

文章编号: 1000-0615(2018)05-0633-13

DOI: 10.11964/jfc.20170510846

中国明对虾 *Dscam* 基因的克隆及其在免疫致敏(类免疫)诱导反应中的表达分析

曹家旺^{1,2}, 孟宪红^{1,2*}, 孔杰^{1,2}, 史晓丽^{1,2},
栾生^{1,2}, 罗坤^{1,2}, 董丽君^{1,2}, 陈宝龙^{1,2}

(1. 中国水产科学研究院黄海水产研究所, 农业部海洋渔业可持续发展重点实验室, 山东青岛 266071;
2. 青岛海洋科学与技术国家实验室, 海洋渔业科学与食物产出过程功能实验室, 山东青岛 266237)

摘要: 为了研究对虾免疫致敏过程中 *Dscam* 基因的表达规律, 实验采用同源克隆和 RACE 技术获得了中国明对虾 *Dscam* 基因 cDNA 全长序列, 并对该序列进行生物信息学分析。结果显示, 中国明对虾 *Dscam* 基因的 cDNA 全长为 6 624 bp, 其中包括 171 bp 的 5' 端非编码区, 459 bp 的 3' 端非编码区, 开放阅读框的长度为 5 994 bp, 编码 1 996 个氨基酸。推测该基因编码的蛋白含有一个信号肽、10 个 Ig 结构域、6 个 FNIII 功能区、1 个胞质尾区和 1 个跨膜结构域。同源性分析及系统进化分析表明, *Dscam* 基因与节肢动物的 *Dscam* 基因首先聚为一类, 且与凡纳滨对虾的同源性最高, 为 92.4%。连续投喂 6 d 热灭活白斑综合征病毒 (white spot syndrome virus, WSSV) 饵料来诱导免疫致敏反应, 在 0、6 和 12 d 及二次感染后的 12、24、48、72 和 168 h 分别取样, 用 RT-PCR 的方法检测中国明对虾 *Dscam* 的相对表达量。结果显示, 诱导感染组 *Dscam* 基因在第 12 天开始上调, 且与阴性对照组和未诱导感染组差异显著; 二次感染后 24 h, *Dscam* 基因的相对表达量达到最大值, 与阴性对照组和未诱导感染组差异显著; 48 h 后基因表达量开始下降, 但表达量仍高于阴性对照组和未诱导感染组。实验表明, 中国明对虾存在 *Dscam* 基因, 并在免疫致敏过程中发挥重要作用。

关键词: 中国明对虾; *Dscam* 基因; 免疫致敏; 表达分析

中图分类号: Q 785; S 945.1

文献标志码: A

在全球范围内, 对虾养殖病害时有发生, 而流行性疾病的大范围暴发给对虾养殖产业带来巨大损失。20 世纪 80 年代中期, 台湾地区首先暴发了斑节对虾 (*Penaeus monodon*) 杆状病毒病 (monodon baculovirus virus, MBV), 80 年代后期至 90 年代初期全球范围内又先后暴发了黄头病毒病 (yellow head virus, YHV)、传染性皮下和造血器官坏死病毒病 (infectious hypodermal and hematopoietic necrosis virus, IHNV) 和桃拉病毒病 (Taura syndrome virus, TSV) 等^[1-3], 其中致病力最强、造成

经济损失最大的病毒性疾病是 90 年代初在亚洲暴发的白斑综合征病毒病 (white spot syndrome virus, WSSV)^[4]。近年来, 虽然众多学者对 WSSV 的发生和传播机制进行了深入有效的研究, 但依然缺少治疗疾病的有效方法, 抑制疾病暴发的关键仍然是有效预防。

长久以来, 人们认为无脊椎动物的免疫防御仅仅依靠先天性免疫, 缺少脊椎动物中的可诱导产生抗体和记忆细胞的获得性免疫系统。然而, 最新研究发现, 无脊椎动物中同样存在

收稿日期: 2017-05-23 修回日期: 2017-11-04

资助项目: 中国水产科学研究院黄海水产研究所基本科研业务费专项(20603022017001); 国家自然科学基金(31372523); 泰山学者种业团队项目; 国家现代农业产业技术体系 (CARS-48)

通信作者: 孟宪红, E-mail: mengxianhong@ysfri.ac.cn

经诱导产生免疫记忆的机制,这种反应使无脊椎动物的免疫防御体系处在一个紧张有效的“致敏”状态^[5-7]。当无脊椎动物受到病毒感染或受到病毒的某一部分刺激后(如囊膜蛋白,灭活病毒粒子),再次感染同样的病毒,其死亡率和病原致病力显著降低,从而导致无脊椎动物产生“免疫记忆”或“免疫致敏”现象^[8-9]。Dscam(down syndrome cell adhesion molecule)蛋白作为免疫球蛋白超家族的一个成员(immunoglobulin superfamily, IgSF),在无脊椎动物的病原识别和免疫防御中发挥至关重要的作用。据报道,在果蝇(*Drosophila*)中, *Dscam*基因可以通过3个Ig结构域的可变剪切产生38 000种以上的亚型,此基因众多的亚型可能支持无脊椎动物特异性识别不同的病原体,并作为模式识别受体参与到病原体的特异性识别及信号转导过程中,与无脊椎动物的免疫致敏反应密切相关^[10]。

本实验采用同源克隆和RACE技术克隆了中国明对虾(*Fenneropenaeus chinensis*)*Dscam*基因的cDNA全长序列,并分析该基因在免疫致敏诱导前后时空表达的变化。以期进一步了解*Dscam*基因在免疫致敏过程中发挥的作用,为阐明中国明对虾等无脊椎动物免疫致敏机制奠定基础,为对虾等养殖业探寻防止疾病暴发的策略和方法。

1 材料与方法

1.1 实验材料

实验所用中国明对虾取自中国水产科学研究院黄海水产研究所遗传育种中心,平均体质量(0.21±0.08)g,平均体长(24.71±3.90)mm。海水经砂滤井过滤,盐度29±1。实验前暂养3d。

1.2 实验试剂

Trizol Reagent购自Invitrogen公司(美国); PrimeScript™ RT reagent Kit with gDNA Eraser, SMART™ RACE cDNA Amplification Kit (Clontech), LA Taq, SYBR® Premix Ex Taq™, pMD18-T载体、DH5α感受态细胞和琼脂糖胶回收试剂盒(TaKaRa MiniBEST Agarose Gel DNA Extraction Kit Ver.4.0)购自宝生物工程(大连)有限公司;引物由生工生物工程(上海)股份有限公司合成,其余试剂为国产分析纯。

1.3 中国明对虾总RNA提取及cDNA合成

取中国明对虾头胸甲和第1、2腹节约0.1g,置于无RNA酶的破碎管中,加入破碎珠,于组织破碎仪破碎10s,采用Trizol法提取对虾总RNA,采用超微量紫外分光光度计和1.0%琼脂糖凝胶电泳检测总RNA的浓度和完整性。参照PrimeScript™ RT reagent Kit with gDNA Eraser试剂盒的步骤进行反转录,合成第一链cDNA,产物用于PCR扩增和实时荧光定量;使用SMART™ RACE cDNA Amplification Kit (Clontech)反转录合成cDNA第一链,作为中国明对虾*Dscam*基因3'和5'端序列快速扩增的模板。

1.4 中国明对虾*Dscam*基因cDNA全长的克隆

根据GenBank中凡纳滨对虾(*Litopenaeus vannamei*)*Dscam*基因(GQ154653.1)、通讯螯虾(*Pacifastacus leniusculus*)*Dscam*基因(HQ596367.1)和红螯螯虾(*Cherax quadricarinatus*)*Dscam*基因(JX295853.1)的保守序列区设计引物(表1)。

用1%琼脂糖凝胶电泳检测片段大小,用TaKaRa胶回收试剂盒回收目的片段,胶回收后与pMD18-T载体连接,转入到大肠杆菌(*Escherichia coli*)DH5α感受态细胞中。涂板后置于37℃恒温培养箱中培养过夜,次日挑取单菌落接种于1mL液体培养基中(氨苄终浓度为100 μg/mL)振荡培养,用M13引物进行阳性克隆的初步鉴定,送生工生物工程(上海)股份有限公司进行测序。

1.5 中国明对虾*Dscam*基因可变剪切区域测序与分析

设计F1/R1, F2/R2, F3/R3三对引物(表1)对*Dscam*基因的3个可变结构域进行PCR扩增后,将片段连接转入大肠杆菌DH5α,每个片段挑取70个阳性单克隆进行测序。对测序结果进行双向拼接后用MEGA 6.0软件进行排列,统计此区域序列可变种类的数量。

1.6 中国明对虾*Dscam*基因生物信息学分析

将测序所得序列与中间片段进行拼接后即可获得完整的中国明对虾*Dscam*基因cDNA序列。使用ORF finder软件进行开放阅读框的寻找和对可能编码的氨基酸序列进行翻译;利用ProtParam工具(<http://web.expasy.org/protparam/>)预测*Dscam*基因所表达蛋白的基本理化性质(包括氨基酸残基分析及蛋白质等电点、分子量预测等);

