

# 魁蚶(*Scapharca broughtonii*)半乳糖凝集素(*SbGal*) 基因 cDNA 的克隆及表达分析\*

郑利兵<sup>1,2</sup> 吴彪<sup>1</sup> 刘志鸿<sup>1</sup> 周丽青<sup>1</sup> 孙秀俊<sup>1</sup>  
田吉腾<sup>1</sup> 杨爱国<sup>1</sup> 郑言鑫<sup>2</sup>

(1. 农业部海洋渔业可持续发展重点实验室 中国水产科学研究院黄海水产研究所 青岛 266071;  
2. 上海海洋大学水产与生命学院 上海 201306)

**摘要** 半乳糖凝集素(Galectin)作为动物凝集素家族的一员,是一种重要的免疫因子,在脊椎动物和无脊椎动物中都参与了重要的免疫防御反应。为明确魁蚶(*Scapharca broughtonii*)半乳糖凝集素基因的结构特征、组织分布及其对病原菌刺激的免疫反应规律,本研究从构建的魁蚶转录组文库中筛选得到半乳糖凝集素基因部分序列,通过 cDNA 末端快速扩增(rapid amplification of cDNA ends, RACE)技术获得该基因 cDNA 全长序列(命名为 *SbGal*)。序列和结构分析表明,该 cDNA 全长 3120bp,编码 554 个氨基酸,是一个含有 4 个糖识别结构域(CRD)的且在双壳贝类中较为独特的半乳糖凝集素;预测蛋白的分子量(Mw)为 63.3158 kDa,理论等电点(pI)为 4.99。同源性及系统分析表明,*SbGal* 基因与美洲牡蛎、合浦珠母贝、海湾扇贝 galectin 的同源性分别为 70%、65%、46%,*SbGal* 基因四个 CRD 彼此的相似性为 36%—41%。实时荧光定量 PCR(qRT-PCR)结果表明:*SbGal* mRNA 在魁蚶血细胞、斧足、闭壳肌、外套膜、鳃和肝胰腺中均有表达,血细胞中的表达量最高,在肝胰腺中的表达量最低;注射鳃弧菌后,相对于对照组,*SbGal* 基因在所检测的每个组织中 mRNA 水平上的表达量都显著上调( $P < 0.01$ ),而且表达量具有显著的时间依赖性和瞬时表达趋势,呈现先升高后降低的趋势。

**关键词** 魁蚶;半乳糖凝集素;基因表达;免疫反应

中图分类号 Q75 doi: 10.11693/hyhz20141100329

凝集素(lectin)是生物体内广泛存在的一类能使细胞发生凝集的糖类结合蛋白。作为软体动物体液免疫的一个重要免疫因子,凝集素在贝类免疫防御中起到异物成分的识别、调理、凝集等作用,是构成机体防御体系的重要组成部分(翟玉梅等, 1998; Weiss *et al*, 2000; 孙虎山等, 2001; 胥炜等, 2005)。

半乳糖凝集素(galectin)是动物凝集素家族的成员之一(Gerardo *et al*, 2009),可特异性识别糖链末端的  $\beta$ -半乳糖苷(贺雪明等, 2012),为非膜整合性的可溶性蛋白(Barondes, 1984; Drickamer *et al*, 1985)。在

贝类等无脊椎动物发生感染时,半乳糖凝集素通过其标志性的糖识别结构域(Carbohydrate Recognition Domain, CRD)特异性结合半乳糖苷来参与对病害的免疫反应(Dumic *et al*, 2006; 陈骋, 2013)。例如,菲律宾蛤仔(*Ruditapes philippinarum*)的半乳糖凝集素可结合派琴虫(*Perkinsus*)和弧菌(*Vibrio*)表面的半乳糖和 N-乙酰半乳糖胺而进行病原菌的识别(Kim *et al*, 2008);海鳗的原型 galectin 参与细胞对寄生虫的包裹作用(Nakamura *et al*, 2012);大西洋鳕鱼 galectin-1 能凝集革兰氏阳性和阴性菌,参与其免疫防御过程

\*山东省自然科学基金项目, ZR2013CQ047 号;科技基础条件平台项目“黄渤海区水生生物种质资源标准化整理、整合与共享”。  
郑利兵, 硕士研究生, E-mail: zhenglibing1988@hotmail.com

通讯作者: 刘志鸿, 研究员, E-mail: liuzh@ysfri.ac.cn

收稿日期: 2014-11-21, 收修改稿日期: 2015-02-19

(Rajan *et al.*, 2013)。因此, 半乳糖凝集素作为模式识别受体(Pattern Recognition Molecular, PRR)发挥作用的功能得到了广泛研究。

Galectin 家族成员众多, 在一种动物中可能含有不同的类型, 即使在同一种细胞中也可能表达出几种不同的 galectin (Colnot *et al.*, 1996)。近年来, 对软体动物半乳糖凝集素的研究中发现一种独特的 quadruple-galectin (Tasumi *et al.*, 2007; Song *et al.*, 2010; Zhang *et al.*, 2011), 其对半乳糖凝集素家族的结构是一个新的补充。但有关该类型 galectin 的报道还很少, 第一个被鉴定的 quadruple-galectin 是来自于美洲牡蛎(*Crassostrea virginica*), 它可以识别并吞噬原生动物鞭孢簇虫(*Perkinsus marinus*)(Tasumi *et al.*, 2007); 之后被鉴定的 quadruple-galectin 还有 PoGal (Zhang *et al.*, 2011)、AiGal1 (Song *et al.*, 2010)和 AiGal2 (Song *et al.*, 2011), 经过病原微生物的免疫刺激, 这些 galectin 的表达量都有明显的上调。然而, 目前尚未有魁蚶(*Scapharca broughtonii*)半乳糖凝集素的相关报道。

魁蚶属于双壳纲, 蚶目, 具有适应性强、生态修复效果显著的特点, 但近年来资源量锐减, 深入挖掘魁蚶免疫因子对其产业的发展具有重要的理论和现实意义。目前, 本课题组已在魁蚶血细胞分类及免疫功能研究(周丽青等, 2013, 2014)及种质资源和遗传多样性分析(吴彪等, 2010, 2012; 梁超等, 2011; 田吉腾, 2012; 周丽青等, 2012)等方面进行了研究, 发现鳃弧菌(*Vibrio anguillarum*)能够启动魁蚶的免疫防御功能。为进一步明确魁蚶半乳糖凝集素的特征和免疫作用, 本研究在魁蚶转录组文库中筛选到的半乳糖凝集素序列的基础上, 经过 RACE 扩增得到 cDNA 全长序列, 并开展基因表达组织分析和鳃弧菌刺激响应规律的研究, 以期为魁蚶分子育种及抗病选育提供依据, 为认识魁蚶半乳糖凝集素和对该分子的进化、结构和功能的研究提供资料。

## 1 材料与方法

### 1.1 实验材料与来源

实验用魁蚶购于青岛台东水产品市场, 壳长约 55mm 左右, 个体健康, 于 20°C 充气海水中暂养一周, 每天换水两次。期间投喂硅藻, 实验处理前两天停止投喂。鳃弧菌为本实验室保存的菌种。

