

基于马氏珠母贝(*Pinctada fucata*)血细胞全长转录组数据补体样组分的鉴定与分析*

王 菁 李桂英 林楷琪 欧 静 张淑瓶 王忠良

(广东海洋大学水产学院 广东湛江 524088)

摘要 补体系统作为先天免疫的重要组成部分,是一种复杂的限制性蛋白水解系统,其在免疫系统中发挥着重要的防御作用。为分析马氏珠母贝补体系统的组成及作用机制,使用血细胞样品进行了全长转录组测序建库、基因比对、功能注释,共挖掘到 212 个潜在补体样组分相关基因。补体样组分基因经同源性比对和结构域检测分析表明,检索到的基因分别编码 89 个含 C1q 结构域蛋白、57 个 C 型凝集素蛋白、33 个纤维胶凝蛋白、11 个纤维蛋白原相关蛋白、8 个甘露糖结合型凝集素关联丝氨酸蛋白酶、2 个含硫酸酯蛋白(1 个 C3 分子,1 个 TEP 分子)、1 个补体受体、2 个补体因子、9 个丝氨酸蛋白酶。随机选择 12 个补体相关基因,使用溶藻弧菌刺激前后的血细胞样品进行实时定量 PCR 检测其表达水平,结果显示 *C1q* (C1q domain containing protein)、*C-lectin*、*MBL* (mannose-binding lectin)、*ficolin*、*MASP* (mannan-binding lectin serine protease)等基因均呈现出显著差异表达,表明马氏珠母贝补体系统是一个复杂的多组分效应系统,且可能通过凝集素途径或类似于凝集素途径激活补体系统的免疫作用。研究结果为进一步验证马氏珠母贝中存在的原始补体系统提供了分子生物学证据,同时对深入了解马氏珠母贝免疫防御机制,丰富和发展海洋无脊椎动物免疫学内容也具有重理论意义。

关键词 马氏珠母贝; 全长转录组; 补体组分; 生物信息学分析; 实时定量 PCR

中图分类号 Q789 **doi:** 10.11693/hyhz20210700158

与其他无脊椎动物一样,马氏珠母贝(*Pinctada fucata*)缺乏适应性免疫系统,其机体的正常生命活动主要依靠先天免疫系统以抵御各种病原和外来物质的入侵(焦钰等, 2020)。先天免疫是机体抵抗和消灭外来病原入侵的第一道防线(吴勇等, 2019)。补体系统作为先天免疫的组成部分及免疫效应系统,是连接先天免疫和适应性免疫的枢纽,具有识别、清除病原体的作用(刘先军等, 2011; Li *et al.*, 2019)。

补体系统是一种复杂的生物效应系统,自 19 世纪末被发现以来,在哺乳动物中展开了广泛的研究

(Fujita *et al.*, 2004)。哺乳动物补体系统是由 30 多种蛋白组成的复杂蛋白水解系统,具有诱导炎症、调理并溶解病原体的作用(Suckale *et al.*, 2005)。补体系统的激活途径主要包括经典途径(classic pathway)、凝集素途径(lectin pathway)和替代途径(alternative pathway)(Fujita *et al.*, 2004)。经典途径和凝集素途径通过利用不同的蛋白质 C1q (C1q domain containing protein)、MBL (mannose-binding lectin)、Ficolins 识别靶标从而激活蛋白酶 C1r、C1s、MASP (mannan-binding lectin serine protease); 而替代途径中激活的 C3

* 广东省自然科学基金-面上项目, 2019A1515011875 号; 广东省省级科技计划项目(国际科技合作领域), 2019A050510044 号; 2021 年度广东省普通高校特色创新类项目, 2021KTSCX044 号; 广东海洋大学 2019 年“冲一流”省财政专项资金建设项目; 国家级大学生创新创业训练计划项目, CXXL2020002 号; 广东海洋大学“南海学者计划”(2017 年度); 广东海洋大学 2021 年度本科生创新团队项目, CXTD2021001 号。王 菁, 硕士研究生, E-mail: a13414942597@126.com

通信作者: 王忠良, 博士, 副教授, E-mail: zhongliangwang@vip.163.com

收稿日期: 2021-07-08, 收修改稿日期: 2021-08-20

(complement 3)可与 Bf (Factor B)结合, 在 Factor D 作用下形成 C3 转化酶(Nonaka *et al*, 2006)。三条补体途径激活后最终均进入终末途径(terminal pathway), 形成膜攻击复合物(membrane attack complex, MAC)裂解靶细胞(Suckale *et al*, 2005)。此外, 缺乏 C3 基因的小鼠可通过凝血酶替代 C3 依赖型 C5 转化酶介导补体途径的激活(Huber-Lang *et al*, 2006)。

事实上, 补体系统是一种古老的防御系统, 其起源早于脊椎动物的发育, 在许多无脊椎动物门中均发现了与哺乳动物相似的补体成分, 包括尾索动物(Al-Sharif *et al*, 1998; Suckale *et al*, 2005)、头索动物(Suzuki *et al*, 2002)、刺胞动物(Ji *et al*, 1997)、节肢动物(Dishaw *et al*, 2005; Liu *et al*, 2019)、软体动物(Gerdol *et al*, 2015; Wang *et al*, 2019)等。早期在海胆(sea urchins)和被囊类动物中鉴定到的与脊椎动物 C3 同源的分子具有调理活性以及潜在的 C3 转化酶作用。因此, 研究者普遍认为补体系统起源于后口动物(Nonaka *et al*, 2004a, 2004b)。然而, 在节肢动物蜚蠊(*Carcinoscorpius rotundicauda*)中发现了脊椎动物 C3 和 C2/Bf 的同源物, 表明在原口动物谱系中存在着以 C3 为中心分子的复杂的调理防御系统(Zhu *et al*, 2005)。从海鞘(*Halocynthia roretzi*)分离得到甘露糖凝集素, 可结合相关的甘露糖结合型凝集素关联丝氨酸蛋白酶, 进而激活 C3 启动补体系统(Sekine *et al*, 2001)。关于补体系统的研究在双壳贝类中同样展开了广泛的研究, 且相关补体分子在机体的免疫响应中发挥着重要作用, 包括栉孔扇贝(*Chlamys farreri*) (Yang *et al*, 2011)、菲律宾蛤仔(*Ruditapes philippinarum*) (Takahashi *et al*, 2008)、海湾扇贝(*Argopecten irradians*) (Yang *et al*, 2014)、光滑双脐螺(*Biomphalaria glabrata*) (Lu *et al*, 2020)、太平洋牡蛎(*Crassostrea gigas*) (Wang *et al*, 2017)、厚壳贻贝(*Mytilus coruscus*) (陈永霞, 2018)等。对太平洋牡蛎的研究发现, 其体内存在 C3、含 C1q 结构域蛋白(C1qDC)、凝集素样蛋白(C-lectin like protein)、纤维蛋白原相关蛋白(fibrinogen-related protein, FREP)、丝氨酸蛋白酶(Serine protease)和补体受体样蛋白, 表明双壳类中存在一个多组分的补体系统(Wang *et al*, 2017)。文蛤(*Ruditapes decussatus*)中 C3 样分子和 Bf 样分子的存在, 表明双壳类补体系统可能通过替代途径激活(Prado-Alvarez *et al*, 2009)。缢蛏(*Sinonovacula constricta*)中 C3、C1qDC、C-type lectin 基因在先天免疫中发挥重要的免疫作用(Peng *et al*, 2017; Li *et al*, 2019; Lan *et al*, 2020)。此外, 补体系统