表 1 基因克隆及荧光定量引物序列

Tab. 1 The primer sequences of gene cloning and Real-time PCR

引物名 primer	引物序列(5'-3') primer sequence(5'-3')	用途 usage
<i>Dscam</i> F1	GGCAACTCGGCCATCCTCAAG	核心片段验证
<i>Dscam</i> R1	CTTCGCCATTACGGTAAACATA	核心片段验证
<i>Dscam</i> F2	TCCAGTCAACCACCACCAATA	核心片段验证
<i>Dscam</i> R2	CTCGTCGCTCACCATCCTTA	核心片段验证
<i>Dscam</i> F3	TGGAACTCGCCTCGCTAAC	核心片段验证
<i>Dscam</i> R3	TCTCCTGCTCTGCCTCACT	核心片段验证
F1-210	TACCGCTGCCARGCCWCCAACCTCC	同源克隆
R1-1057	CACTGGTACATKCCTTTGTCTC	同源克隆
F2-2005	GTACCWCCWCGYTGGATTGT	同源克隆
R2-3537	GACATRACYARAGCCTTGAC	同源克隆
F3-3120	WGTSTACACYGAATACGCTG	同源克隆
R3-4298	CCAATCTTRTTRTADGCHGAAGCAT	同源克隆
3'Race F1	CGGCATTGAGAACTTGAGACTGCGAGAG	3'Race
3'Race F2	CATGGCTCACAGAACACTTCGTCAACT	3'Race
3'Race F3	CACTGCTAAACGCCTTAATGGTGCTCAC	3'Race
5'Race F1	TGCACTTCCTGGCGTAATCCTC	5'Race
5'Race F2	GGAAGATCAAGGTGCCGTTGGAGAG	5'Race
RT-F3	CATTGGTGCTGGACTGTCTACTGTT	RT-PCR
RT-R3	GTGCTGTCTGATTGCGGTACTGAA	RT-PCR
18s <i>rRNA</i> F	TATACGCTAGTGGAGCTGGAA	RT-PCR
18s <i>rRNA</i> R	GGGGAGGTAGTGACGAAAAAT	RT-PCR
M13-F	CGCCAGGGTTTTCCAGTCACGAC	菌落鉴定
M13-R	AGCGGATAACAATTCACACAGGA	菌落鉴定
F1	GACACCTTCCTCCTCACCAT	Ig2、Ig3结构域测序
R1	CTTCGCCATTACGGTAAACATA	Ig2、Ig3结构域测序
F2	TGGAACTCGCCTCGCTAAC	Ig7结构域测序
R2	TCTCCTGCTCTGCCTCACT	Ig7结构域测序
F3	TCCAGTCAACCACCACCAATA	第三段可变序列测序
R3	CTCGTCGCTCACCATCCTTA	第三段可变序列测序

利用SMART和InterProScan预测和分析蛋白结构域;用DNAMAN软件对几种不同物种的*Dscam*基因所编码的氨基酸进行排序和比对;用MEGA 6.0以NJ法(Neighbor-Joining)构建系统发育进化树。

1.7 热灭活WSSV肌肉及毒饵的制作

取感染WSSV且症状明显的中国明对虾腹节

肌肉, 剁碎成泥状后装入研钵中充分研磨, 加入适量食用红色素并混合均匀, 作为二次感染对虾的毒饵, $-80\text{ }^{\circ}\text{C}$ 冰箱冷藏。用荧光定量PCR(RT-PCR)法来检测其病毒载量, 测得用作毒饵的肌肉组织平均病毒载量为 2.05×10^8 copies/mg。

取部分毒饵, 参考Balasubramanian等^[11]的灭活方

法, 于60 °C恒温培养箱中灭活处理1 h, 作为热灭活WSSV饵料, -20 °C冰箱保存。

1.8 热灭活WSSV诱导及人工二次感染

根据是否进行免疫致敏反应的诱导将实验分为3组, 阴性对照组(未诱导未感染)、未诱导感染组(投喂配合饲料)和诱导感染组(投喂热灭活的感染WSSV的肌肉)。每组设置3个重复(30尾/重复), 阴性对照组和未诱导感染组每天投喂3次商业配合饲料, 诱导感染组在实验的第1~6天每天投喂1次热灭活感染WSSV的肌肉, 2次配合饲料。实验第12天, 除阴性对照组外均采用单尾定量感染WSSV的方法进行人工感染。实验期间水温维持在(23±1.8) °C。在实验开始第0、6和12天, 二次感染WSSV病毒后12、24、48、72和168 h时间点取样, 每个时间点取3尾虾, 保存于液氮中, 提取总RNA反转录后进行相对表达量分析; 全程记录实验虾死亡的时间, 统计累积死亡率。

1.9 中国明对虾*Dscam*基因的表达分析

本实验选取对虾18S *rRNA*作为内参基因, 根据克隆出的中国明对虾*Dscam*基因的全长序列, 在保守区域设计RT-PCR引物RT-F3/RT-R3, 使用ABI 7500实时荧光定量PCR仪对免疫致敏过程中*Dscam*基因表达量进行检测, RT-PCR反应程序: 95 °C预变性30 s; 95 °C变性5 s, 60 °C退火34 s, 75 °C延伸5 min, 40个循环; 75 °C延伸10 min; 总反应体系为20 μL: SYBR Premix Ex *Taq* II(2×) 10 μL, 正反向引物(10 μmol/L)各0.8 μL, ROX Reference Dye II (50×) 0.4 μL, ddH₂O 6.0 μL, cDNA模板2 μL。实验中每个样品做3个技术重复, 取其平均值作为最终结果。基因相对表达量的计算采用 $2^{-\Delta\Delta C_T}$ 法进行分析, 数据以平均值±标准差(mean±SD)表示, 并用SPSS 19.0软件进行单因素方差分析(One-Way ANOVA), 用显著水平 $P<0.05$ 和 $P<0.01$ 分别进行分析。

2 结果

2.1 中国明对虾*Dscam*基因的克隆及序列分析

中国明对虾*Dscam*基因的cDNA全长序列为6 624 bp, 其中5'UTR为171 bp, 3'UTR为459 bp, ORF长5 994 bp, 共编码1 996个氨基酸。3'UTR非编码序列包含终止密码子(TGA)(图1),

GeneBank登录号为KY636384。利用简并引物F1-210/R1-1057, F2-2005/R2-3537, F3-3120/R3-4298获得了长度为848, 1 533和1 179 bp的3个可变区域片段。

2.2 结构域分析

SMART在线预测中国明对虾*Dscam*蛋白结构域, 结果显示, 该基因编码的蛋白含有一个信号肽(1~21 aa)、10个Ig(immunoglobulin, Ig)结构域, 6个FNIII(fibronectin type III)功能区, 1个胞质尾区(cytoplasmic tail, 1 829~1 972 aa), 1个跨膜结构域(transmembrane domain, 1 612~1 629 aa)(图1)。

2.3 可变区域测序

设计3对引物F1/R1, F2/R2, F3/R3分别对Ig2、Ig3, Ig7及第三段可变区域进行测序发现*FcDscam*基因编码的蛋白在Ig2结构域的547~698 bp和Ig3结构域的854~909 bp处存在不同的碱基序列, 两个结构域的可变区域共存在46种不同的组合形式(图2-a)。在Ig7结构域的1 982~2 261 bp处存在16种不同的组合形式(图2-b)。在可变低复杂性区域(low-complexity region, LCR)及胞质尾区(cytoplasmic tail)存在着5个可变区域, 这5个区域有11种不同的组合形式(图2-c)。理论上, 中国明对虾*Dscam*基因编码的蛋白可存在 $46 \times 16 \times 11 = 8 096$ 种不同的亚型。

2.4 蛋白质结构预测

利用ProtParam工具预测*Dscam*基因所表达蛋白的基本理化性质, 结果显示, 该蛋白含有20种基本氨基酸, 其中含量最高的是Val(8.5%), 最低的是Cys(1.4%)分子量为220.06 ku, 分子式为C₉₆₈₃H₁₅₁₁₆N₂₇₃₈O₃₀₁₈S₅₉; 蛋白质中含有带负电荷的残基(Asp+Glu)227个, 带正电荷的残基(Arg+Lys)198个; 理论等电点pI=6.02。对*Dscam*蛋白的二级结构进行分析使用SOPMA工具, 结果显示, *Dscam*预测蛋白包含18.33%的α螺旋(alpha helix)、27.54%的延伸链(extended strand)、10.27%的β转角(beta turn)和43.87%的不规则卷曲(random coil)。