### 1.2 魁蚶半乳糖凝集素基因的获得

**1.2.1 总 RNA 提取** 异硫氰酸胍法提取总 RNA: 采用 Solution D (4mol/L 异硫氰酸胍, 17mmol/L 月桂

酰肌苷酸钠, 25mmol/L 柠檬酸钠)与  $\beta$ -巯基乙醇裂解组织, 然后相同体积氯仿/异戊醇(24 : 1)和水饱和酚抽提蛋白质, 异丙醇与醋酸钠沉淀核酸, 75%乙醇洗涤核酸后干燥, 用 DNase(Progema)来去除基因组 DNA, 具体步骤参照董迎辉(2012)。1.0%琼脂糖检测 RNA 完整性, 核酸分析仪(Eppendorf)检测其纯度和浓度。

**1.2.2 全长 cDNA 克隆及测序** 根据获得的魁蚶半乳糖凝集素基因的部分序列, Primer Premier 5.0 设计 3'RACE 和 5'RACE 特异性引物, 利用 SMART<sup>®</sup> RACE Amplification Kit 和 Advantage 2 PCR Kit 用于 3'端和 5'端的扩增。3'RACE 扩增所用引物为 SbGal-3F 和通用引物 UPM, 5'RACE 扩增所用引物为 SbGal-5R 和通用引物 UPM(表 1)。具体步骤如下: 扩增体系为 50 $\mu$ L, 内含 3'(5')RACE cDNA 模板 2.5 $\mu$ L, PCR-Grade water 34.5 $\mu$ L, 10 $\times$  Advantage2 PCR buffer 5.0 $\mu$ L, dNTP Mix(10mmol/L)1.0 $\mu$ L, UPM 5.0 $\mu$ L, SbGal-3F (SbGal-5R) 1.0 $\mu$ L; 反应程序为: 94 °C 30 s, 68 °C 30 s, 72 °C 2 min, 25 个循环。PCR 产物经 1.2%琼脂糖凝胶电泳检测、切胶回收, 回收产物分别与 PMD18-T 连接, 转化大肠杆菌 Top10 感受态细胞, 并挑阳性克隆进行测序。

**1.2.3 序列分析** 将筛选到的序列及推测的氨基酸序列与 GenBank 数据库中的已知序列进行 blast 比对; DNASTar 软件对 cDNA 序列进行开放阅读框搜索及编码氨基酸预测; 利用 ProtParam tool (<http://web.expasy.org/protparam/>)对半乳糖凝集素分子量(Mw)及等电点 pI 进行预测; SignalIP 4.0 Server (<http://www.cbs.dtu.dk/services/SignalIP/>)查找半乳糖凝集素信号肽; TargetP 1.1 (<http://www.cbs.dtu.dk/services/TargetP/>)对半乳糖凝集素进行亚细胞定位; TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>)预测蛋白质跨膜区域; SMART (<http://www.expasy.ch/SMART>)进行蛋白质结构域分析。用 NCBI 的 Blastp (<http://blast.ncbi.nlm.nih.gov/>)对序列进行同源检索, 取不同分类物种以及不同种类半乳糖凝集素氨基酸序列, 用 DNA man8.0 软件与魁蚶铁蛋白氨基酸序列进行同源性比对分析, 在此基础上用 MEGA 5.0 构建系统发育树。

**1.3 半乳糖凝集素组织分布及对细菌刺激的响应模式分析**

**1.3.1 鳃弧菌刺激实验** 挑选健康的暂养魁蚶并随机分为对照组和处理组进行鳃弧菌刺激实验。鳃弧

菌经 2216E 培养基扩大培养后离心取沉淀, 用 PBS 调整浓度至  $OD_{600}=0.4$  ( $1OD=5 \times 10^8$  bacteria/mL)。微量注射器向处理组个体前闭壳肌处注射 50 $\mu$ L 菌悬液后放回原养殖槽内, 分别在 0h、4h、8h、16h、24h、32h、64h 随机选取 3 个个体活体取血(4 $^{\circ}$ C, 800r/min 15min 收集血细胞沉淀)、外套膜、鳃、闭壳肌、肝胰腺、斧足, 用于总 RNA 的提取。对照组注射同剂量的灭菌 PBS(0.1 mol/L, pH 7.4)代替鳃弧菌。

**1.3.2 实时荧光定量 PCR(qRT-PCR)** 运用 qRT-PCR 检测半乳糖凝集素转录本的组织分布情况及对鳃弧菌刺激的反应模式。根据获得的半乳糖凝集素 cDNA 序列, 用 Primer Premier 5.0 软件设计 qRT-PCR 引物 q*SbGal*-F、q*SbGal*-R(引物信息见表 1), 选取  $\beta$ -actin 基因作为内参基因(Li *et al.*, 2012), 反应在

ABI7500 荧光定量 PCR 仪上完成。利用 PrimeScript<sup>TM</sup> RT reagent Kit (TaKaRa)试剂盒进行反转录合成模板 cDNA, 反应体积及反应条件按照说明书进行。反转录产物置于 -20 $^{\circ}$ C 保存备用。qRT-PCR 扩增体系为 20 $\mu$ L, 包括: SYBR Premix Ex Taq<sup>TM</sup> (2 $\times$ ), 10 $\mu$ L; PCR Forward Primer(10 $\mu$ mol/L), 0.8 $\mu$ L; PCR Reverse Primer (10 $\mu$ mol/L), 0.8 $\mu$ L; ROX Reference Dye (50 $\times$ ), 0.4 $\mu$ L; cDNA 模板, 2.0 $\mu$ L; DEPC 水, 6.0 $\mu$ L。反应程序为 95 $^{\circ}$ C 30 s; 95 $^{\circ}$ C 5 s, 60 $^{\circ}$ C 34 s, 40 个循环, 95 $^{\circ}$ C 15 s, 60 $^{\circ}$ C 1 min, 95 $^{\circ}$ C 15s。实验设置 3 个平行组。

**1.3.3 数据分析** 软件获得每个反应的  $C_T$ , 采用  $2^{-\Delta\Delta CT}$  方法(Livak *et al.*, 2001)进行相对定量分析。根据每个样品的 3 个平行实验所得的数据计算相对表达量平均值和标准差。

表 1 实验所用引物序列  
Tab.1 The primers used in the experiment

引物	序列 5'—3'
10 $\times$ Universal Primer A Mix (UPM)	Long, 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3' Short, 5'-CTAATACGACTCACTATAGGGC-3'
SbGal-3F	5'-TTTGCCTGTGACTGTGAGAACGAGA-3'
SbGal-5R	5'-GCTGATCTCTCTCCTCATCTGCCA-3'
$\beta$ -actin-F	5'-GGTTACACTTTCACCACCACAG-3'
$\beta$ -actin-R	5'-ACCGGAAGTTTCCATACCTAAGA-3'
qSbGal-F	5'-GGATGGCAGGATGAG-3'
qSbGal-R	5'-AAGCACCACCAAAGC-3'