中存在具有相似结构的同源分子, 如 C3/C4/C5、C2/Bf 和 C1s/C1r/MASP, 被认为是基因复制的产物(Nonaka *et al*, 2004a)。因此, Nonaka 等(2004a)认为哺乳动物的补体系统似乎是通过基因复制从一个更简单的补体系统逐步发展进化形成的。

全长转录组测序技术的兴起为绝大部分非模式生物的基因研究提供了更全面的基因序列信息(Chao *et al*, 2019), 进而可以全面、系统地开展补体系统进化、功能和作用机制研究。本文基于马氏珠母贝全长转录组测序数据, 系统地检索马氏珠母贝血细胞中存在的潜在补体样组分基因; 同时, 使用溶藻弧菌(*Vibrio alginolyticus*)刺激马氏珠母贝后的血细胞样品进行实时定量 PCR 检测相关补体基因的表达水平变化, 有助于加强对马氏珠母贝免疫相关基因的系统认知, 进一步理解马氏珠母贝补体系统的组成及作用机制, 对丰富和发展海洋无脊椎动物补体系统也具有重要理论意义。

1 材料与方法

1.1 马氏珠母贝的暂养、细菌刺激及总 RNA 提取

马氏珠母贝(平均壳长 70 mm)购自广东省湛江市雷州养殖场, 暂养于室内过滤海水桶中, 水温 25 °C, 饲养期间保持充气, 定时投喂螺旋藻粉; 每天 100%换水一次。室内暂养一周后将马氏珠母贝随机分为两组, 实验组采用闭壳肌注射的方法注射 100 μ L 浓度为 5×10^7 CFU/mL 的溶藻弧菌悬浮液; 对照组于闭壳肌处注射等体积的 PBS 溶液(pH 7.0) (Wang *et al*, 2016)。注射 4 h 后, 于闭壳肌处采集血淋巴, 每只约 0.5 mL, 采集的血淋巴立即分装至 1.5 mL 离心管中, 4 °C, 800 \times g 条件下离心 10 min 收集血细胞, 并立即提取总 RNA。

1.2 全长转录组文库构建及测序

使用马氏珠母贝健康血细胞样品, 利用 Agilent 2100 和 NanoDrop 2000 分光光度计检测总 RNA 的完整性和浓度。按照制造商说明书进行 SMRTbell 文库构建, 采用 MagBead Loading 上机测序获得原始数据, 使用 SMRT Link v6.0 对原始数据进行分析提取 CCS (Circular Consensus Sequence)序列, 对全长序列进行去冗余聚类、校正, 得到的全长转录组数据已提交至 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/sra/SRR9644402>)。

1.3 生物信息学分析

Unigene 在 Nt、Nr、Swiss-Prot、KEGG、COG 及 Pfam 数据库中比对并获得基因注释信息(E 值

$<1.0E^{-5}$)。使用 unigene 在 Nr 库中比对结果最好(E 值最低)的序列为对应同源序列, 确定同源序列所属物种并统计比对到各个物种的同源序列数量。

1.4 补体样组分筛选

基于功能基因注释信息检索含 C1q 结构域蛋白(C1q、C1qDC)、凝集素(凝集素、C 型凝集素、甘露糖结合凝集素)、纤维胶凝蛋白、C3、纤维蛋白原相关蛋白、丝氨酸蛋白酶、含硫酯蛋白(thioester-containing protein, TEP)、补体因子等补体样组分。采用 NCBI 保守结构域数据库(<https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>)及 SMART server(<http://smart.embl-heidelberg.de/>)进行补体样组分的结构域检测。

1.5 实时定量 PCR 检测补体相关基因表达水平

从全长转录组数据中筛选的补体相关基因中随机选择 12 个基因进行实时定量 PCR 检测其表达水平, 使用 One-Step gDNA Removal 试剂盒(北京全式金生物公司)合成 cDNA 第一链。按照 PerfectStart[®] Green qPCR SuperMix (北京全式金生物公司)说明书进行实时定量 PCR, 实验获得数据使用 $2^{-\Delta\Delta Ct}$ 方法进行计算, 并使用 SPSS 24 进行独立样本 T 检验, $P < 0.05$ 为显著差异, $P < 0.01$ 为极显著差异。

2 结果与分析

2.1 全长转录组测序数据

利用 SMRT 测序技术, 共获得 14 048 434 585 总碱基数(bp), 原始平均序列长度为 1 838 bp, N50 为 2 563 bp。使用 Permissive CCS2 软件进行序列提取和质量过滤后, 共得到 915 060 427 个循环一致序列(CCS), 其中全长非嵌合序列(full-length non-chimeric, FLNC)具有 636 357 960 个 CCS, 去掉 5'primer、3'primer 和 polyA 后获得原始的全长转录本序列。序列去冗余后获得 82 381 unigenes, N50 为 2861 bp (表 1)。

表 1 马氏珠母贝全长转录组测序结果
Tab.1 The full-length transcriptome sequence result of *P. fucata*

转录本	SMRT 结果
数据总长/bp	14 048 434 585
平均原始数据长度/bp	1 838
原始数据 N50/bp	2 563
CCS	915 060 427
FLNC	636 357 960
Unigene	82 381
N50/bp	2 861

2.2 马氏珠母贝基因注释及功能分类

使用 BLAST 工具将所有的 Unigene 与 NR、NT、Pfam、Swiss-Prot 等蛋白序列数据库进行比对, 结果共得到 59 621 个注释基因和 22 760 个未注释基因(图 1), 表明全长转录组基因注释的水平较好。Unigene 注释到同源基因的物种分布如图 2 所示, 其中匹配到太平洋牡蛎的基因数量最多, 有 42 014 条基因。其他同源序列匹配较高的物种有马氏珠母贝、鸭嘴舌形贝(*Lingula anatina*)、美国加州海兔(*Aplysia californica*)等物种。