2.5 中国明对虾*Dscam*基因同源性分析和系统进化树构建

氨基酸序列同源性比对结果显示, 中国明对虾*Dscam*基因与甲壳动物亚门动物的同源性较高, 与凡纳滨对虾的同源性为92.4%, 与红螯螯

1 ccagtgaag agtgaacct cagcttacc gctcccgtt caccacggt cagtgccag gcctatatt ttgggctaa tccaagtga tctctagtt acacccccg tggatagat 108
 signal peptide
 M G T T Y M V W A I L L A L T
 109 tgtcgtgga gatatagtg tatggtgtt ataatttcc tccttggcg acaacttcc tccttcacc ATGGGCACT ACCTATATG GTGTGGGCG ATCCTCCTC GCTCTCAACA 216
 L H K A V C D E S G P V I V E E P D N R V D F S N S T G A N I H C S V R
 217 CTGCACAAA GCTGTGTGT GATGAGAGC GGACCCGTC ATCGTAGAA GAACAGAC AACCCGCTG GACTTCAGC AACTCAACC GGCGCCAAC ATTCCTGTC TCTGTTCGA 324
 G R P A P S V V W V R A D N G S A I G V V P G L R M V L S N G T L I F P
 325 GGACGACCC GCCCCTCC GTCTGTTGG VRAA DNG SA I G V V P G L R M V L S N G T L I F P 432
 P F R A E D Y R Q E V H A Q V Y R C Q A S N S H G T V H S R D V H V R A
 433 CCCTTCGCG GCCGAGGAT TACCGCCAG GAAGTGCAC GCCCAGGTC TACCGCTGC CAGGCTCC AACTCCAC GGCACCGTT CACTCCCGC GAGCTCCAC GTGAGGGCA 540
 V V A Q H Y E T E V N N E F V I R G N S A I L K C N I P S F V A D F V S
 541 GTGGTGGCC CAGCACTAC GAGACGGAG GTGAACAAC GAATTCGTC ATCAGGGGC AACTCGGCC ATCCTCAAG TGCAACATT CCCTCCTTC GTCCCGGAC TTCGTCACT 648
 V Q A W V D N D G T A I Y P S K S Y D G K Y L V L P S G E L H I R S V S
 649 GTGCAGGCG TGGTGGAT AATGACGCG ACAGCGATC TATCCCTCC AAGTCTTAT GAACCGAAG TACTGGTGC CTTCCTCC GGCGLACT CACATCCGC TCCGTCAAG 756
 S E D G F K S Y K C R T V H R L T Q E T R L S A T A G R L V I S E P M G
 757 TCCGAGGAC GGCTTCAAG AGTACAAA TGCCGCACC GTGCACCGC CTCACCCAG GAAACCCGC CTCTCCGCC ACCGCCGGA CGTCTGTG ATCTCCGAG CCGATGGGG 864
 S S A P R L P S K D K G D I V E H A A G S V V P L F C E A Q S N P V P R
 865 AGCAGCGCC CCCGCGCTG CCCAGCAA GATAAAGGG GACATTGTG GAGCAGCA GCGGGCTCG GTTGTTCCT CTGTTTTC GAGGCGCAG AGCAACCCC GTCCCTCCG 972
 F R W F K V P E G G R K A A V E L G E R V K Q V G G T L I I R E A K V E
 973 TTCAGTGG TTCAGGTC CCCGAAGGA GGACGCAAG GCAGCTGTG GAATCCGCG GAACCGCTG AAGCAAGT GGTGGAACC CTCATCAT CGTGAAGCC AAGTGGAG 1080
 D S G K Y L C V V N N S V G G E S V E T V L T V T A P L S A Q V E P K V
 1081 GACTCTGCC AAATACCTG TGCCTGTC AACAACCTCT GTTGGAGGC GAGAGGCTC GAGACTGTG CTCATCTGC ACTGCTCC CTAAGTGC CAAGTGA CCACAGGTC 1188
 Q T V E F G R P A T F T C T Y R G N P V K S V T W L K D G I P L N H K E
 1189 CAGACGTT GAATTTGA CGCCAGCA ACCTTTACC TGCACATAC CGAGGCAAC CCTGTAG TCTGTCACT TGGCTCAAG GATGGAAT CCACCTAAC CACAAGGAA 1296
 A V L R I D T V G R E D K G M Y Q C F V R N D Q E S A Q G T A C C L G
 1297 GCTGTCTT CGATGDAC ACCGTTGTT CGTGAGGAC AAAGGAATG TACCAGTGC TTTGTAGG AATGACCAG GAATCAGCA CAAGAACT GCCGAAGCT AAATCTGGA 1404
 G R F E P P Q L I Y T F Q T N T L Q P G P A V F L K C V A A G N P T P E
 1405 GGACGTTT GAGCCACCA CAGCTGATC TACACCTTC CAGACCAAC ACTCTCCAA CCGTAGCA GCAGTATC CTTAAGTGT GTCCAGCT GGCAATCT ACCCCTGAG 1512
 I T W E L D G T R L A N S E R M Q V G Q Y V T V N G E V V S H L N I S A
 1513 ATCACCTGG GAACCTGAT GGAACCTGC CTCGCTAAC TCAGAACGT ATGCAAGT GGACAGTAT GTTACCGTG AATGGCGAA GTTGTCTCC CACTGAAC ATCTCTGCA 1620
 V H T N D G G L Y A C V A S S T V G S V R H A A R L N V Y G L P Y I R P
 1621 GTTCACACA AATGATGA GGCCTCTAT GCTTGTGTT GCTCTCC ACTGTTGGC TCTGTCCGT CATGCTGCC AGGCTTAA AGCTTATGA TCCGCTAC ATCAGGCT 1728
 M D K A A V V A G E N M V V H C P V A G Y P I D S I V W E K N G R M L P
 1729 ATGGACAAG GCTGCAGTT GTTGTGGA GAAACATG GTTGTACAC TGCCCGGT GCTGGATAT CCCATTGAC TCTATTGT TGGAAAAG AATGGCCCG ATGCTGCC 1836
 I N R R Q K T F P N G T L I V E A V Q R S T D Q G R Y T C V A R N S Q A
 1837 ATCAACGCG CGCCAGAG ACATTCCTC AACGGCAC CTCATTGTT GAAGCTGTC CAACGCTCC ACTGACCAA GGAAGATAC ACCTGTGTT GCCCGTAAC AGCCAGGCT 1944
 Y T A R G D L D V Q V M E P P Q L L P I E F G S E A F Y E G D M A O A N
 1945 TACACTGCC CGTGGADL CTGATGTA CAGTCTATG GAACCACT CAGCTAITA CCTATGAA TTTGGTCT GAAGCATT TATGAGGT TATGAGGT CCAGCGCC CAAGCAAC 2052
 R R L R K G D R P V T F S W L Y N G I L L V N T D D T Y I D H M G S R T
 2053 CGTAGGCTG CGTAAGGTT GACCGCCA GTAACATT AGTTGGTA TACAATGG ATACTACTG GTGAACACA GATGACACA TACATTGAC CATATGGC AGCAGGACG 2160
 S I L T L D P V R A H H Q G N Y S C R A V N K A G F T Q V D T T I I V N
 2161 AGCATTCTG ACTCTGAC CCAAGTCCG GCACACCAT CAGGGCAAC TACAGCTGT AGGGCTGTT AATAAGCTT GGATTTACC CAAGTAGAT ACAACTATA ATTGTCAAT 2268
 V P P R W I V E P A D K A F A L G S D A R L E C K A D G F P R P S L G W
 2269 GTACCACCA CGTGGATT GTGGAACCT GCAGACAGC GCCTTGTCT CTGCGAGT CTAGTAGG CTAGAAATG AAAGCAGAT GGTITCCCA CGACCTCT CTCCGATGG 2376
 K K A A G H T P G D Y R D L D V N N P N I K V T G D G T L H I S S I Q K
 2377 AAGAAGCT GCAGGACAT ACCCTGTT GACTATGTT GATCTGAT GTAACAAC CCCAATAT AAGGTACT GGTGATGGA ACTCTCAT ATCAGCAGT ATTCAGAAA 2484
 S H E G Y Y L C E A N N G I G A G L S T V I Y V R V Q A P P Q F K I Q Y
 2485 TCTCATGAA GGATATTAC TTGTGTGAA GCTAACAA GGCATTGTT GCTGGACTG TACTAGTGT ATCTACGTT AGGGTGCA GCTCCCCA CAGTTCAAG ATTCAGTAC 2592
 R N Q T A R H G D D A V L E C E A G G E T P I G I L W S K D K H S V D Q
 2593 CGCAATCAG ACAGCACGC CATGGAGAT GATGCTGTC TTGGAATGT GAGGCGGA GGAAEAC CCCATTGGC ATCCTCTG AGCAAGGAC AAGCAGST GTTGACCAA 2700
 A A E P R Y T I R E E M R G G S V H S S L S I K T T D R T D S A V Y T C
 2701 GCAGTGAA CCAAGTAC ACAATCCGT GAAGAAATG CGTGGTGGC AGTGTTCAC AGCAGCTTT AGCATCAAG ACAACTGAT CGCACTGAC TCTGCTGTC TATACTGT 2808
 V A T N A F G S A D T N I N L I I Q E H P E Q P S S L K V L D K S G R S
 2809 GTGGCTACC AATGCTTTT GGAAGTGA GATACCAAT ATTAACCTC ATCATTCAA GAACACCA GAACAACC AGCAGCTT AAGTCTCT GATAAGAGT GGCCGATCA 2916
 V E L S W T S P Y D G N S P I T R Y I V E Y K L S R R N W E N D G E R M
 2917 GTGGAAGCT TCTTGGAC TCTCCCTAT GATGGAAC TCCCCTAT ACCCGATAC ATTGTGGAA TATAAGCTT AGCCGAGCT AACTGGGAA AATGATGT GAACGCATG 3024
 3025 ATGGTGCT GGAACACG AACATGGCA GCTGTCTG GATCTTCG CCAAGCACT ACCTATCAC CTCGTTATT GTTGCAAGG AATGAGATT GGTGATTCT GAACCATCA 3132
 D I V T I I T A E E A P S G S P R D L K V E A V D Q T S L R V K W K P P
 3133 GACATTGTG ACCATCATC ACTGCCGAG AAGCACCA AGTGCTCA CCAAGTGC CTCAAAGT GAAGCTGTT GACCAGCC TCCTCCGT GTTAAGTG AAGCCACT 3240
 L R E W N G D I Q G Y Q V G Y R L A S S N T S Y V Y E T V E F S K E V
 3241 CTTAGGAAA GAATGAAAT GGAGACATC CAGGGCTAC CAAGTAGGC TACCGTCTC GCATCATCT AATACTCT TATGTCTAT GAACTGTG GAATTCAGC AAGGAAGT 3348
 G K E H Q L T I K K L Q V Y T E Y A V V V S A F N K I G Q G P T T E E I
 3349 GGCAAGAA CATCAACT ACTATCAAG AAGCTGCA GTGTACTT GAATACGCT GTTGTGTT TCTGTCTT ACAAGATT GGCAAGGA CCCACAAC GAGGAAATC 3456
 R S Y T A E G T P Q Q P P H D V T C T T L T S Q T I R V S W A S P P L E
 3457 CGCTCTAC ACAGTGAA GGAACCCC CAACAACC CCTCATGAT GTCACCTGC ACAACTCTG ACTTCTCAG ACCATCCG GTGTCTGG GCCTCACT CCATCTGAG 3564

