋白的功能、表达载体的构建、分子进化、分子定向改造、免疫原性研究和疫苗研发奠定了基础。

## 参考文献

- [1] BRENNER D J, KRIEG N R, STALEY J T. Bergey's manual of systematic bacteriology[M]. London:Springer, 2005.
- [2] YE S G, LI H, QIAO G, et al. First case of *Edwardsiella ictaluri* infection in China farmed yellow catfish *Pelteobagrus fulvidraco*[J]. Aquaculture, 2009, 292 (1-2): 6-10.
- [3] 梁万文,陈明,余晓丽,等. 斑点叉尾鲷肠败血症病原菌的分离与鉴定[J]. 西南农业学报,2007, 20(5): 1124-1129.
- LIANG Wanfang, CHEN Ming, YU Xiaoli, et al. Isolation and identification of causative pathogen for enteric septicemia of catfish (ESC)[J]. South West China Journal of Agricultural Sciences, 2007, 20(5): 1124-1129.
- [4] KESKIN O, SECER S, IZGÜ R M, et al. *Edwardsiella ictaluri* infection in rainbow trout (*Oncorhynchus mykiss*) [J]. Turkish Journal of Veterinary and Animal Sciences, 2004, 28(4): 649-653.
- [5] 肖克宇,江为民,舒新华,等. 爱德华氏菌变异株 C9605 及对鳖的致病性研究[J]. 微生物学通报,1998,25(5): 262-265.
- XIAO Keyu, JIANG Weimin, SHU Xinhua, et al. Studies on *Edwardsiella ictaluri* variation strain (C9605) and its pathogenicity to soft-shelled turtles [J]. Microbiology China, 1998, 25(5): 262-265.
- [6] 熊静,关瑞章,郭松林,等. 鱼类病原菌外膜蛋白及其免疫原性研究进展[J]. 水生生物学报, 2011, 35(1): 163-169.
- XIONG Jing, GUAN Ruizhang, GUO Songlin, et al. A Review on the immunogenicity of fish pathogenic bacterial outer membrane proteins [J]. Acta Hydrobiologica Sinica, 2011, 35(1): 163-169.
- [7] LI Ningqiu, BAI Junjie, WU Shuqin, et al. An outer membrane protein, *OmpK*, is an effective vaccine candidate for *Vibrio harveyi* in Orange-spotted grouper (*Epinephelus coioides*) [J]. Fish & Shellfish Immunology, 2008, 25(6): 829-833.
- [8] 黄辉,毛芝娟,陈吉刚. 哈维氏弧菌外膜蛋白 *OmpU* 的克隆、表达与免疫原性研究[J]. 华中农业大学学报, 2010, 29(3): 346-350.
- HUANG Hui, MAO Zhijuan, CHEN Jigang. Cloning, expression and immunogenicity analysis of *OmpU* of *Vibrio harveyi* ZJ 2008 [J]. Journal of Huazhong Agricultural University, 2010, 29(3): 346-350.
- [9] 张亚宁,李晓,耿晓娜,等. 迟缓爱德华氏菌 HB01 外膜蛋白 *OmpS<sub>2</sub>* 基因的克隆表达及其免疫原性研究[J]. 细胞与分子免疫学杂志, 2011, 27(10): 1075-1082.
- ZHANG Yaning, LI Xiao, GENG Xiaona, et al. Cloning and expression of a outer membrane protein gene (*OmpS<sub>2</sub>*) of *Edwardsiella tarda* HB 01 and its immunogenicity [J]. Chi-

- na Journal of Cell Molecular Immunology, 2011, 27(10): 1075-1082.
- [10] 刘明智, 叶星, 田园园, 等. 嗜水气单胞菌外膜蛋白 W 基因的表达及其免疫原性分析[J]. 微生物学通报, 2011, 38(3): 437-445.
- LIU Mingzhi, YE Xing, TIAN Yuanyuan, et al. Expression and immunogenicity analysis of the outer membrane protein W gene of *Aeromonas hydrophila* [J]. Microbiology China, 2011, 38(3): 437-445.
- [11] LOU H B, CHEN M, BLACK S S, et al. Altered antibiotic transport in OmpC mutants isolated from a series of clinical strains of multi-drug resistant *E. coli* [J]. PLoS One, 2011, 6(10): 1-15.
- [12] 熊伟, 杨勇琴, 张海洋, 等. 人线粒体转录终止因子 1 (hMTERF 1) 蛋白的生物信息学分析[J]. 生物信息学, 2015, 13(1): 23-30.
- XIONG Wei, YANG Yongqin, ZHANG Haiyang, et al. Bioinformatic analysis of human mitochondrial transcription termination factor 1 (hMTERF 1) [J]. Chinese Journal of Bioinformatics, 2015, 13(1): 23-30.
- [13] PETERSEN T N, BRUNAK S, VON H G, et al. SignalP 4.0: discriminating signal peptides from transmembrane regions[J]. Nature Methods, 2011, 8(10): 785-786.
- [14] NIELSEN M, LUNDEGAARD C, LUND O, et al. CPH-models-3.0-Remote homology modeling using structure-guided sequence profiles [J]. Nucleic Acids Research, 2010, 38(20103811): 576-581.
- [15] HOSSAIN M J, RAHMAN KH S, TERHUNE J S, et al. An outer membrane porin protein modulates phage susceptibility in *Edwardsiella ictaluri* [J]. Microbiology, 2012, 158(2): 474-487.
- [16] 朱珊丽, 尤孙武, 娄崇洁, 等. 沙眼衣原体主要外膜蛋白生物信息学分析[J]. 温州医学院学报, 2009, 39(1): 5-7.
- ZHU Shanli, YOU Sunwu, LOU Chongjun, et al. Bioinformatics analysis of the major outer membrane protein of *Chlamydia trachomatis* serovar E [J]. Journal of Wenzhou Medical College, 2009, 39(1): 5-7.
- [17] KUMAR A, HAJJAR E, RUGGERONE P, et al. Structural and dynamical properties of the porins OmpF and OmpC: insights from molecular simulations [J]. Journal of Physics: Condensed Matter, 2010, 22(45): 454125.
- [18] BEKHIT A, FUKAMACHI T, SAITO H, et al. The role of OmpC and OmpF in acidic resistance in *Escherichia coli* [J]. Biological and Pharmaceutical Bulletin, 2011, 34(3): 330-334.
- [19] BENZ R, BAUER K. Permeation of hydrophilic molecules through the outer membrane of gram-negative bacteria. Review on bacterial porins [J]. European Journal of Biochemistry, 1988, 176(1): 1-19.
- [20] JAPAL B K, WALIANAL P J. Biophysics of the structure and function of porins [J]. Quarterly Reviews of Biophysics, 1990, 23(4): 367-403.