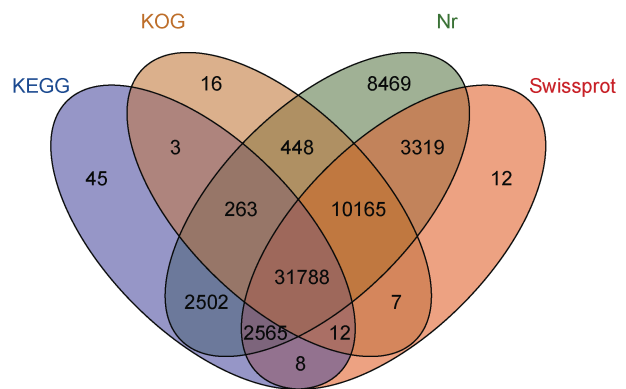


图 1 马氏珠母贝基因功能注释 venn 图

Fig.1 The Venn diagram of gene functional annotation in *P. fucata*

2.3 马氏珠母贝的补体样组分

通过检索马氏珠母贝全长转录组测序数据, 共获得 212 条补体样成分 unigene, 分别编码 89 个含 C1q 结构域蛋白、57 个 C 型凝集素蛋白、33 个纤维胶凝蛋白、11 个纤维蛋白原相关蛋白、8 个甘露糖结合型凝集素关联丝氨酸蛋白酶(MASP)、2 个含硫酯蛋白(1 个 C3 分子, 1 个 TEP 分子)、1 个补体受体、2 个补体因子、9 个丝氨酸蛋白酶。

2.4 含 C1q 结构域蛋白

在检索到的 89 个含 C1q 结构域蛋白(C1qDC)中, 有 81 个 C1qDC 含有 C 末端球状结构域, 10 个 C1qDC 蛋白还具有 smc 结构域; 6 个 C1q 亚成分结合蛋白含有 MAM33 结构域, 该结构域可与 C1q 的球状结构域结合, 参与线粒体的氧化磷酸化和细胞核-线粒体的相互作用, 但均未能检测到 N 末端的胶原结构域(表 2)。

2.5 纤维胶凝蛋白、纤维蛋白原相关蛋白及 C 型凝集素

甘露糖结合凝集素(MBL)和纤维胶凝蛋白(ficolin)分子的结构类似, 仅在 C 端的放射状球形结

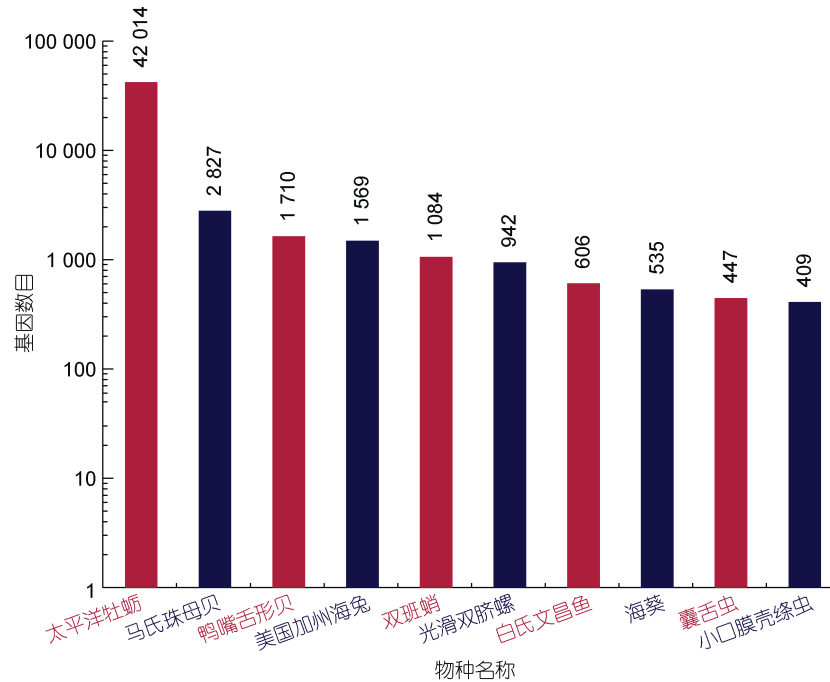


图 2 马氏珠母贝全长转录组 unigene 与 NR 数据库匹配物种分布

Fig.2 NR annotated species distribution of transcriptome unigenes of *P. fucata*

表 2 马氏珠母贝 C1q/C1qDC 保守结构域分析

Tab.2 Analysis of the conserved domains in C1q/C1qDC of *P. fucata*

基因 ID	注释信息	保守结构域
Isoform0004975	Complement C1q-like protein 4	—
Isoform0072432	Complement C1q-like protein 4	—
Isoform0002888	Complement component 1 Q subcomponent-binding protein	MAM33
Isoform0008172	Complement component 1 Q subcomponent-binding protein	MAM33
Isoform0008699	Complement component 1 Q subcomponent-binding protein	MAM33
Isoform0011565	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0015471	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0016222	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0019455	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0023169	Complement component 1 Q subcomponent-binding protein	MAM33
Isoform0025205	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0025376	Complement component 1 Q subcomponent-binding protein	MAM33
Isoform0026227	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0031059	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0047238	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0047673	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0049451	Complement component 1 Q subcomponent-binding protein	MAM33
Isoform0075709	Complement C1q-like protein 4	C-球状结构域、smc
Isoform0055352	complement 1 q protein	C-球状结构域

构域上各自有差别。在本文中, 共检索到 33 个 ficolin, 其中 31 个分子具有 C 末端 FBG 结构域。同时, 检测到 11 个具有 FBG 结构域的纤维蛋白原相关蛋白, 57

个 C 型凝集素基因(C-type lectin, CTL)中有 26 个基因具有 CTLD 结构域(表 3)。且在全长转录组数据中未检索到甘露糖结合凝集素。

表 3 马氏珠母贝 C 型凝集素保守结构域分析
Tab.3 Analysis of the conserved domains in C-type lectin of *P. fucata*