(图1 Fig.1)

3565 T V Q G V I K G Y K V I Y G P S D K W Y D E E R K D T K I T S S T E T H 3672
 ACTGTACAG GGAGTGATC AAGGGATAC AAAGTAATC TATGGACCA TCTGACAAG TGGTATGAT GAAGAGAGA AAGGACACA AAAATTACT AGTTCCACT GAGACTCAC

3673 L H G L Q K Y T N Y S L Q V L A F T S G G E G V R S Q P I H C Q T D Q D 3780
 CTCCATGGA CTCAGAAAG TACACTAAC TACAGTCTC CAGGTTTTG GCTTTCACA TCAGGTGGT GAAGGAGTA CGATCACAG CCTATTAC TGTACAGACT GATCAGGAT

3781 V P D A P T S V K A L V M S A D S I L V S W L P P D R P N G I I T Q Y T 3888
 GTCCTGAT GCCCCACA TCTGTCAAG GCTCTGGT ATGTCAGCT GACTCCATC CTGTTATCT TGGCTACCC CCAGACCGT CCTAATGGT ATCATTACT CAGTATACT

3889 V Y F K E E G K S D S E A E Q E K L L S S Q L N Y E A N G L K Q R D E Y 3996
 GTATACTC AAGGAAGAG GAAAGAGT GACAGTGAG GCAGAGCAG GAGAAACTT CTATCATCT CAACTGAAC TACGAAGCA AATGGACTT AAACAGCGG GATGAATAT

3997 V F W V T A S T T I G E G E K S E S V H L K L S S K V P A K I A S F D D 4104
 GTCTTTTGG GTGACAGCC TCTACAACC ATTGGAGAA GGAGAGAAG TCTGAATCT GTCCATCTG AAGCTCTCA AGCAAAGTC CCTGCCAAG ATTGCATCC TTTGATGAT

4105 E Y V A T Y K E D V K L H C Q A V G L P T P D I R W T I R G E P F T P N 4212
 GAGTATGTT GCCACATAC AAGGAAGAT GTCAAGCTC CACTGCCAG GCTGTAGGT CTCCCCACC CCTGATATC AGGTGGACA ATCCGTGGT GAACCAATT ACCCCCCAAT

4213 D R M R L L T E G S L L I R E V S R D D A G E Y T C H V E N P Y G Q D T 4320
 GATCGTATG CGTCTTTT ACTGAGGTT TCCTACTG ATTCTGAA GTGTCTCGT GATGATGCT GGAGAGTAC ACCTGCCAT GTTGAGAAT CCCTATGGA CAAGACACA

4321 V T H T L L I Q A P P H P P E I Q F Q S T T T N S I E V K L K P S V I D 4428
 GTCACACAC ACACTCTC ATTCAGGCC CCACCTCAC CCACCTGAG ATTCAAATC CAGTCAACC ACCACCAAT AGCATTGAG GTTAAGCTG AAACCTTCT GTTATTGAT

4429 D T T P I H G Y T V Y Y K P E F S T W E S V Q V P A S T R S Y N L E G L 4536
 GACACCACT CCCATTCAT GGTACTACT GTCTACTAC AAGCTCAG TTTGATACC TGGGAGTGT GTGCAGGTT CCTGCTTCC ACAAGGTC AACAATTTG GAAGCTTGG

4537 W C G S R Y Q I Y A S A Y N K I G T G E S S E I L N T R T K G K K P E I 4644
 TGGTGGT TCTCGTTAC CAGATTAT GCTTCGGCT TACAACAAG ATTGGCACT GGTGAATCT TCTGAGATC TTGAACACG CGCACAAG GCACAAG CCGTGAAT

4645 P E V H R F V E V S S V S I N L H L K A W Q D G G C P M N Y F V V E Y K 4752
 CCTGAAGTT CATCGTTT GTAGAAGT TCCTCTGTC TCTATCAAT CTCCACCTT AAAGCCTGG CAAGATGGT GGTGCCCC ATGAATTAC TTTGTGGT GAATACAAG

4753 P R H Q T E W I M A D N Q V S P T G N Y G I M E L T P A T W Y N L R I S 4860
 CCAAGGCAC CAGACTGAA TGGATAATG GCAGACAAC CAAGTGAG CCAACAGGA AACTATGGT ATCATGGAA CTGACCCCT GCTACCTGG TATAATCTT CGTACTCT

4861 A H N N A G S S V A E Y E V A T L T L T G A T L P P T V S D S R V T W L 4968
 GCTCATAA AATGCTGA TCTCAGTT GCTGAATAT GAAGTTGCT ACATTGACT CTTACTGGA GCCACGCTG CCGCCAAT GTTTCGGAG AGCCGAGTA ACCTGGCT

4969 P D W W P K W L D L N I V P V I A T I V V I I V G I V V I C V A V T R 5076
 CCTGATTGG TGGCCGAG TGGCTGGAT CTGAATGTC CTGGTCCA GTTATTGCC ACTATCTGTC GTCATCATI GTGGGCATT GTTGTATC TGTGTGCA GTAACTCGT

5077 R K N G I E N L R L R E E V Y Q Q Y Q Y N A S M P P P S T M D K R H P G 5184
 CGAAGAAC GGCATTGAG AACTTGACA CTGCGAGAG GAAGTGATC CAGCAGTAC CAATACAAT GCAAGCATG CCCCCACA TCCACACTG GACAAACGC CACCCAGGA

5185 F R E E L G Y I P P P N R K L P P V P G S Q Y N T C D R I K R G G G S G 5292
 TTCGAGAA GAACCTGGC TACATTCA CCAACCAAC CGCAAGCTT CCACTGTC CTGGATCC CAGTATAAC ACCTGTGAC ACCTGTAAG CGAGGTGG GGGTCGGT

5293 R G T H A T W D P R R P M Y E E L S L H P P P G R R I P L G G P P Q P L 5400
 CCGGCACA CATGCCACC TGGGACCA AGACGACC ATGTACGAG GAACCTCC CTCCATCCC CCCCCGGA CGACGTATC CCCCTGGA GGGCCCCA CAACCCCT

5401 G S Q D T L R S G G D D E I C P Y A T F H L L G F R E E M D P Q Q A G N 5508
 GGCTCTCAG GACTCTC AGATCAGGT GGCGATGAC GAGATCTGC CCTATGCT ACCTTCCAC CTCTGGGC TTCCTGAG GAAATGAC CCCCAACA GCTGGCAC

5509 N F Q T F P H Q N G H G S Q Q H F V N S P A S R S M P P S S T Y Y S T V 5616
 AACITCCAG ACTTTCCC CACCAGAAC GGCCATGGC TCACAGCAA CACTTCGTC AACTCCCT GCCTCAAGG AGCATGCC CCAAGCAGC ACCTACTAC TCTACTGTT

5617 P G D M T A S R M S N S T F S P T Y D D P A R S D E E S D Q Y G G S T Y 5724
 CCTGGCGAT ATGACTGCT TCCCGCAT AGCAACTCA ACTTCTCT CCAACTTAC GATGACCT CCCCCTCT GATGAAGAG AGTGACAA TACGGTGA TCTACTTAT

5725 S G G G P Y A R A I D S V S Q S G T A K R L N G A H P P G A P V S G P Q 5832
 TCTGGTGGT GGACCCAT GCTCGAGCC ATTGACTCT GTGTACAG TCTGGCACT GCTAAACGC CTTAATGGT GCTCACCC CCTGGAGCT CCTGTGCT GGGCCTCAA

5833 P S N H R F I S N R G S T S G S A G Q G S P E P P P P P P R N G D L P 5940
 CCCAGCAAC CATAGATT ATATCTAAC CGTGAAGC ACTTCTGA AGCGTGA CAAGGATCC CCTGAGCC CCCCCTCT CCACTCCC CGCAACGA GACCTCCC

5941 L D S S G L G S S L N D S N N S T A S N Q F S E A E C D H D L V Q R N Y 6048
 CTTGATTC TCTGGCTC GTTTCATCT CTGAACGAC TCAAACAC TCCACAGCA TCTAATCAG TTCTCAGAG GCTGAGTGT GATCATGAT CTTGTGCA GGTAAITAT

6049 G V K A T K S T E E M R K L L D K N E A A A H I Q N G G L R M V S D E M 6156
 GGAGTGAAG GCTACCAAG AGCACAGAG GAGATGCG AAGCTCCTG GACAAAAAT GAAGCAGCT GCCCCATT CAAAATGGA GGCCTAAGG ATGGTGAGC GACGAGAT

6157 N V AATGTTGA ggctctcca gaccccact gcaggacct cccaaacc cgtgccata caggcagc ctgattacc gccccatg tcactcac gctgattat tgattctga 6264

6265 acaccaag ctgaacct gagacaaga tctgtact agatatgat atgattata tatatatgt atagacct caccctag acgattac tgaatgag cacagagga 6372

6373 cattaatag cggagcca tatactaca aggaagaag gtagtattt gtgataaaa tgotataga tatcagagg aaattgtc atcatttaa atcttagct ctctaatt 6480

6481 aaaatagt ccagatacc tcattaat acatgata ctccaagca ctg aaataa aggtaaat cagcagcag ttgttctct tgtgtgaa gatcgaata tctgattct 6588