## 2 结果

### 2.1 *SbGal* 基因 cDNA 序列分析

3'RACE 和 5'RACE 扩增产物分别测序后与转录组文库中筛选出的序列进行拼接, 获得魁蚶半乳糖凝集素全长 cDNA 序列, 命名为 *SbGal*(GenBank 登录号为 KP123598)。*SbGal* 基因 cDNA 全长 3120bp, 包含一个 213 bp 的 5'-端非翻译区(5'-UTR), 1342 bp 的 3'-UTR 和 1565 bp 的开放阅读框(ORF); 其起始密码子为 ATG, 终止密码子为 TGA, Poly(A)序列上游有 4 个规则的加尾信号 AATAAA, 很明显, 该 cDNA 具有完整的蛋白质编码区(图 1)。

氨基酸序列分析可知, *SbGal* 基因编码一个含有 554 个氨基酸残基组成的蛋白质, 预测的分子量大小为 63.3158kDa, 理论等电点为 4.99。结构特征分析结果表明: 该氨基酸序列无信号肽, 也非线粒体靶向肽, 无明显的跨膜区域; 该蛋白具有多个功能化位点, 第 186 位 NETC 和 319 位 NGTD 存在潜在糖基化位点;

在 137、218、278、279、382 处有脾酪氨酸激酶磷酸化位点, 279、382、388、526bp 处有可激活胰岛素受体激酶磷酸化位点, 225 位有共济失调毛细血管扩张突变激酶磷酸化位点, 374 位有蛋白激酶 C 磷酸化位点, 218 位有表皮生长因子受体磷酸化位点; 结构域的预测分析结果表明 *SbGal* 有四个糖识别结构域(CRD), 且每个结构域通过一含有 6 个氨基酸的短肽链连接在一起, 每个 CRD 都含有保守的糖结合位点, 第一、第三和第四 CRD 均含有 132 个氨基酸残基(12—143, 286—417, 424—555), 第二 CRD 含有 130 个氨基酸残基(150—279)。综上分析结果表明, 本实验成功获得了魁蚶半乳糖凝集素基因全长 cDNA。

### 2.2 *SbGal* 基因同源性分析

同源性在线搜索结果显示: *SbGal* 编码的氨基酸序列与其它双壳类具有较高的相似性, 例如, 与美洲牡蛎 *Crassostrea virginica* 的相似性为 70%, 合浦珠母贝 *Pinctada fucata* 为 65%, 海湾扇贝 *Argopectens irradians* 为 46%; 与腹足类的相似性相对较低, 与光

1 AAATAACCGAAAATGAGATAACATTCCTTTCAAAGTAAAGATACGTGTGGATAAAAAGGTGTGTAATAACAATCACATGGTTGTCAGTATGGTTATTAGATATCTTAG  
111 TCGGCCATACATATCATATATAAGCTTTAATATCCGAGCACACACCCCTACCAGATAACAGCACAAGACTGGCCATACGACGTCGAGGAGGAGGAAACACTGCTCGTCCAC  
M S Y Q I S R V Y T P V S Y Y I P N G L H P  
221 AGTTCGGTACAGTTTAAATAGAAGGTTGCTAAAACAGGGATAAGATGTCCTACCAAAATTAGCAGAGTTTATACCCCTGCTCTTATTACATCCCAAATGGACTCCATCCA  
G K H I I L R G R V T W G T E A F A I N L Q E A A D A G D G C V A F H F N  
331 GGAAACACATCATCTCAGAGGGAGAGTCACATGGGTACAGAAGCATTGCCATAAAATTACAAGAAGCAGCAGATGCTGGAGATGGGTGTTGCTTCCATTTTAA  
P R P S E G V C V R N H C D G D W G G E E R D Q P H F P F D D G R S F L L  
441 CCCTCGCCCTAGTAGAGGAGTGTGTCCGTAATCACTGCGAGGACTGGGAGGTGAGGAACGTGATCAACCACATTCCCGTTTGTATGACGGAAGAAGTTCTCTGC  
R I E V T D S E F R T Y V H G K P Y I D F G H R V D L S R V H Y L H L T  
551 TCCGATTGAGGTACAGATAGTGAATTTAGGACATATGTCATGGAAAACCATACATAGATTTGGCCACAGAGTGGACCTTAGCCGTGTCATTATCTCCACTTGACC  
D G A E Y Y D I T F Q D R Y S L P Y R T E I P N G M Q V G K A V R I T G A  
661 GATGAGCAGAATACTATGACATCACATCCAGGACAGATATAGTCTACCATACAGAACAAGAAATCCCTAATGGTATGCAGGTTGGTAAAGCTGCTCCGATACATGGTGC  
A Q D N D G F S I N F A C D C E N E T C A F H F N P R P N E G V V I R N A  
771 TGCACAGGACATGATGGATTACGACATTAATTTGCTGTGACTGTGAGAAGCAGACCTGTGCTTCCATTTAAACCCCGCTCTAATGAGGAGTGTGTATCAGGAATG  
N L G G W G E E R D Y D A E F P F S P G N F F D A L F V C T D D Q Y C  
881 CTAACCTAGGAGGTTGGGGTGAAGAAGAACGGGATATGATGACAGAGTCCCAATTTCTCCAGGAAATTTCTTGTATGCCCTGTTGTATGACTGATGACCGATTGT  
V Y I N E K Y F A N F N H R C G V S D V N H F H V Q G N M D I R S V E Y Y  
991 GTTTACATTAATGAGAAATATTTGCCAACTCAACCATCGTTGGTGTATCAGATGCAACCATTCCATGTCAGGGCAACATGGACATAAGAAGTGGAGTATTA  
E P M D D D F V K P I P S G L E K G D V L L F K G F M R P G G D T F S I N  
1101 TGAACCAATGGATGATGATTTGTTAAACCTATCCATCAGGCTTAGAGAAAGGTGATGTTCTCTGTTTAAAGGATTCATGAGACTGGAGGAGATACATTTCCATCA  
L M N G T D I G S D I A F H I N P R V G E G Q V V M N C C M D G G W G E  
1211 ATCTGATGAATGGCACAGACATTGGATGACATTCGCTTCCACATCAACCTCGTGTGGGAGGGACAGGTCGTTATGAACCTGCTATGGACGGAGGATGGGGTGA  
E E R E D I P S P F T D R V P F E V K I V T K R N K F K V Y V N G K K Y M  
1321 GAGGAAAGAGAAGACATCCCATACCCCTTTACAGACAGAGTACCATTGAGGTCAAGATCGTAACCAAGAGAAACAAGTCAAGGCTATGCAATGGAATAAATACAT  
K F R A R G N V E D I K G V N V K G D A F I Y Q T K L L R K L D K P V W E  
1431 GAAGTTCGGTCCCGGGGAAACGTAGAAGACATCAAAGTGTATGTTAAGGAGACGCTTTATCTATCAGACAAAAGTCTCAGAAAAGTATGATAAAGCAGTCTGGG  
R I P G N F R E G G W V V V S G I P K K E S G G F A I N F R C A D D D D  
1541 AAAGAATCCAGAAATTTGCGAAGGTGGATGGGTTGTTGTTCAAGAAATCCAAAGAAAGAAATCTGGAGGTTTGGCATCAATTCGGTGTGCTGATGATGATGAC  
S D I A F H F N P R L N E G D T V R N A C V G G G W Q D E E R D Q P C F P  
1651 AGTGACATTGCCTCCATTTAAACCAAGTTTAAATGAGGGTGACACCGTACGTAATGCATGTGGGTGGTGGATGGCAGGATGAGGAGAGAGATCAGCCTGCTCCCC  
F E E K D T F E V A I N A W P D K F I T Y V N G K H Y I D Y N H R L P V D  
1761 ATTGGAAGAAAAGACACATTTGAAGTAGCCATCAATGCCTGGCCAGATAAAATTTATACATATGTAATGGTAAACATTACATAGATTACAACCACAGACTCTCTGTAG  
S V C H I Q L T G D A D F F E P E F F \*  
1871 ACAGTGTGTGCACATACAGTGCAGGGAGATGCTGATTCTTGAACAGAAATCTCTGAAAACCTTTGTTTAAATTAATGAACCTGCTAGTGGATAGTAAATCTTAGT  
1981 TGTGTAAGTGTATTATATAACTAGCAATATAAGAGATATTGCTTTGGTGGTCTGTATTAAATTAAGACAATCAGCATTCTGTGATTTTAAATTAATCAATGATC  
2091 AATGATTTTACGCAATATAATTTAATATGATGAAACGACTTAATGGTGACCAAAATGGCTGACAGATTGACAGATTTTCAGTTTTTATAAAATTTTAAACCGTC  
2201 GATTACAAGATAGCAACGAAGACGTTCAAGTTCATGTACACATATAGTTGATTTGTTTTTCAAAAACGGATTAATAAAGTTCATTTCAATGGGAACAGATTTT  
2311 GAGTTTGAAGAAAAATCTTTAGAAACATACAAGAGAAACAAGAAAGTTCCAAGATTTAAAATATTTATGAGAATCTAAAATGATTTGAAAATGGACTAAAATGGGA  
2421 TGTTTATATTGGAATGTACATAGTGATACTATAAAATCCATCATTGACATATGACTTTTAAAGTTGTGACTGTTAACTATTATTTAAATATAAGAATACTACA  
2531 TTTTGACACTAGATAGATTTTGGAAATACAGTTGCAATTCCTGATAACATATCAACATTAATTTGTTACATCTCGCTGTGTGATAATAACTAGTCAGACAAATTC  
2641 TAGCCAAACAGTTCATACAGGAGAATTTGAAATCAAGAAATCAGCACTACTATGAAACATATCATTCTATGACAATATTACAATGTCACTGTCACTTTAGTAAAG  
2751 ATGATCAACAGCAGTATTGTAATGATTCGGATTCTTTGAAAGAGTTATGGCAAACCATAGTATTCTGGTAAACAGCATTGAAATATAATGATTTGTCAGCAAA  
2861 AACACTGTACATATGGAGTTACTGTAGCATGTTATCATCATGATGCAAACTGACTGGAAAAACACATAGATGCGGCTGCCGACTGAAAAGTGAAGAACACATG  
2971 AATGCTGCTGGAAAAATGCTACACAACTGCTGATGACTAGTCTATAAAAAAATATATGTAACAGCTTCAAGTTTTATAAAATCTAATAATTGATGATTTATA  
3081 TTTTATAGTTATGATAAAACAACAATTTATGTGATATAATCTAGCAATTTTACATGACCCGTTACAATGTGATGATGATTTACGTAACATAAAATTTACT  
3191 GGGGGTTTTTAAAGCTGGTTTTACTTATTCAAAAAATAA