基因 ID	注释信息	保守结构域
Isoform0000057	C-type lectin-1	CTLTD 结构域
Isoform0000588	C-type lectin 8	CTLTD 结构域
Isoform0000836	C-type lectin 8	CTLTD 结构域
Isoform0016929	C-type lectin 7	CTLTD 结构域
Isoform0028928	C-type lectin lectoxin-Lio3-like	CTLTD 结构域
Isoform0050423	C-type lectin domain family 17	CTLTD 结构域
Isoform0066707	C-type lectin isoform 3	REJ
Isoform0080026	C-type lectin isoform 1	CTLTD 结构域
Isoform0015097	C-type lectin domain family 17	FTP
Isoform0001111	C-type lectin	CTLTD 结构域
Isoform0006712	C-type lectin	CTLTD 结构域
Isoform0009334	C-type lectin	CTLTD 结构域
Isoform0009711	C-type lectin	—
Isoform0010194	C-type lectin	—
Isoform0010594	C-type lectin	CTLTD 结构域
Isoform0010611	C-type lectin	CTLTD 结构域
Isoform0011410	C-type lectin	—
Isoform0011681	C-type lectin	CUB 结构域
Isoform0013319	C-type lectin	CTLTD 结构域
Isoform0013683	C-type lectin	—
Isoform0014936	C-type lectin	CUB 结构域
Isoform0016790	C-type lectin	CTLTD 结构域
Isoform0017277	C-type lectin	CCP 结构域
Isoform0017960	C-type lectin	CTLTD 结构域
Isoform0018879	C-type lectin	CUB 结构域
Isoform0019968	C-type lectin	CCP 结构域
Isoform0020006	C-type lectin	CTLTD 结构域
Isoform0021924	C-type lectin	CCP 结构域
Isoform0028029	C-type lectin	CTLTD 结构域
Isoform0031196	C-type lectin	CTLTD 结构域
Isoform0032523	C-type lectin	CTLTD 结构域
Isoform0034067	C-type lectin	CUB 结构域
Isoform0034757	C-type lectin	CUB 结构域
Isoform0036042	C-type lectin	CTLTD 结构域
Isoform0041354	C-type lectin	CTLTD 结构域
Isoform0042024	C-type lectin	CCP 结构域
Isoform0042186	C-type lectin	CA 结构域
Isoform0042950	C-type lectin	CCP 结构域
Isoform0047481	C-type lectin	CTLTD 结构域
Isoform0048787	C-type lectin	—
Isoform0049720	C-type lectin	CTLTD 结构域
Isoform0050311	C-type lectin	CTLTD 结构域
Isoform0050510	C-type lectin	—
Isoform0051990	C-type lectin	GAIN 结构域

续表

基因 ID	注释信息	保守结构域
Isoform0053723	C-type lectin	CUB 结构域
Isoform0053790	C-type lectin	CUB 结构域
Isoform0054956	C-type lectin	CUB 结构域
Isoform0057321	C-type lectin	CCP 结构域
Isoform0057716	C-type lectin	CUB 结构域
Isoform0058976	C-type lectin	CUB 结构域
Isoform0061957	C-type lectin	CUB 结构域
Isoform0064154	C-type lectin	CCP、EGF 结构域
Isoform0068521	C-type lectin	CCP 结构域
Isoform0074531	C-type lectin	CTLD 结构域
Isoform0077736	C-type lectin	CTLD 结构域
Isoform0080116	C-type lectin	CCP 结构域
Isoform0082337	C-type lectin	—

2.6 丝氨酸蛋白酶

补体系统三条激活途径具有不同的丝氨酸蛋白酶分子, MASP/C1r/C1s 在凝集素途径和经典途径中发挥作用, 而替代途径中则是 Bf/C2。在马氏珠母贝全长转录组数据中共检索到 17 个丝氨酸蛋白酶基因,

其中, 6 个甘露糖结合型凝集素丝氨酸蛋白酶含有 IG_like、Tryp_SPc 结构域; 2 个甘露糖结合型丝氨酸蛋白酶含有 CUB、IG_like、Tryp_SPc 结构域; 2 个丝氨酸蛋白酶含有 Trypsin_2 结构域; 7 个丝氨酸蛋白酶基因含有 N 端 β -螺旋结构域(表 4)。

表 4 马氏珠母贝丝氨酸蛋白酶保守结构域分析
Tab.4 Analysis of the conserved domains in serine protease of *P. fucata*

基因 ID	注释信息	结构域
Isoform0009961	mannan-binding lectin serine protease 1	IG_like、Tryp_SPc
Isoform0017290	mannan-binding lectin serine protease 1	IG_like、Tryp_SPc
Isoform0021114	mannan-binding lectin serine protease 1	IG_like、Tryp_SPc
Isoform0026673	mannan-binding lectin serine protease 1	IG_like、Tryp_SPc
Isoform0038356	mannan-binding lectin serine protease 1	IG_like、Tryp_SPc
Isoform0058600	mannan-binding lectin serine protease 1	CUB、IG_like、Tryp_SPc
Isoform0059099	mannan-binding lectin serine protease 1	IG_like、Tryp_SPc
Isoform0080823	mannan-binding lectin serine protease 1	CUB、IG_like、Tryp_SPc
Isoform0031455	Serine protease	Trypsin_2
Isoform0032197	Serine protease	Trypsin_2
Isoform0044225	Serine protease	N 端 β -螺旋结构域
Isoform0049848	Serine protease	N 端 β -螺旋结构域
Isoform0054163	Serine protease	N 端 β -螺旋结构域
Isoform0065400	Serine protease	N 端 β -螺旋结构域
Isoform0073592	Serine protease	N 端 β -螺旋结构域
Isoform0074112	Serine protease	N 端 β -螺旋结构域
Isoform0075574	Serine protease	N 端 β -螺旋结构域

2.7 其他补体样成分

马氏珠母贝全长转录组数据中共检索到 2 个含硫酯结构域蛋白, 包括 C3 分子和 TEP (thioester-containing protein) 分子。保守结构域分析发现 C3 具

有 complement_C3_C4_C5、A2M_comp 结构域; TEP 分子具有 A2M_2、A2M_comp、A2M_recep 结构域(表 5)。D 因子具有 Tryp_SPc 结构域, 而 H 因子和补体受体 2 (Cr2) 未检测到相应的结构域(表 5)。

表 5 马氏珠母贝其它补体样成分保守结构域分析
Tab.5 Analysis of the conserved domains in other complement-like components of *P. fucata*

基因 ID	注释信息	结构域
Isoform0018832	Complement C3	Complement_C3_C4_C5、A2M_comp
Isoform0081916	thioester-containing protein 1	A2M_2、A2M_comp、A2M_recep
Isoform0001492	Complement factor D	Tryp_SPc
Isoform0050510	Complement factor H	—
Isoform0065377	Complement receptor type 2	—

2.8 补体相关基因表达水平分析

随机挑选 12 个补体样相关基因进行实时荧光定量 PCR, 检测其在溶藻弧菌刺激前后的表达水平变化(图 3)。结果显示, 与对照组相比, 11 个基因的表达水平上调, 其中 10 个基因表现出极显著差异, 包括 *C1q*、*C-lectin*、*MBL*、*ficolin*、*MASP*、*Hf-like* 等补体基因, 暗示马氏珠母贝的补体系统是一个复杂的免疫效应系统, 包括并不限于这些基因参与到马氏珠母贝的免疫响应过程中。