6589 gtgttcata gsgtgacc cacgtctc ttaattgt 6624

图 1 中国明对虾 *Dscam* 基因核酸以及预测氨基酸序列

黑线、红线和蓝线分别标出 Ig2、Ig3 和 Ig7 三个可变结构域；椭圆、红框及菱形分别标出跨膜结构域、胞质尾区及终止密码子(TGA)

Fig. 1 Nucleotide and deduced amino acid sequence of the *Dscam* gene

The three variable structural domains of Ig2, Ig3 and Ig7 were respectively marked by black, red and blue lines; the transmembrane domain, the cytoplasmic tail and the termination codon (TGA) were marked by the ellipse, red box and rhombus

虾、通讯螯虾、普通滨蟹(*Carcinus maenas*)和中华绒螯蟹(*Eriocheir sinensis*)的同源性分别为 87.3%、85.7%、84.3%、82.7%, 而与昆虫纲其他动物的同源性则相对较低(图3)。系统进化树的

结果显示, 中国明对虾先与凡纳滨对虾聚为一支, 然后与甲壳动物亚门的红螯螯虾、通讯螯虾、普通滨蟹和中华绒螯蟹聚为一大支, 之后再和昆虫纲的其他动物聚为一类(图3)。

```

1 HSRD VHRVAVVHQD QLDVPEYV I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
2 HSRD VHRVAVVHQD QLDVPEYV I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
3 HSRD VHRVAVVHQD QLDVPEYV I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
4 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
5 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
6 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
7 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
8 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
9 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
10 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
11 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
12 HSRD VHRVAVVHQD QLDVPEYV I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
13 HSRD VHRVAVVHQD QLDVPEYV I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
14 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
15 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 100
16 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
17 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
18 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
19 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
20 HSRD VHRVAVVQQD LLDV SLE Y I GNSA I LKCN I P F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
21 HSRD VHRVAVVQAQAE I R S D E F I R G N A A I L K C S I P S F V D F V T V Q A W I N D Q S S Y T N K H D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
22 HSRD VHRVAVVQAQAE I R S D E F I R G N A A I L K C S I P S F V D F V T V Q A W I N D Q S S Y T N K H D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
23 HSRD VHRVAVVQAQAE I R S D E F I R G N A A I L K C S I P S F V D F V T V Q A W I N D Q S S Y T N K H D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
24 HSRD VHRVAVVQAQAE I R S D E F I R G N A A I L K C S I P S F V D F V T V Q A W I N D Q S S Y T N K H D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
25 HSRD VHRVAVVHQE E T H V P E S V I R G N A A I L K C N I P S F V D F V K V S V W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
26 HSRD VHRVAVVHQE E T H V P E S V I R G N A A I L K C N I P S F V D F V K V S V W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
27 HSRD VHRVAVVHQE E T H V P E S V I R G N A A I L K C N I P S F V D F V K V S V W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
28 HSRD VHRVAVVQAQ L L D S M E N V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
29 HSRD VHRVAVVQAQ L L D S M E N V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
30 HSRD VHRVAVVQAQ L L D S M E N V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
31 HSRD VHRVAVVQAQ L L D S M E N V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
32 HSRD VHRVAVVQAQ L L D S M E N V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
33 HSRD VHRVAVVQAQ E T S R H D E V L I R G N A A I L K C V I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
34 HSRD VHRVAVVQAQ D L L Q L D E A V I R G N A A I L K C V I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
35 HSRD VHRVAVVQAQ D L L Q L D E A V I R G N A A I L K C V I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
36 HSRD VHRVAVVQAQ D L L Q L D E A V I R G N A A I L K C V I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
37 HSRD VHRVAVVQAQ D L L Q L D E A V I R G N A A I L K C V I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
38 HSRD VHRVAVVHQD P D S D N K E Y V I R G N A A I L K C S I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
39 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
40 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
41 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
42 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
43 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
44 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 100
45 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
46 HSRD VHRVAVVQQD F E V G D E Y V I R G N A A I L K C N I P S F V D F V S V A W T I K E T F Y S E N I D G K Y L V L P S G E L H I R S V S S E D G F K S Y K R T V H R L T Q 99
hsrdvhravv q e v g n a i k c p f v d f v w d g k y l v l p s g e l h i r s v s s e d g f k s y k r t v h r l t q

```

```

1 ETRL SATAGRLVISE V G S S P R P S E E . . . Q L V N K V Q S A T S N A I L C N V I S F V P Q H W F K V P E G G 163
2 ETRL SATAGRLVISE V G S S P R P S E E . . . R G D I L T K E R K S S E A I T C K V A F P L P E H W F K V P E G G 162
3 ETRL SATAGRLVISA V G S S P R P P R D . . . T V S K E V T Q R G S S E S L F C A G A F P P K V W F K V P E G G 163
4 ETRL SATAGRLVISA V G S S P R P L Q A . . . L T N V V R S G I R E L A I C D A A F P L P S E W F K V P E G G 163
5 ETRL SATAGRLVISA V G S S P R P S E E V . . . K L L Q Y T K H R Q Q G C F C K A G Y V P A W F K V P E G G 163
6 ETRL SATAGRLVISE V G S S P R P R E W . . . P I P S S I P A R S S K S V C N L G F F I P E L W F K V P E G G 163
7 ETRL SATAGRLVISA V G S S P R P T I R D . . . K H V T E D K A F G S S A L F C T V A S F V P Q H W F K V P E G G 163
8 ETRL SATAGRLVISA V G S S P R P L Q A . . . L T N V V R S G I R E L A I C D A A F P L P S E W F K V P E G G 163
9 ETRL SATAGRLVISA V G S S P R P S L S E . . . K M G L A S A A D V G A A A L C F A A H P T P G E W F K V P E G G 163
10 ETRL SATAGRLVISA V G S S P R P L Q A . . . L T N V V R S G I R E L A I C D A A F P L P S E W F K V P E G G 163
11 ETRL SATAGRLVISA V G S S P R P A S R D . . . R M G T D R L F G S S E S L F C A G A F P P K V W F K V P E G G 163
12 ETRL SATAGRLVISA V G S S P R P P R D . . . T V S K E V T Q R G S S E S L F C A G A F P P K V W F K V P E G G 163
13 ETRL SATAGRLVISA V G S S P R P V Q A . . . R L N Q F I K A T G E S A L F C B A G Y V P V W F K V P E G G 163
14 ETRL SATAGRLVISA V G S S P R P P R D . . . T V S K E V T Q R G S S E S L F C A G A F P P K V W F K V P E G G 163
15 ETRL SATAGRLVISA V G S S P R P P R D . . . T V S K E V T Q R G S S E S L F C A G A F P P K V W F K V P E G G 163
16 ETRL SATAGRLVISE V G L A P R S L S E . . . K V D A T G T E V N T A T L T C R A A H P P E T W F K V P E G G 163
17 ETRL SATAGRLVISE V G S S P R P S E D . . . L I R T K H P R G S D E A L F C Q S G Y V P G E W F K V P E G G 163
18 ETRL SATAGRLVISE V G S S P R P S E D . . . L I R T K H P R G S D E A L F C Q S G Y V P G E W F K V P E G G 162
19 ETRL SATAGRLVISA V G S S P R P T R T R . . . V P E V T I E T S S A S A L F C P A A Y P P K V W F K V P E G G 163
20 ETRL SATAGRLVISE V G S S P R P T E G E . . . K I D I F K R A R G S S E A L T C R A G F F V P E W F K V P E G G 162
21 ETRL SATAGRLVISE V G L A P R S L S E . . . A T G T E V N T A T L T C R A A H P P E T W F K V P E G G 163
22 ETRL SATAGRLVISE V G L A P R P I L D E . . . L A G T H E G A F A T L H C R A A H P P T P E W F K V P E G G 163
23 ETRL SATAGRLVISE V G S S P R P S E D . . . L I R T K H P R G S D E A L F C Q S G Y V P G E W F K V P E G G 163
24 ETRL SATAGRLVISE V R S L T A P R A L S D . . . K V T G L A S A T A V N D G V L T C D A A H P P T I W F K V P E G G 165
25 ETRL SATAGRLVISA V G S S P R P L Q A . . . L T N V V R S G I R E L A I C D A A F P L P S E W F K V P E G G 163
26 ETRL SATAGRLVISA V G S S P R P P G S D . . . L S R V M G L R G H V T L L C P L G F F V P S H W F K V P E G G 162
27 ETRL SATAGRLVISA V G S S P R P L Q A . . . L T N V V R S G I R E L A I C D A A F P L P S E W F K V P E G G 163
28 ETRL SATAGRLVISE V G S S S P R R S K D . . . T I N A D V E A G T P I S L L C A A H P A P G W F K V P E G G 163
29 ETRL SATAGRLVISE V G S S S P R R S K D . . . T I N A D V E A G T P I S L L C A A H P A P G W F K V P E G G 163
30 ETRL SATAGRLVISA V G L S A P R A S R D . . . R M G T D R L F G S S E S L F C A G A F P P P V T W F K V P E G G 162
31 ETRL SATAGRLVISE V G S S P R P V E V E . . . R I V N L M R A S G S S E A L P C K G G F F V P E W F K V P E G G 163
32 ETRL SATAGRLVISA V G S A A P R P S G K . . . K I E F P S V G F G S A T L L C B A S H P T P K W F K V P E G G 163
33 ETRL SATAGRLVISA V G L S A P R A S R D . . . R M G T D R L F G S S E S L F C A G A F P P P V T W F K V P E G G 162
34 ETRL SATAGRLVISE V G S S P R P A S G A . . . R G D I L T K E R K S S E A I T C K V A F P L P E H W F K V P E G G 162
35 ETRL SATAGRLVISA V G T S P A L S S R E . . . K L I N E V R R S E T D E A L L C N V A N P A P E H W F K V P E G G 163
36 ETRL SATAGRLVISA V G S S P R P T W E E R V H A I L L L K P Q E T A S F C B A G Y V P P E T W F K V P E G G 166
37 ETRL SATAGRLVISA V G T S P A L S S K E . . . K L I N E V R R S E T D E A L L C N V A N P A P E H W F K V P E G G 163
38 ETRL SATAGRLVISE V G L A A P R I S V S E . . . A S A A B V S Q P A T I T C S A A F P P Q H W F K V P E G G 163
39 ETRL SATAGRLVISE V G L A A P R I S V S E . . . R F D V S E A G V G E P S T I T C R A A H P I P E W F K V P E G G 162
40 ETRL SATAGRLVISA V G S S E S P R V S G E . . . Q G V A M V R Q E G A G E S P C P A G C F F I P S E W F K V P E G G 163
41 ETRL SATAGRLVISA V G Y A A P R P N K D . . . V S R T I L V D S G I S S L F C A G A F P I P R W F K V P E G G 163
42 ETRL SATAGRLVISE V G S S E S P K L A G A . . . R G D I F F R P Q G S S E A L P C R V G F F V P E W F K V P E G G 162
43 ETRL SATAGRLVISA V G S S P R P S D D . . . R L H S V T K A S G A D E G L L C B A A Y P P P S E W F K V P E G G 163
44 ETRL SATAGRLVISA V G S S P R P T T R S . . . R S Q T F D V A A G L P T I L C B A A S P P P S E W F K V P E G G 164
45 ETRL SATAGRLVISA V G T L A P R T I R D . . . K H V T E D K A F G S S E A L F C T V A S F V P Q H W F K V P E G G 163
46 ETRL SATAGRLVISA V G L S A P R P N I D . . . L S R T Q R L A G S A F L L C A G A H P A P S E W F K V P E G G 163