图 1 *SbGal* 核苷酸序列及其编码氨基酸序列

Fig.1 Nucleotide and deduced amino acid sequence of *SbGal*

阴影部分为 *SbGal* 的四个结构域, 加粗带下划线的为糖结合位点, 双下划线分别代表起始密码子和终止密码子, 带框的为加尾信号, 波浪线表示多聚尾

滑双脐螺 *Biompalaria glabrata* 两个半乳糖凝集素氨基酸的相似性分别是 30%、29%; 魁蚶半乳糖凝集素的四个 CRD 彼此的相似性在 36%—41%之间, 且 CDR1 与 CDR4 的相似性最大为 41%。

DNAMAN8.0 软件对 *SbGal* 基因编码的氨基酸序列与海湾扇贝、美洲牡蛎和合浦珠母贝的氨基酸序列进行比对(图 2), 在 *SbGal* 四个 CRD 上都有保守的糖结合位点, 不同物种的同一种 galectin 在糖结合位点上都很保守, 在 *SbGal* 的第一 CRD 中, 这些保守序列

分别是 His<sup>57</sup>, Asn<sup>59</sup>, Arg<sup>61</sup>, Asn<sup>70</sup>, Trp<sup>76</sup>, Glu<sup>79</sup> 和 Arg<sup>81</sup>, 在其它三个 CRD 中这些氨基酸残基也同样存在。

在 GenBank 上搜索选取了具有代表性物种的半乳糖凝集素序列, 利用 Mega5.0 构建 NJ 系统进化树如图 3, 整个系统发育树分成 4 大支, 原型 galectin 分成原型 1 和 2 两个分支, quadruple-galectin 居于这两原型分支间, 且与原型 galectin 1 首先聚在一起; 在 quadruple-galectin 分支中, *S. broughtonii* 首先与 *C. virginica* 聚在一起后再与 *P. fucata* 聚类。