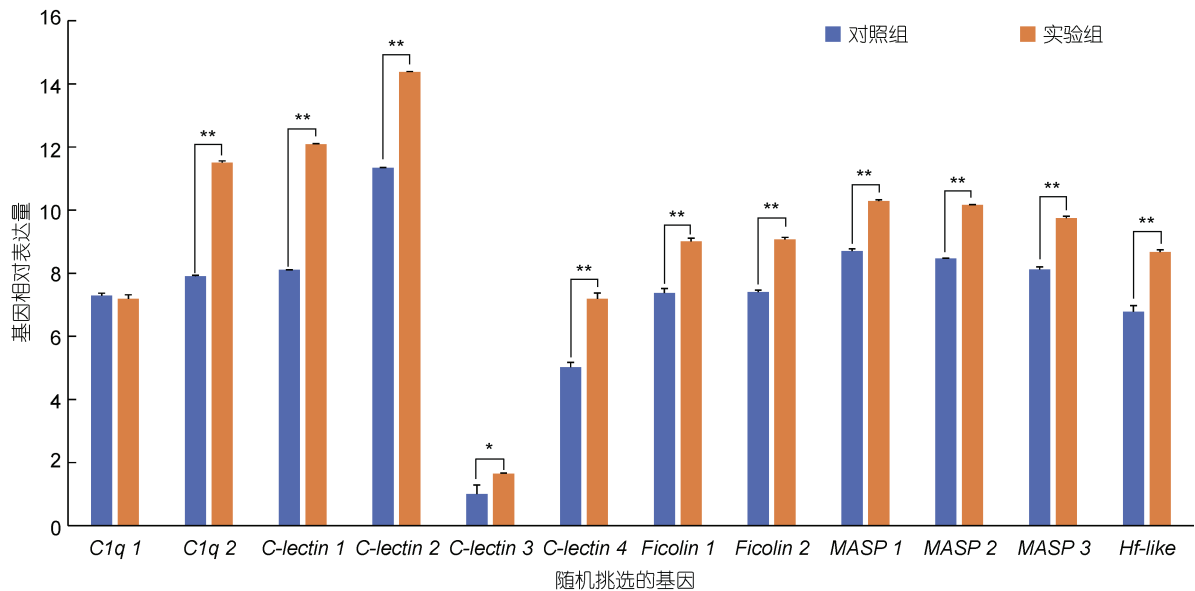


图 3 马氏珠母贝补体样组分基因表达水平

Fig.3 Expression level of complement component gene of *P. fucata*

目前, 关于无脊椎动物补体系统的研究主要集中在海胆(Al-Sharif *et al*, 1998)、海鞘(Ji *et al*, 1997)、鲎(Zhu *et al*, 2005)、珊瑚(Coral) (Dishaw *et al*, 2005)、太平洋牡蛎(*C. gigas*) (Wang *et al*, 2017)等物种中, 而关于马氏珠母贝补体系统的研究尚且处于起步阶段。本研究基于马氏珠母贝全长转录组数据, 系统检索补体相关组分, 结果显示共挖掘到 212 条补体样成分 unigene, 分别编码 89 个含 C1q 结构域蛋白、57 个 C

3 讨论

补体系统是一种古老的免疫防御机制, 在无脊椎动物抵御病原入侵中发挥重要的免疫作用(Nonaka *et al*, 2004a)。补体组分是具有特定结构域分子, 研究表明在哺乳动物补体成分中发现的大多数结构域普遍存在于后口动物, 这表明后口动物体内已经存在由补体分子共同发挥免疫作用的效应系统(Patthy, 2003)。研究表明无脊椎动物的补体系统与脊椎动物的补体系统存在一定差异(Nonaka *et al*, 2006)。

型凝集素蛋白、33 个纤维胶凝蛋白、11 个纤维蛋白原相关蛋白、8 个甘露糖结合型凝集素关联丝氨酸蛋白酶、2 个含硫酯蛋白(1 个 C3 分子, 1 个 TEP 分子)、1 个补体受体、2 个补体因子、9 个丝氨酸蛋白酶。这些补体同源分子的发现暗示它们参与构成马氏珠母贝原始的补体系统, 可能具有类似于脊椎动物补体组分相似的免疫学功能。

补体 C3 作为补体系统的中心组分, 在经典途径、

凝集素途径和替代途径中具有枢纽作用, 因此大部分研究人员将 C3 分子或类似物的出现作为补体系统产生的主要标志(Nonaka *et al*, 2004b)。在无脊椎动物海参(*Apostichopus japonicus*) (Zhou *et al*, 2011)、珊瑚(*Swiftia exserta*) (Dishaw *et al*, 2005)、海葵(*Nematostella vectensis*) (Kimura *et al*, 2009)、文蛤(Prado-Alvarez *et al*, 2009)中发现的 C3 分子具有原始 C3 的结构特征。补体 C3 具有独特的链内硫酯键, 可与入侵微生物形成共价键, 促进吞噬作用(Wang *et al*, 2021)。节肢动物鲎、乌贼、珊瑚中 C3 样分子的发现, 将其起源推至 13 亿年前的寒武纪, 表明这类分子可能在后生动物进化的最早期就已具备调理作用(张岷, 2010)。在鲎中发现的 C3 属于脊椎动物 C3 的同源分子, 其硫酯键水解后能够促进与金黄色葡萄球菌(*Staphylococcus aureus*)、大肠杆菌(*Escherichia coli*)等微生物表面结合(Zhu *et al*, 2005)。南美白对虾(*Litopenaeus vannamei*) C3 在细菌溶解和溶血方面起着至关重要的免疫作用, 激活包括 C3 在内的补体途径可能会限制 WSSV 在对虾体内的复制(Liu *et al*, 2019)。在前期研究中, 我们成功从马氏珠母贝中克隆得到完整的 C3 分子, 其能够在体外条件下激活 NF- κ B 信号通路, 特异性的 siRNA 敲降 C3 的表达能够显著降低马氏珠母贝血细胞对溶藻弧菌的吞噬作用(Wang *et al*, 2021)。

替代途径可通过 C3 分子的激活而启动, 在没有抗体的情况下可通过补充调节因子(H 因子和 P 因子)的相互作用直接结合病原体表面(Wang *et al*, 2017)。哺乳动物 Bf 和 C2 在补体激活过程中具有重要的作用, 其具有相似结构域, 是通过基因复制产生的嵌合蛋白, 包括 short consensus repeat (SCR), von Willebrand factor type A (vWA), serine protease domain-containing (SPC) (张岷, 2010)。研究发现 Bf 和 C2 仅在两栖动物和哺乳动物中同时存在(Ohta *et al*, 2006)。目前在无脊椎动物体内发现的具有类似结构域分子均是 Bf 和 C2 的同源分子, 在结构上同哺乳动物存在一定差异。文蛤中 Bf 因子相比脊椎动物缺少 1 个 SCR 结构域, 而海胆 Bf 具有 5 个 SCR 结构域(Smith *et al*, 1998); 玻璃海鞘(*Ciona intestinalis*) Bf 因子除了具有脊椎动物典型的结构域, 还额外具有 2 个 Low-density lipoprotein receptor (LDLR)和 1 个 SCR (Yoshizaki *et al*, 2005)。文昌鱼(*Branchiostoma belcheri*) Bf/C2 分子在 SCR 重复结构域中包含 1 个表皮生长因子(EGF)结构域, LPS 刺激后在肝脏中的表达量显著上升(He

et al, 2008)。虽然在马氏珠母贝全长转录组数据中发现了大量的 SCR、SPC 结构域, 但是没有发现同时具有 SCR、SPC 结构域的基因, 表明没有典型的 Bf/C2 分子存在于马氏珠母贝中, 其分子功能仍需要进一步验证。