```

(a) (图2 Fig.2)


```

1 DVQVMKPEILFAE.DGEIRSKDQIVCFVSDVYPLKIRWYFQEQSHHMGVTTVKLQSPSSILSSEHHTHGSGVITCVSNANGEKRFQASLTVHPPR 104
2 DVQVMKPEILFDHGEPTFFEDLQATCLKRGDLPVIFSNMNGVELSQDDEICLVKGRRTSILLDPQAHQGTICRTVNEGTREVELEIVNPPR 105
3 DVQVMKPEILFAE.GSEAFYEDMAGANCLRKGRDPIFSWLNGLVNTDDTYDHDMSRDLITLDPRAHQGNSCRANVKGFGIVDVTIVNPPR 105
4 DVQVMKPEILFAE.PEEVEAFQIVQVSCAVSGDDPILQWIKDGLPINSANFIDNNVSLPLPLSSDAEHSGLTCLAFNPGVTREASLTKVKNPPR 104
5 DVQVMKPEILFDHGEPTFFEDLQATCLKRGDLPVIFSNMNGVELSQDDEICLVKGRRTSILLDPQAHQGTICRTVNEGTREVELEIVNPPR 105
6 DVQVMKPEILFDHGEPTFFEDLQATCLKRGDLPVIFSNMNGVELSQDDEICLVKGRRTSILLDPQAHQGTICRTVNEGTREVELEIVNPPR 105
7 DVQVMKPEILFDHGEPTFFEDLQATCLKRGDLPVIFSNMNGVELSQDDEICLVKGRRTSILLDPQAHQGTICRTVNEGTREVELEIVNPPR 105
8 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
9 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
10 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
11 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
12 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
13 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
14 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
15 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 105
16 DVQVMKPEILFAE.PDEVASGTPQLQITCVLVGDEPTLQWFKRNLHSDHETLHSLTRVRSLLPSASHSAGLTCRASNMGVQVSSQLKVKVPPR 104
dvqvm p q c p w g c n g v ppr

```

(b)

```

1 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 104
2 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 105
3 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 105
4 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 104
5 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 102
6 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 105
7 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 105
8 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 105
9 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 104
10 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 104
11 NNAGSSVAEYEVATLLTGCILAPAREVP..AFcAGDIPILNLLNLPVWSAIVVILAVIICYLGRG.....NTFEEVYQQYQNASMPPPPSTM 104
nagssvaeeyevatliltg p pl ln vpv vvi v c r n eevyqqyqnasmpppstm

```

```

1 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 192
2 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 192
3 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 192
4 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 192
5 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
6 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
7 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
8 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
9 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
10 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
11 DKRHPGFREELGYIPPPNRKLLPVPGSQYNTCDRIKRGGSGRGTHATWDRPRMYEELSLHPPGRRIPLGPPQPLGSDTLRSGGDEICFYATFHL 200
dkrhpgrfrelgyippnrkllpvpgsqyntcdrikrggsgrgthatwdrprmyeelshppgrriplgppqplgsdrtlrsggdeicfyatfhl

```

```

1 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 265
2 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 258
3 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 251
4 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 233
5 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 233
6 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 265
7 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 285
8 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 241
9 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 241
10 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 258
11 LGFREEMDQQAGNNFQTFPHQNGHGSQQHFVNSPASRSMPRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG 258
lgfreemdqqagnnfqtphqngghsqqhfvnspasrsmpRHGSG.....NYYSVCVAGIYGP.....TCGHTPNEGHPRHGSG

```

```

1 NYYSVCVAGIYGP.....GCPSSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLT..... 348
2 .....GCPSSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLTNGAHPGAPV..... 338
3 .....GCPSSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLTNGAHPGAPV..... 331
4 .....PSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLTNGAHPGAPV..... 311
5 .....PSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLT..... 301
6 .....HQPSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLTNGAHPGAPV..... 346
7 .....GCPSSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLT..... 356
8 .....PSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLT..... 309
9 .....PSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLTNGAHPGAPV..... 319
10 .....GCPSSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLTNGAHPGAPV..... 339
11 .....GCPSSSTYYSTVPGDMTASRMSNSTFSPTYDDPARSDEESDQYGGSTYSGGGPYARAIDSVSQGTAKRLT..... 329
psstyystvpgdmtasrmsnstfsptyddparsdeesdqyggstysgggpyaraidsvsqgtakrlt

```

```

1 .....NRGSTSGSAGGQSPPEPPPPPP 370
2 GPOFSNHRFISNRGSTSGSAGGQSPPEPPPPPP 371
3 GPOFSNHRFISNRGSTSGSAGGQSPPEPPPPPP 364
4 GPOFSNHRFISNRGSTSGSAGGQSPPEPPPPPP 344
5 .....NRGSTSGSAGGQSPPEPPPPPP 323
6 GPOFSNHRFISNRGSTSGSAGGQSPPEPPPPPP 379
7 .....NRGSTSGSAGGQSPPEPPPPPP 378
8 .....NRGSTSGSAGGQSPPEPPPPPP 331
9 GPOFSNHRFISNRGSTSGSAGGQSPPEPPPPPP 352
10 GPOFSNHRFISNRGSTSGSAGGQSPPEPPPPPP 372
11 .....NRGSTSGSAGGQSPPEPPPPPP 351
nrgstsgsagqgspepppppp

```

(c)

图2 中国明对虾Dscam蛋白可变区域多氨基酸序列排列

(a)为Ig2和Ig3结构域, (b)为Ig7结构域, (c)为可变低复杂性区域及胞质尾区结构域

Fig. 2 Multiple amino acid sequence alignments of the Dscam protein variable regions

(a) the Ig2 and Ig 3 domains, (b) the Ig7 domain, (c) the low-complexity region and the cytoplasmic tail