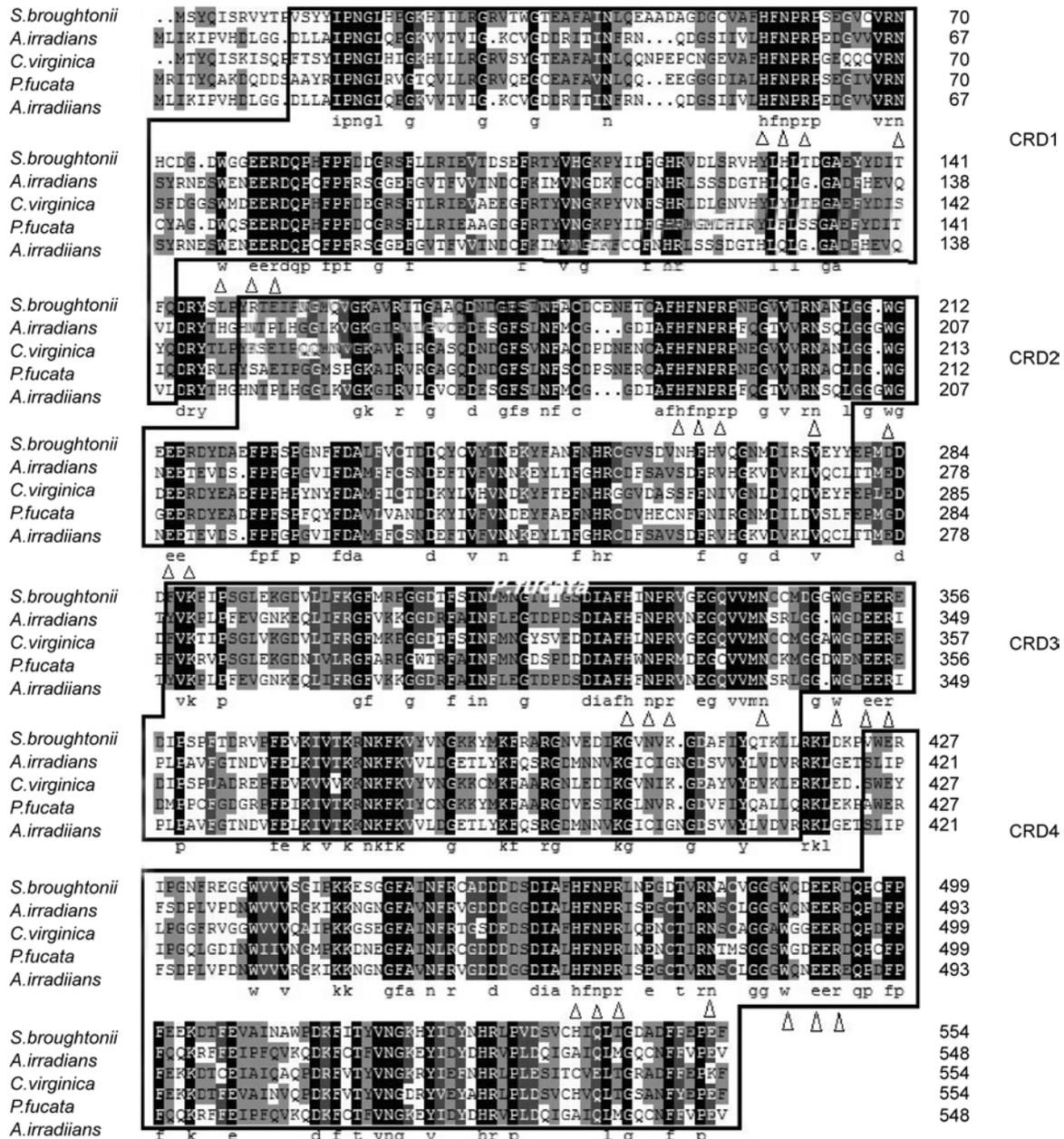


图 2 不同物种 Galectin 氨基酸序列比对

Fig.2 Multiple sequence alignment by DNAMAN 8.0

框线内为半乳糖凝集素的四个糖识别结构域, 用方框标出, 糖结合位点用三角标出, 各物种 GenBank 登录号为: 海湾扇贝 *A. irradians* (ACS72240, FJ469998); 美洲牡蛎 *C. virginica* (ABG75998); 合浦珠母贝 *P. fucata* (ACO36044)

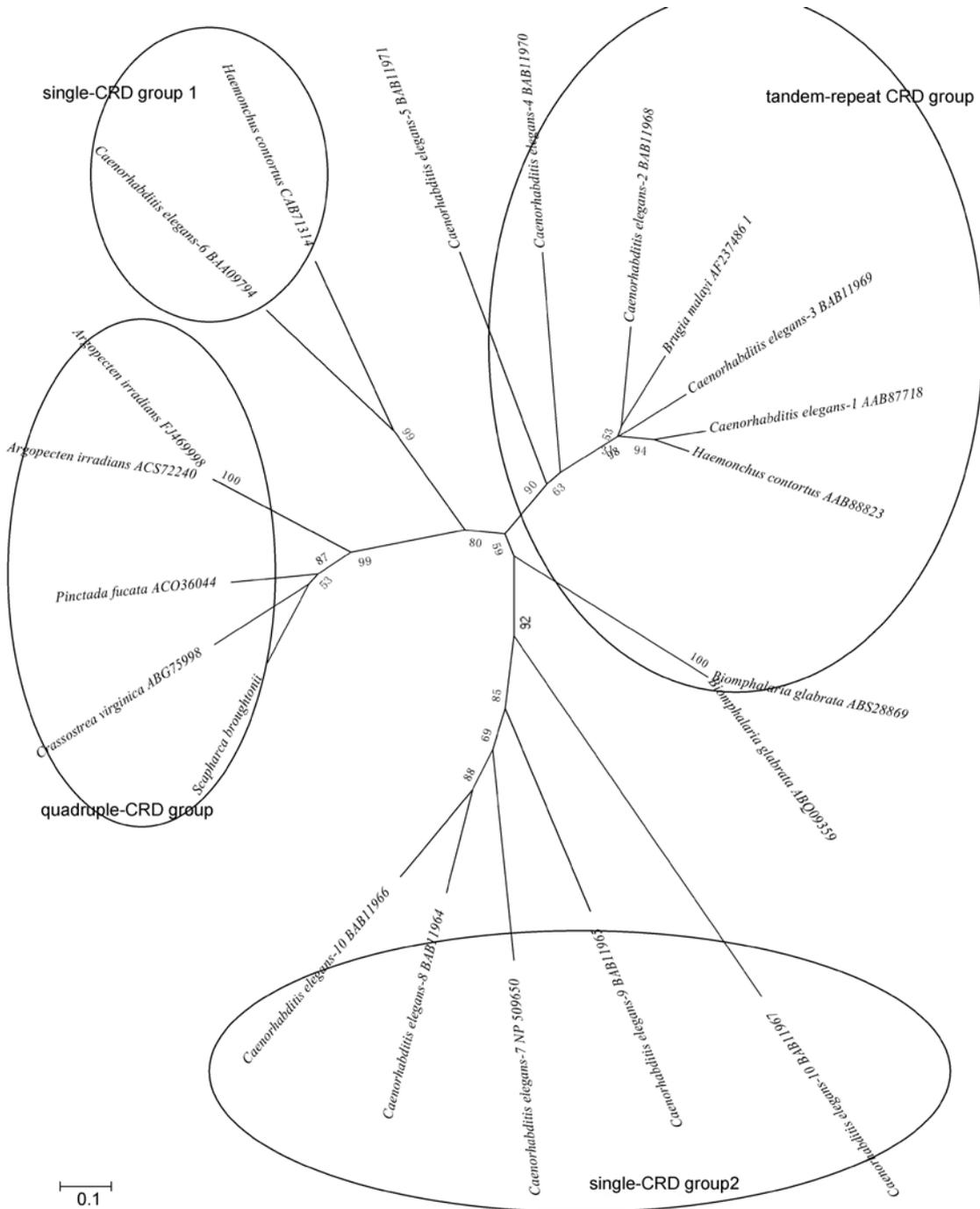


图 3 半乳糖凝集素的 NJ 系统进化树

Fig.3 Phylogenetic tree of CRDs from 21 galectins

各物种 galectin 同源氨基酸序列的 GenBank 登录号如下: 海湾扇贝 *A. irradians* (ACS72240, FJ469998); 美洲牡蛎 *C. virginica* (ABG75998); 合浦珠母贝 *P. fucata* (ACO36044); 秀丽隐杆线虫 1-9, *Caenorhabditis elegans*-1-9 (AAB87718, BAB11968, BAB11969, BAB11970, BAB11971, BAA09794, NP-509650, BAB11964, BAB11965); 秀丽隐杆线虫 *Caenorhabditis elegans*-10 (BAB11966, BAB11967); 马来布鲁丝虫, *Brugia malayi* (AF237486-1); 捻转血矛线虫 *Haemonchus contortus* (AA88823, CAB71314); 光滑双脐螺 *Biomphalaria glabrata* (ABS28869, ABQ09359)

### 2.3 *SbGal* 基因组织表达差异分析

以  $\beta$ -actin 为内参基因检测了 *SbGal* mRNA 在健康个体不同组织中的表达差异性(图 4), 结果表明:

*SbGal* mRNA 在斧足、鳃、外套膜、肝胰腺、闭壳肌、血细胞中都有表达, 但表达量存在明显的差异。以肝胰腺的表达量为参考, *SbGal* 在血细胞中的表达量最

高, 为肝胰腺的 87.05 倍( $P<0.01$ ), 其次是在斧足当中, 为 72.86 倍( $P<0.01$ ).