原始补体组分的功能进化及整合促进了补体系统调理作用的产生(Zhu *et al*, 2005)。凝集素途径中 MASP 代替替代途径中 Bf/C2 的作用直接激活 C3 (Fujita *et al*, 2004)。纤维胶凝蛋白(Ficolin)和甘露糖结合型凝集素(MBL)是识别外源病原体并激活凝集素途径中的关键分子(Gadjeva *et al*, 2001)。光滑双脐螺和文昌鱼的 FREP 具有凝集素的活性, 能识别并结合病原微生物, 在先天免疫中发挥着 PRR 的作用(Fan *et al*, 2008; Lu *et al*, 2020)。本研究从全长转录组中检索到 57 个 C-lectin 和 8 个 MASP。C-lectin 含有一个或多个 C 型凝集素结构域(CTLD), 能够与多种配体结合, 包括碳水化合物、脂质、蛋白质、CaCO₃ 等(Huang *et al*, 2008)。MBL 的 C 型凝集素结构域可识别病原体表面上的甘露糖或 N-乙酰氨基葡萄糖结构, 与活化的 MASP 形成复合酶, 进而激活补体系统(Ikeda *et al*, 1987)。缢蛭 C 型凝集素具有介导凝集与吞噬作用的免疫功能(Lan *et al*, 2020)。缢蛭 C-type lectin 具有促进血细胞吞噬细菌的能力, 且这种促吞噬活性是依靠 CTLD 结构域对细菌的结合实现的(蓝天一, 2018)。凝集素途径在缺乏免疫球蛋白的无脊椎动物中具有重要的免疫作用, 海鞘补体系统的凝集素途径的激活产生 C3 片段, 有助于吞噬细胞上结合受体进行吞噬作用(Miyazawa *et al*, 2001)。事实上, 凝集素-MASP 复合物在结构和功能上类似于 C1 复合物, 基于凝集素的补体系统已逐步演变形成经典途径(Matsushita *et al*, 2004)。

在全长转录组数据中共鉴定到 89 个 C1qDC 基因, 其中有 81 个 C1qDC 基因含有 C 末端球状结构域(Globular C1q domain, gC1q), 该结构域作为 C1qDC 的核心组成部分, 主要功能是作为 PRR 直接与细胞外膜上的受体相互作用, 介导细胞吞噬和凝集(Gerdol *et al*, 2011)。大量的基因复制事件导致了 C1qDC 家族成员的扩增, 这可能是促进经典途径产生的重要原因(Nonaka, 2001)。C1qDC 蛋白是经典补体经典途径的起始分子, 是构成 C1 复合物的重要组分; 当 C1qDC 分子与免疫球蛋白 IgM 或 IgG 的 Fc 段结合后, C1qDC 结构发生变化, 进而导致 C1r 和 C1s 的活化, 启动补体经典激活途径(Kishore *et al*,

2004; Yan *et al*, 2020)。研究显示 C1qDC 出现在脊椎动物进化的早期阶段, 在适应性免疫建立和免疫球蛋白出现之前, 原始的 C1qDC 作为补体系统的初始识别分子发挥免疫作用 (Matsushita *et al*, 2004; Nonaka *et al*, 2006)。七鳃鳗 (*Lampetra japonicum*) C1qDC 分子能够结合广泛分布于细菌的细胞壁中的 N-乙酰葡糖胺, 其功能上可能类似于凝集素, 是一种模式识别受体 (Matsushita *et al*, 2004)。太平洋牡蛎 C1qDC-5 可以作为调理素来识别和结合细菌, 并能与血细胞表面的整联蛋白相互作用, 增强牡蛎血细胞的吞噬作用 (Lv *et al*, 2019); C1qDC-6 可以作为模式识别受体, 并介导血细胞的吞噬作用和迁移, 以消除入侵的病原体 (Lv *et al*, 2019); C1qDC-7 对 LPS 和 PGN 以及革兰阴性细菌具有很高的结合亲和力 (Zong *et al*, 2019)。七鳃鳗 C1q-MASP-A 复合物可结合病原体上的碳水化合物导致 C3 的活化及 C3b 的形成, 沉积在病原体的表面上的 C3b 具有调理素的作用 (Nonaka *et al*, 1984)。

随机挑选 12 个补体相关基因进行荧光定量 PCR 检测, 结果显示其中 10 个基因的表达水平经藻弧菌刺激后显著上升, 包括 *C1q*、*C-lectin*、*MBL*、*ficolin*、*MASP*、*Hf-like* 等补体基因, 暗示马氏珠母贝的补体系统是一个复杂的免疫效应系统, 包括并不限于这些基因参与到马氏珠母贝的免疫响应过程中。

4 结论

本研究基于马氏珠母贝全长转录组数据检索到多种 C1qDC、C 型凝集素、纤维胶凝蛋白、甘露糖凝集素、甘露糖结合型凝集素关联丝氨酸蛋白酶以及 C3 等补体相关分子。经实时定量 PCR 检测显示, 多种补体样基因参与到溶藻弧菌刺激后的免疫响应, 暗示马氏珠母贝补体系统是一个复杂的多组分效应系统。*C1q*、*C-lectin*、*MBL*、*ficolin*、*MASP* 等基因呈现出显著差异表达, 表明马氏珠母贝可能倾向于通过凝集途径或类似于凝集素途径执行补体系统的先天免疫作用。以上研究结果为探讨马氏珠母贝体内可能存在的补体系统的激活机制提供了研究基础, 对于进一步加深对马氏珠母贝的免疫防御机制的理解、丰富和发展无脊椎动物补体系统的进化过程具有理论意义。