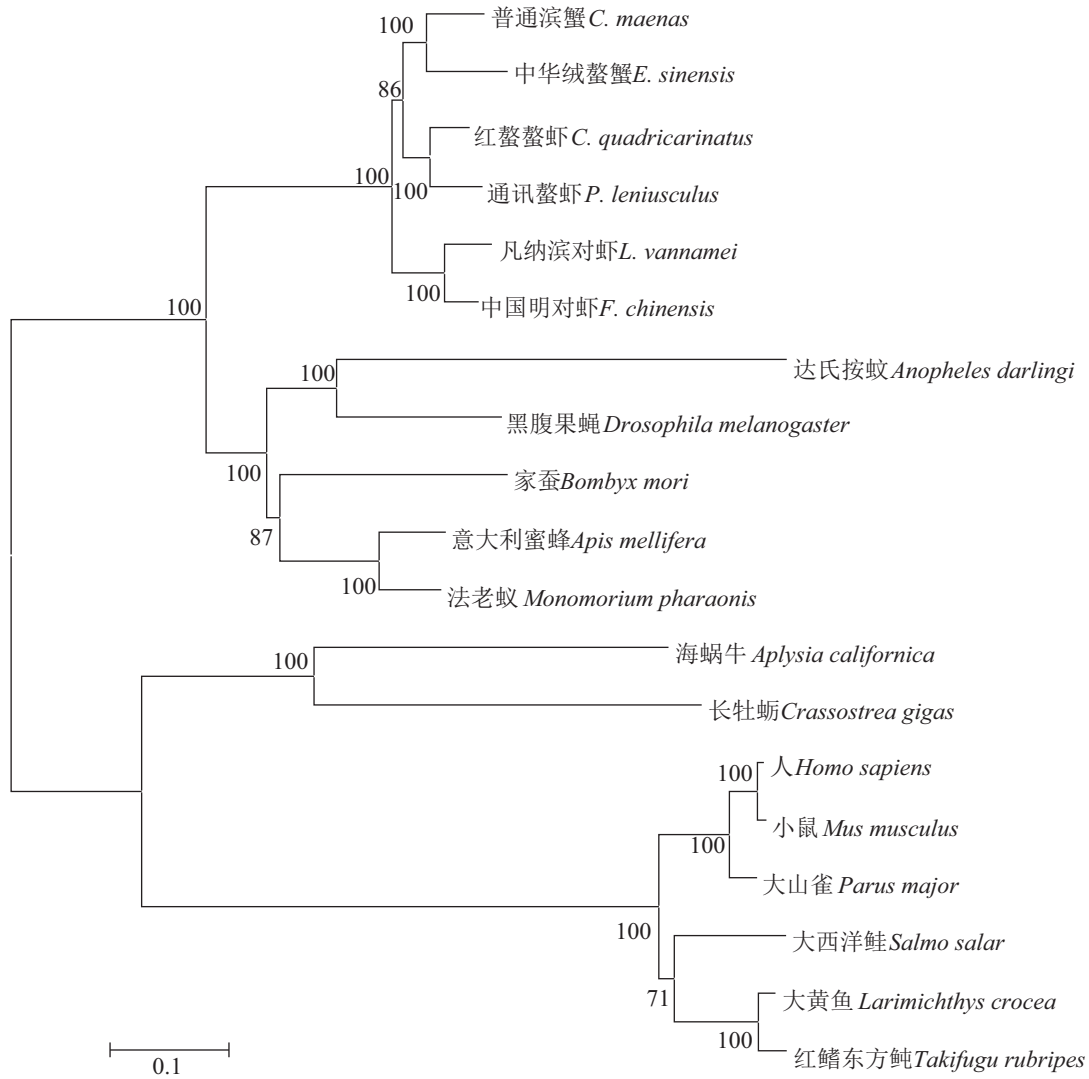


图 3 *Dscam* 氨基酸序列的系统进化树

物种名称及 Genebank ID: 普通滨蟹 *C. maenas* CDO91660.1; 中华绒螯蟹 *E. sinensis* AGL39311.1; 红螯螯虾 *C. quadricarinatus* AGK90306.1; 通讯螯虾 *P. leniusculus* AEC50084.1; 凡纳滨对虾 *L. vannamei* ACZ26466.1; 中国明对虾 *F. chinensis* KY636384; 达氏按蚊 *Anopheles darlingi* ETN63184.1; 黑腹果蝇 *Drosophila melanogaster* AAF71926.1; 家蚕 *Bombyx mori* XP_012548607.1; 意大利蜜蜂 *Apis mellifera* AAT96374.1; 法老蚁 *Monomorium pharaonis* XP_012526097.1; 海蜗牛 *Aplysia californica* ABS30432.1; 长牡蛎 *Crassostrea gigas* AFW98341.1; 人 *Homo sapiens* AAF27525.1; 小鼠 *Mus musculus* NP_112451.1; 大山雀 *Parus major* XP_015480990.1; 大西洋鲑 *Salmo salar* XP_013993288.1; 大黄鱼 *Larimichthys crocea* KKF19481.1; 红鳍东方鲀 *Takifugu rubripes* XP_003968332.1

Fig. 3 Phylogenetic tree of *Dscam* amino acid sequences

The reference sequences were as follows (GenBank ID): *Carcinus maenas* (CDO91660.1); *Eriocheir sinensis* (AGL39311.1); *Cherax quadricarinatus* (AGK90306.1); *Pacifastacus leniusculus* (AEC50084.1); *Litopenaeus vannamei* (ACZ26466.1); *Fenneropenaeus chinensis* (KY636384); *Anopheles darlingi* (ETN63184.1); *Drosophila melanogaster* (AAF71926.1); *Bombyx mori* (XP_012548607.1); *Apis mellifera* (AAT96374.1); *Monomorium pharaonis* (XP_012526097.1); *Aplysia californica* (ABS30432.1); *Crassostrea gigas* (AFW98341.1); *Homo sapiens* (AAF27525.1); *Mus musculus* (NP_112451.1); *Parus major* (XP_015480990.1); *Salmo salar* (XP_013993288.1); *Larimichthys crocea* (KKF19481.1); *Takifugu rubripes* (XP_003968332.1)

2.6 中国明对虾 *Dscam* 基因免疫致敏诱导及二次感染后表达量变化

用 RT-PCR 分析中国明对虾 *Dscam* 基因在热灭活 WSSV 诱导免疫致敏反应阶段和二次感染后表达量变化。在肝胰腺混合部分肌肉的样品的

组织中, 各个时期均有表达。

在 0~12 d 免疫致敏反应的诱导阶段, 未诱导感染组和阴性对照组中国明对虾 *Dscam* 基因的相对表达量维持相对恒定, 诱导感染组在连续投喂 6 d 经热灭活感染 WSSV 的肌肉后, 中国明对虾

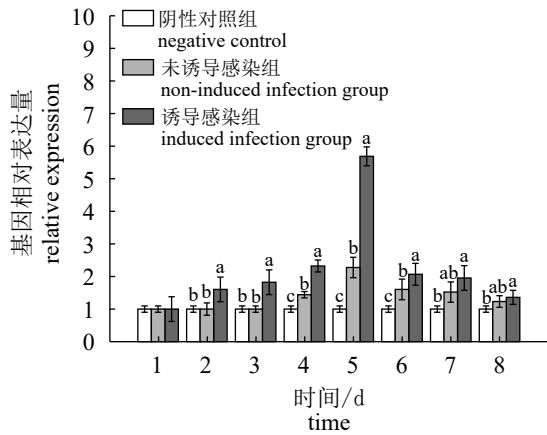


图4 免疫诱导阶段及二次感染后*Dscam*基因的表达
诱导阶段: 1~3. 分别表示诱导0、6和12 d; 感染阶段: 4~8. 分别表示感染0.5、1、2、3和7 d

Fig. 4 The expression profiles of *Dscam* after immune induction and subsequent rechallenge

Infection stage: 1~3. represent inducing 0, 6 and 12 days, respectively; infection stage: 4~8. represent infected 0.5, 1, 2, 3 and 7 days

*Dscam*基因的相对表达量逐渐提高。实验第6天, 诱导感染组的相对表达量高于其余2组 ($P>0.05$); 实验第12天(二次感染0 h), 诱导感染组*Dscam*基因的相对表达量为未诱导感染组的1.94倍, 差异显著 ($P<0.05$)(图4)。

二次经口感染WSSV后12 h, 诱导感染组*Dscam*基因的相对表达量是未诱导感染组的1.64倍 ($P<0.05$)。二次感染24 h后, 诱导感染组为未诱导感染组的2.49倍 ($P<0.05$)。48 h诱导感染组的表达量为未诱导感染组的1.28倍 ($P<0.05$)。72 h, 诱导感染组和未诱导感染组*Dscam*基因相对表达量都降至较低水平 ($P>0.05$)(图4)。

在免疫诱导阶段, 饲喂感染经热灭活WSSV的肌肉的诱导感染组未出现死亡, 阴性对照组和未诱导感染组死亡率同样为0。二次感染后24 h, 未诱导感染组出现死亡; 48 h后, 诱导感染组开始出现死亡。自二次感染WSSV后, 在同一时间点, 未诱导感染组的累积死亡率始终高于诱导感染组。二次感染后168 h, 未诱导感染组的累积死亡率为91.11%, 诱导感染组的累积死亡率为66.67%, 两组累积死亡率存在显著性差异 ($P<0.05$)(图5)

3 讨论

在无脊椎动物中, 普遍被接受的观点是它们缺少适应性免疫, 但是有大量的证据说明在这些无脊椎动物中存在着特异性的“免疫致敏”反

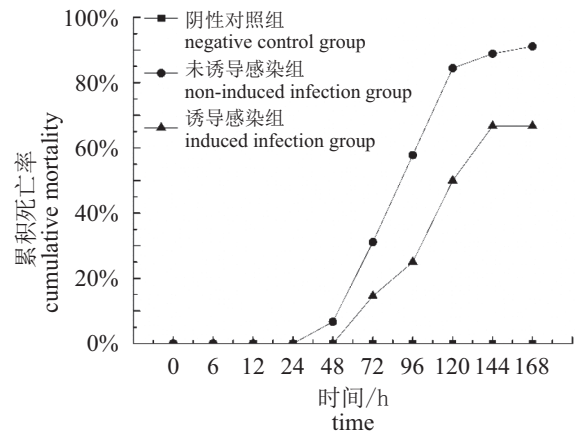


图5 二次感染后各组累积死亡率

Fig. 5 The cumulative mortality of the groups after subsequent rechallenge

应, 也有学者称之为先天性免疫记忆(innate immune memory)或者特异性先天免疫(innate immune specificity)^[9, 12]。研究发现, 唐氏综合征细胞黏附分子*Dscam*可能支撑了这种特异性免疫。无脊椎动物的*Dscam*基因首先在果蝇S2细胞中被发现, 起初发现它与神经细胞轴突的形成有密切关系^[13], 后来的研究证明*Dscam*基因可以被不同的病原体诱发产生不同的亚型, 可以增加免疫细胞吞噬的效率, 同时被认为与病原体的免疫识别有关系^[14-16]。

本研究通过同源克隆和RACE的技术方法获得了中国明对虾的*Dscam*基因, 由于*Dscam*基因的自身特性及其功能的特殊性, 尤其是可以被不同病原体诱发产生多样性, 决定了其cDNA及编码氨基酸的特殊性。