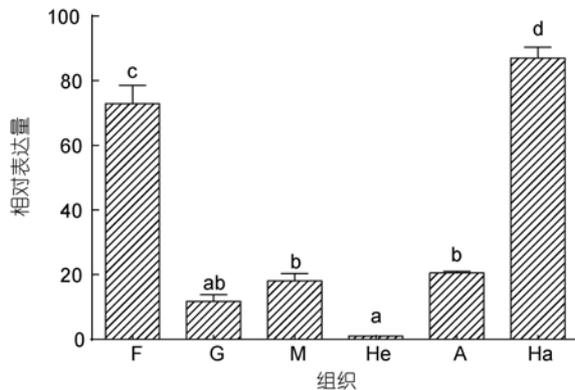


图 4 *SbGal* 在不同组织中的表达分布

Fig.4 Distribution of *SbGal* gene in different tissues of *S. broughtonii*

所有数据按照平均值 $\pm$  S.D. ( $n=3$ )表示。不同字母表示两组织之间 *SbGal* 基因表达量差异极显著( $P<0.01$ )。F. 斧足; G. 鳃; M. 外套膜; He. 肝胰腺; A. 闭壳肌; Ha. 血淋巴细胞

#### 2.4 *SbGal* 在鳃弧菌感染后的表达模式

鳃弧菌感染后, *SbGal* 在斧足、鳃、外套膜、肝胰腺、闭壳肌、血细胞 6 种组织中的表达模式如图 5, 结果表明: 鳃弧菌感染后, 魁蚶 6 种组织中的 *SbGal* 的表达量都明显上调。感染 4h 时, 在各组织中的表达量开始明显高于对照组, 并且在肝胰腺和闭壳肌中, 其表达量达到最大值, 分别为对照组的 1430.48 倍( $P<0.01$ )、1038.23 倍( $P<0.01$ ); 在斧足和鳃中的表达量于感染后 8h 达到最高值, 分别为对照组的 128.21 倍( $P<0.01$ )、1502.70 倍( $P<0.01$ ); 在外套膜和血细胞中的表达量于感染后 16h 达到最高值, 分别为对照组的 361.56 倍( $P<0.01$ )、354.87 倍( $P<0.01$ )。随后, 各组织中的表达量开始回落, 且在鳃和闭壳肌中其表达量的回落表现出明显的瞬时趋势; *SbGal* 在肝胰腺中的表达量于感染后 32h 出现二次表达高峰, 为对照组的 261.72 倍( $P<0.01$ )。总之, 经鳃弧菌感染后, *SbGal* 表现出显著的时间依赖性, 除血细胞外还表现出较显著地瞬时表达趋势, 总体表现为先升高后降低的趋势。

### 3 讨论

特异性结合  $\beta$ -半乳糖苷的半乳糖凝集素是动物凝集素的一种代表类型。半乳糖凝集素作为 PRR 参与免疫防御, 病原微生物的刺激会调节半乳糖凝集素在生物体内的表达, 进而加强免疫应答过程(Vasta,

2009)。近来, 有报道称半乳糖凝集素结合在细胞表面和体内潜在病原微生物细胞外基质的多糖上, 在先天免疫过程中以识别效应器的角色发挥作用(Vasta, 2009)。目前为止, 对无脊椎动物 galectin 的研究指出, quadruple-galectin 仅在软体动物中发现, 如 AiGal1、PfGal、AiGal2 已被成功鉴定并做结构功能分析。本研究通过构建魁蚶高通量测序转录组文库并利用 RACE 技术获得了魁蚶 galectin 基因全长 cDNA; 检测了其 mRNA 在魁蚶中各组织的分布情况和在外源微生物刺激下表达量的变化, 对探讨半乳糖凝集素在固有免疫防御中发挥的作用具有重要意义。

现在已鉴定分析的 galectin 无信号肽, 通过非经典的分泌方式到达细胞外(Almkvist *et al*, 2002)。*SbGal* 编码氨基酸序列的 N 端也不存在信号肽, 推测 *SbGal* 也是通过非经典分泌方式到达细胞外, 但还需要进一步的实验验证。*SbGal* cDNA 的 3'-UTR 存在多个加尾信号的现象可能与脊椎动物 galectin 一样存在多种剪切方式可产生多个转录变体, 皱纹盘鲍 galectin 基因 cDNA 的 3'-UTR 存在 3 处多聚腺苷酸加尾位点(郑明刚, 2007)。

同源性比对分析发现 *SbGal* 的四个结构域高度保守, 且每个结构域中与糖结合的氨基酸残基也都高度保守, 符合半乳糖凝集素家族成员的具有保守糖识别位点(Kasai *et al*, 1996; Leffler, 2001)的结构特征; NJ 系统进化树看到 *SbGal* 与同属于 quadruple-galectin 的其它物种的 galectin 聚在一起, 总之, 成功获得的 *SbGal* 也是一种 quadruple-galectin。

综上, *SbGal* 与已知的无脊椎动物半乳糖凝集素结构相似, 没有信号肽且只含有典型的半乳糖凝集素识别结构域, 这一点与脊椎动物的半乳糖凝集素结构特征相一致(宋小燕, 2009)。特征性序列分析、较高的同源性、保守氨基酸残基的存在都说明 *SbGal* 是半乳糖凝集素家族中的一个新成员, 其是在软体动物中发现的又一个具有四个明显结构域的半乳糖凝集素。因此, 本研究不仅填补了魁蚶半乳糖凝集素研究的空白, 而且还丰富了无脊椎动物半乳糖凝集素研究的相关数据。

qRT-PCR 检测发现 *SbGal* 与其它双壳贝类 galectin 一样在健康个体的不同组织都有表达(Tasumi *et al*, 2007; Song *et al*, 2010, 2011; Zhang *et al*, 2011), 但表达量存在差异, 说明 *SbGal* 基因的表达不具有组织特异性, 但表达量具有一定的组织特异性。有研究表明细菌、病毒、真菌或寄生虫感染后的牡蛎