参 考 文 献

刘先军, 杨丽诗, 黄建华, 等, 2011. 斑节对虾 C1q 结合蛋白 (PmC1qBP) 的克隆及表达特征分析 [J]. 中国水产科学,

- 18(4): 774-781.
- 吴勇, 张野, 许尤厚, 等, 2019. 马氏珠母贝 (*Pinctada fucata*) 血细胞 RNA-Seq 转录组数据中补体样组分分析 [J]. 海洋与湖沼, 50(6): 1343-1353.
- 张岷, 2010. 扇贝补体样成分的基因及其功能研究 [D]. 青岛: 中国科学院研究生院 (海洋研究所): 32-34.
- 陈永霞, 2018. 脂多糖及铜离子胁迫下厚壳贻贝补体 C3 分子表达研究 [D]. 舟山: 浙江海洋大学: 8-12.
- 焦钰, 梁金婵, 谷泽丰, 等, 2020. 脂多糖刺激对马氏珠母贝血细胞 DNA 甲基化修饰的影响 [J]. 广东海洋大学学报, 40(3): 7-13.
- 蓝天一, 2018. 缢蛏 C 型凝集素基因 ScCTL-1 和 ScCTL-2 的克隆鉴定与免疫功能研究 [D]. 上海: 上海海洋大学: 5-6.
- AL-SHARIF W Z, SUNYER J O, LAMBRIS J D, *et al*, 1998. *Sea urchin* coelomocytes specifically express a homologue of the complement component C3 [J]. *Journal of Immunology*, 160(6): 2983-2997.
- CHAO Y H, YUAN J B, GUO T, *et al*, 2019. Analysis of transcripts and splice isoforms in *Medicago sativa* L. by single-molecule long-read sequencing [J]. *Plant Molecular Biology*, 99(3): 219-235.
- DISHAW L J, SMITH S L, BIGGER C H, 2005. Characterization of a C3-like cDNA in a coral: phylogenetic implications [J]. *Immunogenetics*, 57(7): 535-548.
- FAN C X, ZHANG S C, LI L, *et al*, 2008. Fibrinogen-related protein from amphioxus *Branchiostoma belcheri* is a multivalent pattern recognition receptor with a bacteriolytic activity [J]. *Molecular Immunology*, 45(12): 3338-3346.
- FUJITA T, ENDO Y, NONAKA M, 2004. Primitive complement system-recognition and activation [J]. *Molecular Immunology*, 41(2/3): 103-111.
- GADJEVA M, THIEL S, JENSENIUS J C, 2001. The mannan-binding-lectin pathway of the innate immune response [J]. *Current Opinion in Immunology*, 13(1): 74-78.
- GERDOL M, MANFRIN C, DE MORO G, *et al*, 2011. The C1q domain containing proteins of the Mediterranean mussel *Mytilus galloprovincialis*: a widespread and diverse family of immune-related molecules [J]. *Developmental & Comparative Immunology*, 35(6): 635-643.
- GERDOL M, VENIER P, 2015. An updated molecular basis for mussel immunity [J]. *Fish & Shellfish Immunology*, 46(1): 17-38.
- HE Y N, TANG B, ZHANG S C, *et al*, 2008. Molecular and immunochemical demonstration of a novel member of Bf/C2 homolog in amphioxus *Branchiostoma belcheri*: implications for involvement of hepatic cecum in acute phase response [J]. *Fish & Shellfish Immunology*, 24(6): 768-778.
- HUANG S F, YUAN S C, GUO L, *et al*, 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity [J]. *Genome Research*, 18(7): 1112-1126.
- HUBER-LANG M, SARMA J V, ZETOUNE F S, *et al*, 2006. Generation of C5a in the absence of C3: a new complement activation pathway [J]. *Nature Medicine*, 12(6): 682-687.
- IKEDA K, SANNOH T, KAWASAKI N, *et al*, 1987. Serum lectin

- with known structure activates complement through the classical pathway [J]. *Journal of Biological Chemistry*, 262(16): 7451-7454.
- JI X, AZUMI K, SASAKI M, *et al*, 1997. Ancient origin of the complement lectin pathway revealed by molecular cloning of mannan binding protein-associated serine protease from a urochordate, the Japanese ascidian, *Halocynthia roretzi* [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12): 6340-6345.
- KIMURA A, SAKAGUCHI E, NONAKA M, 2009. Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, *Nematostella vectensis* [J]. *Immunobiology*, 214(3): 165-178.
- KISHORE U, GABORIAUD C, WATERS P, *et al*, 2004. C1q and tumor necrosis factor superfamily: modularity and versatility [J]. *Trends in Immunology*, 25(10): 551-561.
- LAN T Y, LI Z, PENG M X, *et al*, 2020. A four-CRD C-type lectin from razor clam *Sinonovacula constricta* mediates agglutination and phagocytosis [J]. *Gene*, 728: 144287.
- LI Y, NIU D H, BAI Y Q, *et al*, 2019. Identification of a novel C1q complement component in razor clam *Sinonovacula constricta* and its role in antibacterial activity [J]. *Fish & Shellfish Immunology*, 87: 193-201.
- LIU Y C, SONG Q Z, LI D L, *et al*, 2019. A novel complement C3 like gene (*Lv-C3L*) from *Litopenaeus vannamei* with bacteriolytic and hemolytic activities and its role in antiviral immune response [J]. *Fish & Shellfish Immunology*, 91: 376-387.
- LU L J, LOKER E S, ADEMA C M, *et al*, 2020. Genomic and transcriptional analysis of genes containing fibrinogen and IgSF domains in the schistosome vector *Biomphalaria glabrata*, with emphasis on the differential responses of snails susceptible or resistant to *Schistosoma mansoni* [J]. *Plos Neglected Tropical Diseases*, 14(10): e0008780.
- LV Z, WANG L L, JIA Z H, *et al*, 2019. Hemolymph C1qDC promotes the phagocytosis of oyster *Crassostrea gigas* hemocytes by interacting with the membrane receptor β -integrin [J]. *Developmental & Comparative Immunology*, 98: 42-53.
- MATSUSHITA M, MATSUSHITA A, ENDO Y, *et al*, 2004. Origin of the classical complement pathway: lamprey orthologue of mammalian C1q acts as a lectin [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 101(27): 10127-10131.
- MIYAZAWA S, AZUMI K, NONAKA M, 2001. Cloning and characterization of integrin α subunits from the solitary ascidian, *Halocynthia roretzi* [J]. *The Journal of Immunology*, 166(3): 1710-1715.
- NONAKA M, 2001. Evolution of the complement system [J]. *Current Opinion in Immunology*, 13(1): 69-73.
- NONAKA M, FUJII T, KAIDOH T, *et al*, 1984. Purification of a lamprey complement protein homologous to the third component of the mammalian complement system [J]. *Journal of Immunology*, 133(6): 3242-3249.
- NONAKA M, KIMURA A, 2006. Genomic view of the evolution of the complement system [J]. *Immunogenetics*, 58(9): 701-713.
- NONAKA M, YOSHIZAKI F, 2004a. Evolution of the complement system [J]. *Molecular Immunology*, 40(12): 897-902.
- NONAKA M, YOSHIZAKI F, 2004b. Primitive complement system of invertebrates [J]. *Immunological Reviews*, 198(1): 203-215.
- OHTA Y, GOETZ W, HOSSAIN M Z, *et al*, 2006. Ancestral organization of the MHC revealed in the amphibian *Xenopus* [J]. *Journal of Immunology*, 176(6): 3674-3685.
- PATTHY L, 2003. Modular assembly of genes and the evolution of new functions [J]. *Genetica*, 118(2/3): 217-231.
- PENG M X, NIU D H, CHEN Z Y, *et al*, 2017. Expression of a novel complement C3 gene in the razor clam *Sinonovacula constricta* and its role in innate immune response and hemolysis [J]. *Developmental & Comparative Immunology*, 73: 184-192.
- PRADO-ALVAREZ M, ROTLLANT J, GESTAL C, *et al*, 2009. Characterization of a C3 and a factor B-like in the carpet-shell clam, *Ruditapes decussatus* [J]. *Fish & Shellfish Immunology*, 26(2): 305-315.
- SEKINE H, KENJO A, AZUMI K, *et al*, 2001. An ancient lectin-dependent complement system in an ascidian: novel lectin isolated from the plasma of the solitary ascidian, *Halocynthia roretzi* [J]. *Journal of Immunology*, 167(8): 4504-4510.
- SMITH L C, SHIH C S, DACHENHAUSEN S G, 1998. Coelomocytes express SpBf, a homologue of factor B, the second component in the sea urchin complement system [J]. *Journal of Immunology*, 161(12): 6784-6793.
- SUCKALE J, SIM R B, DODDS A W, 2005. Evolution of innate immune systems [J]. *Biochemistry and Molecular Biology Education*, 33(3): 177-183.
- SUZUKI M M, SATOH N, NONAKA M, 2002. C6-like and C3-like molecules from the cephalochordate, amphioxus, suggest a cytolytic complement system in invertebrates [J]. *Journal of Molecular Evolution*, 54(5): 671-679.
- TAKAHASHI K G, KURODA T, MUROGA K, 2008. Purification and antibacterial characterization of a novel isoform of the Manila clam lectin (MCL-4) from the plasma of the Manila clam, *Ruditapes philippinarum* [J]. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 150(1): 45-52.
- WANG L L, ZHANG H, WANG L L, *et al*, 2017. The RNA-seq analysis suggests a potential multi-component complement system in oyster *Crassostrea gigas* [J]. *Developmental & Comparative Immunology*, 76: 209-219.
- WANG N, QIN M T, CHEN X H, *et al*, 2019. Molecular cloning of complement component C3 gene from pearl mussel, *Hyriopsis cumingii* and analysis of the gene expression in response to tissue transplantation [J]. *Fish & Shellfish Immunology*, 94: 288-293.
- WANG Z L, LIANG X R, LI G Y, *et al*, 2021. Molecular