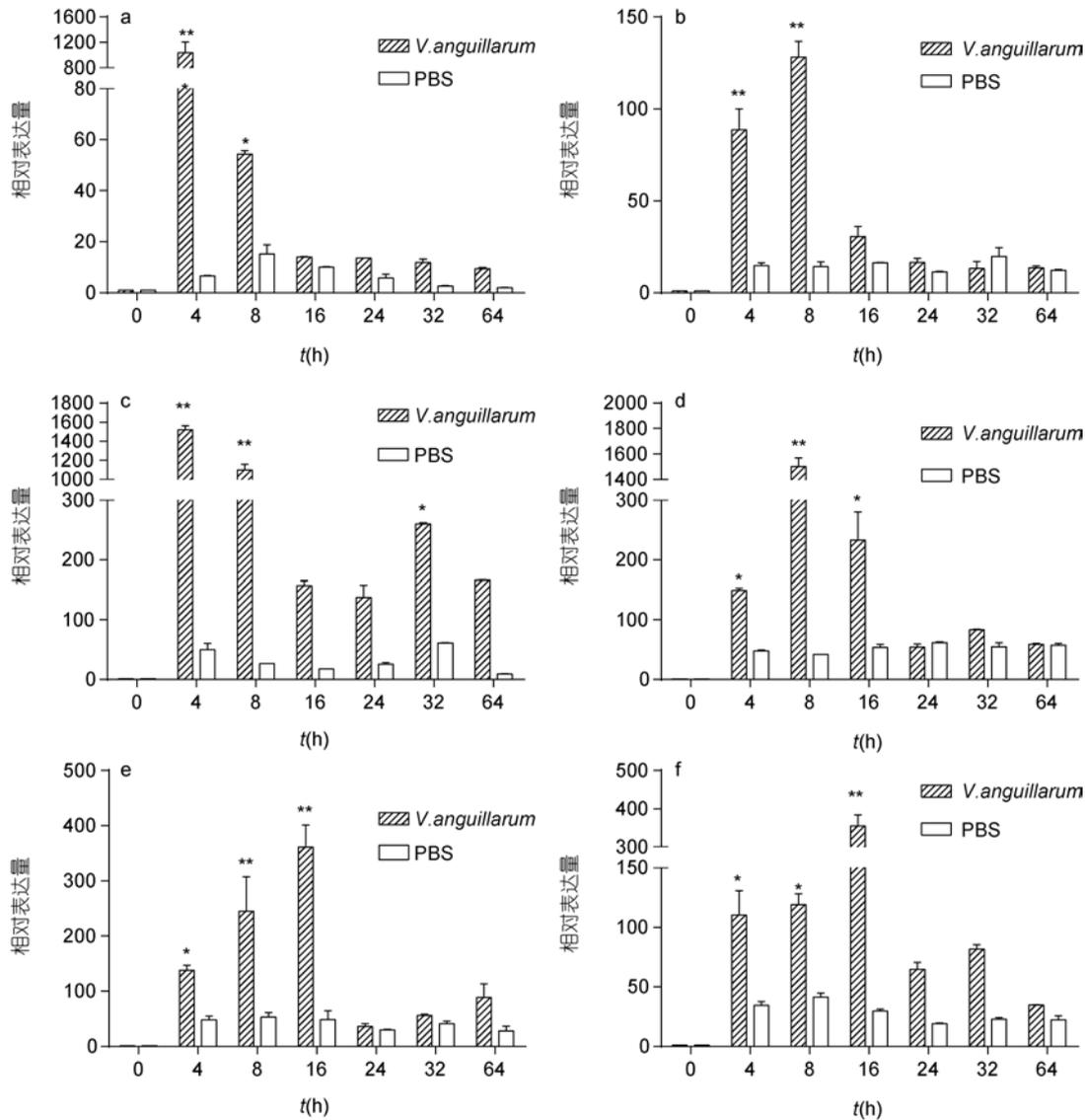


图5 鳃弧菌感染后魁蚶6种组织 *SbGal* 基因的表达情况

Fig.5 *SbGal* mRNA expression level after *V. anguillarum* challenge in five tissues

a. 闭壳肌; b. 斧足; c. 肝胰腺; d. 鳃; e. 外套膜; f. 血细胞。与0h相比,两个星号表示组间差异极显著( $P<0.01$ ),一个星号表示组间差异显著( $P<0.05$ )

oyster(Zhang *et al.*, 2011)、文昌鱼 *Branchiostoma lanceolatum*(禹艳红, 2007)和文蛤 *Meretrix meretrix*(Kim *et al.*, 2008)体内的半乳糖凝集素基因 mRNA 水平上调。经过鳃弧菌和藤黄微球菌(*Micrococcus luteus*)感染后的海湾扇贝, AiGal1 mRNA 的表达量显著上调,然而,经酵母感染后,其表达量基本保持不变(Song *et al.*, 2010);在正常情况下,合浦珠母贝半乳糖凝集素 PoGal2 持续表达,当被溶藻弧菌(*Vibrio alginolyticus*)感染后,其表达量也显著上调(Zhang *et al.*, 2011)。本研究中鳃弧菌感染魁蚶后, *SbGal* 在各组织的表达量都有升高,在肝胰腺中上升幅度最大,可

能 *SbGal* 通过肝胰腺发挥对病原菌入侵的防御作用较大(郑明刚, 2007),或者肝胰腺是软体动物先天免疫中合成蛋白的重要器官(Kong *et al.*, 2010; Prokhorova *et al.*, 2010; Zhao *et al.*, 2010)。鳃弧菌感染后, *SbGal* 前期的表达是逐步上调增强的,再经过一定的时间即恢复到感染前的水平,由此可见 *SbGal* 的表达在时间和空间上都有一定的规律,这为进一步研究 *SbGal* 在魁蚶免疫反应中的相关免疫功能研究提供了资料。

#### 参 考 文 献

田吉腾, 2012. 魁蚶微卫星富集文库的构建及四个地理群体的

- 遗传多样性分析. 上海: 上海海洋大学硕士学位论文, 1—59
- 孙虎山, 李光友, 2001. 双壳贝类参与免疫防御的体液因子. 海洋科学, 25(4): 34—36
- 吴 彪, 杨爱国, 刘志鸿等, 2010. 魁蚶两个不同群体形态性状对体质量的影响效果分析. 渔业科学进展, 31(6): 54—59
- 吴 彪, 梁 超, 杨爱国等, 2012. 基于 SSR 标记的魁蚶(*Scapharca broughtonii*)不同群体遗传结构的分析. 海洋与湖沼, 43(4): 863—869
- 宋小燕, 2009. 海湾扇贝凝集素基因和栉孔扇贝亲环素 A 基因的克隆与表达. 咸阳: 西北农林科技大学硕士学位论文, 1—84
- 陈 骋, 2013. 半滑舌鳎(*Cynoglossus semilaevis*)免疫相关基因 CsG3BP 的克隆、表达及功能研究. 北京: 中国科学院研究生院(海洋研究所)硕士学位论文, 1—72
- 周丽青, 杨爱国, 王清印等, 2012. 魁蚶四个地理群体 ITS 序列变异及系统发生分析. 渔业科学进展, 33(5): 79—84
- 周丽青, 杨爱国, 王清印等, 2013. 魁蚶血细胞分类及其免疫功能的初步分析. 水产学报, 37(4): 599—606
- 周丽青, 杨爱国, 王清印等, 2014. 鳃弧菌对魁蚶血细胞形态及免疫功能的影响. 海洋与湖沼, 45(3): 536—541
- 郑明刚, 2007. 皱纹盘鲍表达序列标签及免疫相关基因的克隆与表达研究. 青岛: 中国海洋大学博士学位论文, 1—124
- 禹艳红, 2007. 文昌鱼凝集素基因家族和其他补体相关的模式识别分子的克隆表达、进化分析稚功能研究. 广州: 中山大学博士学位论文, 1—192
- 胥 炜, 王 昊, 宋林生等, 2005. 栉孔扇贝 C 型凝集素基因的克隆与表达研究. 高技术通讯, 15(1): 83—88
- 贺雪明, 唐 丽, 贺福初, 2012. 半乳糖凝集素 1 的免疫功能. 生物技术通讯, 23(6): 896—899
- 梁 超, 杨爱国, 刘志鸿等, 2011. 4 个地理群体魁蚶(*Scapharca broughtonii*)的形态差异与判别分析. 海洋科学, 35(11): 108—113
- 董迎辉, 2012. 泥蚶高通量转录组分析及生长相关基因的克隆与表达研究. 青岛: 中国海洋大学博士学位论文, 1—144
- 翟玉梅, 丁秀云, 李光友, 1998. 软体动物血细胞及体液免疫研究进展. 海洋与湖沼, 29(5): 558—562
- Almkvist J, Karlsson A, 2002. Galectins as inflammatory mediators. Glycoconj J, 19(7—9): 575—581
- Barondes S H, 1984. Soluble lectins: a new class of extracellular proteins. Science, 223(4642): 1259—1264
- Colnot C, Ripoche M A, Scaerou F *et al*, 1996. Galectins in mouse embryogenesis. Biochem Soc Trans, 24(1): 141—146
- Drickamer K, 1985. Two distinct classes of carbohydrate-recognition domains in animal lectins. J Biol Chem, 263(20): 9557—9560
- Dumic J, Dabelic S, Flögel M, 2006. Galectin-3: an open-ended story. Biochem Biophys Acta, 1760(4): 616—635
- Kasai K-I, Hirabayashi J, 1996. Galectins: a family of animal lectins that decipher glyco-codes. J Biochem, 119(1): 1—8
- Kim J Y, Kim Y M, Cho S K *et al*, 2008. Noble tandem-repeat galectin of Manila clam *Ruditapes philippinarum* is induced upon infection with the protozoan parasite *Perkinsus olseni*. Dev Comp Immunol, 32(10): 1131—1141
- Kong P F, Zhang H, Wang L L *et al*, 2010. AiC1qDC-1, a novel gC1q-domain-containing protein from bay scallop *Argopecten irradians* with fungi agglutinating activity. Dev Comp Immunol, 34(8): 834—846
- Leffler H, 2001. Galectins structure and function—A synopsis. In: Crocker P R ed. Mammalian Carbohydrate Recognition System. Berlin: Springer, 57—83
- Li M, Zhu L, Zhou C Y *et al*, 2012. Molecular characterization and expression of a novel big defensin (*Sb-BDef1*) from ark shell, *Scapharca broughtonii*. Fish Shellfish Immunol, 33(5): 1167—1173
- Livak K J, Schmittgen T D, 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods, 25(4): 402—408
- Nakamura O, Watanabe M, Ogawa T *et al*, 2012. Galectins in the abdominal cavity of the conger eel *Conger myriaster* participate in the cellular encapsulation of parasitic nematodes by host cells. Fish Shellfish Immunol, 33(4): 780—787
- Prokhorova E E, Tsybalenko N V, Ataev G L, 2010. Expression of genes encoding defense factors in the snail *Planorbium corneum* (Gastropoda, Pulmonata) infested with trematodes. Parazitologiya, 44(4): 310—325
- Rajan B, Kiron V, Fernandes J M *et al*, 2013. Localization and functional properties of two galectin-I proteins in Atlantic cod (*Gadus morhua*) mucosal tissues. Dev Comp Immunol, 40(2): 83—93
- Song X Y, Zhang H, Wang L L *et al*, 2011. A galectin with quadruple-domain from bay scallop *Argopecten irradians* is involved in innate immune response. Dev Comp Immunol, 35(5): 592—602
- Song X Y, Zhang H, Zhao J M *et al*, 2010. An immune responsive multidomain galectin from bay scallop *Argopecten irradians*. Fish Shellfish Immunol, 28(2): 326—332
- Tasumi S, Vasta G R, 2007. A galectin of unique domain organization from hemocytes of the Eastern oyster (*Crassostrea virginica*) is a receptor for the protistan parasite *Perkinsus marinus*. J Immunol, 179(5): 3086—3098
- Vasta G R, 2009. Roles of galectins in infection. Nat Rev Microbiol, 7(6): 424—438
- Weiss I M, Kaufmann S, Mann K *et al*, 2000. Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusk *Haliotis laevigata*. Biochem Biophys Res Commun, 267(1): 17—21
- Zhang D C, Jiang S G, Hu Y T *et al*, 2011. A multidomain galectin involved in innate immune response of pearl oyster *Pinctada fucata*. Dev Comp Immunol, 35(1): 1—6
- Zhao J M, Qiu L H, Ning X X *et al*, 2010. Cloning and characterization of an invertebrate type lysozyme from *Venerupis philippinarum*. Comp Biochem Physiol B Biochem Mol Biol, 156(1): 56—60

## CLONING AND EXPRESSION ANALYSIS OF GALECTIN FROM *SCAPHARCA BROUGHTONII* (*SbGal*)

ZHENG Li-Bing<sup>1,2</sup>, WU Biao<sup>1</sup>, LIU Zhi-Hong<sup>1</sup>, ZHOU Li-Qing<sup>1</sup>, SUN Xiu-Jun<sup>1</sup>,  
TIAN Ji-Teng<sup>1</sup>, YANG Ai-Guo<sup>1</sup>, ZHENG Yan-Xin<sup>2</sup>

(1. Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China; 2. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China)

**Abstract** Galectins are a family of  $\beta$ -galactoside-binding lectins that play important roles in innate immunity of both vertebrates and invertebrates. In the present study, the cDNA of a galectin with multiple carbohydrate recognition domains (CRDs) was cloned from ark shell *Scapharca broughtonii* (designated *SbGal*) by partial cDNA sequence selected from the transcriptome library by rapid amplification of cDNA ends (RACE) technique. The full length cDNA of *SbGal*, another specific galectin in bivalve, was 3120 bp in length encoding a polypeptide of 554 amino acids, and contained four carbohydrate-recognition domains. The predicted molecular weight was 63.3158 kDa and the theoretical isoelectric point was 4.99. The results of multiple sequences alignment and phylogeny analysis show that the identity of *SbGal* shared with *Crassostrea virginica*, *Pinctada fucata*, and *Argopecten irradians* was 70%, 65%, and 46%, respectively. And the four domains of *SbGal* shared 36%—41% identities with each other. Quantitative Real-time PCR (qRT-PCR) was used to assess the mRNA distribution in all tested tissues, including hemocytes, foot, adductor muscle, mantle, gill, and hepatopancreas. The highest expression level was observed in hemocytes, while the lowest in hepatopancreas. After being challenged by *Vibrio anguillarum*, the relative mRNA expression level of *SbGal* was up-regulated significantly ( $P < 0.01$ ). The timing of maximum expression occurrence and returning to the initial level were different among the tissues. Therefore, *SbGal* expression is of obvious time dependence and instantaneous expression pattern, and it could function as a pattern recognition receptor in immune defense of ark shells.

**Key words** *Scapharca broughtonii*; galectin; gene expression; immune response