- characterization of complement component 3 (C3) in the pearl oyster *Pinctada fucata* improves our understanding of the primitive complement system in bivalve [J]. *Frontiers in Immunology*, 12: 652805.
- WANG Z L, WANG B, CHEN G, *et al*, 2016. Transcriptome analysis of the pearl oyster (*Pinctada fucata*) hemocytes in response to *Vibrio alginolyticus* infection [J]. *Gene*, 575(2): 421-428.
- YAN F F, ZHOU E X, LIU S, *et al*, 2020. Complement C1q subunit molecules from *Xenopus laevis* possess conserved function in C1q-immunoglobulin interaction [J]. *Developmental & Comparative Immunology*, 103: 103532.
- YANG C Y, WANG L L, ZHANG H, *et al*, 2014. A new fibrinogen-related protein from *Argopecten irradians* (AiFREP-2) with broad recognition spectrum and bacteria agglutination activity [J]. *Fish & Shellfish Immunology*, 38(1): 221-229.
- YANG J L, WANG L L, ZHANG H, *et al*, 2011. C-type lectin in *Chlamys farreri* (CfLec-1) mediating immune recognition and opsonization [J]. *PLoS One*, 6(2): e17089.
- YOSHIZAKI F Y, IKAWA S, SATAKE M, *et al*, 2005. Structure and the evolutionary implication of the triplicated complement factor B genes of a urochordate ascidian, *Ciona intestinalis* [J]. *Immunogenetics*, 56(12): 930-942.
- ZHOU Z C, SUN D P, YANG A F, *et al*, 2011. Molecular characterization and expression analysis of a complement component 3 in the sea cucumber (*Apostichopus japonicus*) [J]. *Fish & Shellfish Immunology*, 31(4): 540-547.
- ZHU Y, THANGAMANI S, HO B, *et al*, 2005. The ancient origin of the complement system [J]. *The EMBO Journal*, 24(2): 382-394.
- ZONG Y N, LIU Z Q, WU Z J, *et al*, 2019. A novel globular C1q domain containing protein (C1qDC-7) from *Crassostrea gigas* acts as pattern recognition receptor with broad recognition spectrum [J]. *Fish & Shellfish Immunology*, 84: 920-926.

IDENTIFICATION AND ANALYSIS OF COMPLEMENT-LIKE COMPONENTS IN *PINCTADA FUCATA* BASED ON THE FULL-LENGTH TRANSCRIPTOME SEQUENCE DATASETS

WANG Jing, LI Gui-Ying, LIN Kai-Qi, OU Jing, ZHANG Shu-Ping, WANG Zhong-Liang
(Fisheries College, Guangdong Ocean University, Zhanjiang 524088, China)

Abstract The complement system in animals is a complex and restricted proteolytic system, and plays a defensive role in the immune system. To explore the composition and mechanism of the complement system in *Pinctada fucata*, full-length transcriptome sequencing, gene alignment, and functional annotation were performed. Homology comparison and conserved domains analysis showed a total of 212 genes of potential complement components were found in *P. fucata*, including 89 C1q domain-containing proteins, 57 C-type lectins, 33 ficolins and 11 fibrinogen-related proteins, 8 mannan-binding lectin-associated serine proteases, 2 thioester-containing proteins (1 C3 and 1 TEP), 1 complement receptor, 2 complement factors and 9 serine proteases. Real-time PCR analysis showed that the expression levels of *C1q*, *C-lectin*, *MBL*, *ficolin*, and *MASP* were up-regulated significantly after *Vibrio alginolyticus* challenge, indicating that the complement system of *P. fucata* is a complex multi-component effective system, and lectin pathway or lectin-like pathway may be used to active the complement system with the immune responses. This study provided molecular evidence for further understanding on the original complement system and the immune defense mechanism of *P. fucata*, as well as the marine invertebrate immunology.

Key words *Pinctada fucata*; full-length transcriptome sequence; complement components; bioinformatical analysis; real-time PCR