3.1 中国明对虾*Dscam*基因的可变剪切

在*Dscam*基因的克隆过程中, 我们发现其Ig2、Ig3、Ig7等区域具有不同的可变剪切, 这与Wang等^[14]的研究结果类似, 在中华绒螯蟹的*Dscam*基因中, 这3种结构域分别检测出了44、39、31种可变剪切, 理论上可以产生53 196($44 \times 39 \times 31 = 53 196$)种*Dscam*亚型。Chou等^[17]发现在斑节对虾*Dscam*基因中, 这3种结构域中分别有28、43、19种可变剪切, 理论上产生22 876($28 \times 43 \times 19 = 22 876$)种亚型。于爱清等^[18]不仅在这3种结构域中发现了30、33、19种可变剪切, 而且还在胞质尾区结构域中发现2种不同的可变剪切, 在水蚤(*Daphnia*)的胞质尾区也发现了4种不同的亚型。但本实验对多个可变区域整合到一个片段

中进行测序, 结果发现, Ig2和Ig3区域同时进行分析, 降低了可变区域可变剪切的数量。Ig2和Ig3区域只存在46种不同的可变剪切。而且实验还发现, 除Ig2、Ig3和Ig7等3个区域存在可变剪切外, 在基因末端的低复杂性区域和胞质尾区还存在11种不同的可变剪切组合。纵观整个*FcDscam*基因的可变剪切区域, 理论上可形成 $46 \times 16 \times 11 = 8\ 096$ 种亚型, 此研究结果与Wang等^[14]的结论有差异, 但这些研究结果都说明*Dscam*基因可产生大量不同的亚型, 且这仅仅是在RNA水平上分析可以产生的不同的信使RNA, 还未曾分析蛋白质形成过程中一级结构及高级结构形成过程中复杂的变化; 且*Dscam*基因不同亚型的形成还受*caper*, *IRSF1*, *serine/arginine (SR)-rich proteinB52*, *hrp36*等基因的调节, 这些基因通过调节可变剪切来影响*Dscam*的多样性^[19-20]。这些因素都决定了*Dscam*基因作为无脊椎动物免疫相关基因的多变性和复杂性。

3.2 中国明对虾*Dscam*基因的结构域

对中国明对虾*Dscam*基因ORF区进行生物信息学分析, 发现中国明对虾*Dscam*基因编码的蛋白质为典型的跨膜结构形式, 包括一个胞外结构域, 一个跨膜结构域和一个胞质尾区, 这与红螯螯虾和黑腹果蝇的*Dscam*蛋白结构一致^[14, 18]。但凡纳滨对虾、中华绒螯蟹和斑节对虾缺少胞质尾区和跨膜区域^[14, 17, 21]。说明*Dscam*蛋白可能存在2种不同的形式, 一种是以膜结合的方式固定在血细胞膜中, 另一种是以分泌型分布在血清中。研究认为, 这2种存在方式都是*Dscam*的亚型, 它们在免疫防御中行使不同的功能。我们推测, 膜结合型*Dscam*通过胞外区域特异性的与病原体或亲同种抗原的亚型结合, 识别不同的病原体, 膜结合型*Dscam*的胞质尾区通过与尾部相连的结构域调节后续的信号转导^[17]。在经过复杂的识别和效应机制后, 膜结合蛋白被水解成为分泌型的*Dscam*蛋白进入血淋巴, 参与病原体的吞噬作用^[22-23]。这与脊椎动物中IgM的存在状况类似, IgM也存在着两种形式: 膜结合型和分泌型^[24-25], 这更充分说明*Dscam*基因与免疫致敏的相关性。

3.3 中国明对虾*Dscam*基因时空表达规律

热灭活WSSV饵料连续对诱导感染组进行6d的投喂后, *Dscam*基因的表达量开始上调, 截止

到实验第12天第二次感染WSSV前, 诱导感染组和未诱导感染组*Dscam*基因的相对表达量出现显著差异($P < 0.05$)。二次感染WSSV后, 诱导产生免疫致敏反应的诱导感染组在12 h左右, 基因的相对表达量显著上升, 至24 h后达到最高值。研究发现, 用WSSV悬液注射感染红螯螯虾, 在72 h内*Dscam*的表达量与注射PBS缓冲溶液的对照组没有显著性差异, 5 d后两者的基因表达水平出现显著性差异, 实验组明显高于注射PBS缓冲溶液的对照组^[12]。本研究中, 连续6 d饲喂热灭活WSSV肌肉诱导免疫致敏反应, 诱导结束后6 d(实验第12天), 诱导感染组相对基因表达量显著高于未诱导感染组。这与在红螯螯虾上的研究结果类似, 虽然红螯螯虾*Dscam*可以被病原体刺激表达, 但是大量表达是在感染病原体2~5 d。与此相比, 对虾抗菌肽、crustins和C型凝集素等基因在感染WSSV后, 24~36 h就达到了表达量的最高峰且持续时间较短^[26-27]。Chiang等^[19]的研究发现, 在斑节对虾受到WSSV感染后, 膜结合和分泌型*Dscam*在24 h时表达量都增加了10倍以上, 但在72 h后达到了最大表达量^[19], 这也说明*Dscam*不同于直接发挥作用的先天性免疫因子, 在中国明对虾的免疫防御过程中发挥其他重要作用。

本研究获得了中国明对虾*Dscam*基因cDNA序列全长并分析了该基因在经热灭活WSSV诱导免疫致敏反应后的表达变化, 为探究甲壳动物的免疫致敏机制奠定了理论基础。

参考文献:

- [1] Flegel T W. Detection of major penaeid shrimp viruses in Asia, a historical perspective with emphasis on Thailand[J]. *Aquaculture*, 2006, 258(1-4): 1-33.
- [2] Flegel T W, Nielsen L, Thamavit V, *et al.* Presence of multiple viruses in non-diseased, cultivated shrimp at harvest[J]. *Aquaculture*, 2004, 240(1-4): 55-68.
- [3] Lightner D V. Virus diseases of farmed shrimp in the Western Hemisphere (the Americas): a review[J]. *Journal of Invertebrate Pathology*, 2011, 106(1): 110-130.
- [4] Wang Q, Poulos B T, Lightner D V. Protein analysis of geographic isolates of shrimp white spot syndrome virus[J]. *Archives of Virology*, 2000, 145(2): 263-274.
- [5] Loker E S, Adema C M, Zhang S M, *et al.* Invertebrate immune systems-not homogeneous, not simple, not well understood[J]. *Immunological Reviews*, 2004, 198(1):

- 10-24.
- [6] Kurtz J, Franz K. Innate defence: Evidence for memory in invertebrate immunity[J]. *Nature*, 2003, 425(6953): 37-38.
- [7] Kurtz J. Specific memory within innate immune systems[J]. *Trends in Immunology*, 2005, 26(4): 186-192.
- [8] Little T J, Kraaijeveld A R. Ecological and evolutionary implications of immunological priming in invertebrates [J]. *Trends in Ecology & Evolution*, 2004, 19(2): 58-60.
- [9] Sadd B M, Schmid-Hempel P. Insect immunity shows specificity in protection upon secondary pathogen exposure[J]. *Current Biology*, 2006, 16(12): 1206-1210.
- [10] Wojtowicz W M, Flanagan J J, Millard S S, *et al.* Alternative splicing of *drosophila* dscam generates axon guidance receptors that exhibit isoform-specific homophilic binding[J]. *Cell*, 2004, 118(5): 619-633.
- [11] Balasubramanian G, Sudhakaran R, Musthaq S S, *et al.* Studies on the inactivation of white spot syndrome virus of shrimp by physical and chemical treatments, and seaweed extracts tested in marine and freshwater animal models [J]. *Journal of Fish Diseases*, 2006, 29(9): 569-572.
- [12] Ng T H, Hung H Y, Chiang Y A, *et al.* WSSV-induced crayfish Dscam shows durable immune behavior[J]. *Fish & Shellfish Immunology*, 2014, 40(1): 78-90.
- [13] Schmucker D, Clemens J C, Shu H, *et al.* *Drosophila* dscam is an axon guidance receptor exhibiting extraordinary molecular diversity[J]. *Cell*, 2000, 101(6): 671-684.
- [14] Wang J J, Wang L L, Gao Y, *et al.* A tailless Dscam from *Eriocheir sinensis* diversified by alternative splicing[J]. *Fish & Shellfish Immunology*, 2013, 35(2): 249-261.
- [15] Chou P H, Chang H S, Chen I T, *et al.* The putative invertebrate adaptive immune protein *Litopenaeus vannamei* Dscam (LvDscam) is the first reported Dscam to lack a transmembrane domain and cytoplasmic tail[J]. *Developmental & Comparative Immunology*, 2009, 33(12): 1258-1267.
- [16] Watthanasurorot A, Jiravanichpaisal P, Liu H P, *et al.* Bacteria-induced dscam isoforms of the crustacean, *Pacifastacus leniusculus*[J]. *PLoS Pathogens*, 2011, 7(6): e1002062.
- [17] Chou P H, Chang H S, Chen I T, *et al.* *Penaeus monodon* Dscam (PmDscam) has a highly diverse cytoplasmic tail and is the first membrane-bound shrimp Dscam to be reported[J]. *Fish & Shellfish Immunology*, 2011, 30(4-5): 1109-1123.
- [18] 于爱清. 虾蟹类免疫相关基因的研究[D]. 上海: 华东师范大学, 2014.
- Yu A Q. Study on immune related genes of crayfish and crabs[D]. Shanghai: East China Normal University, 2014 (in chinese)
- [19] Chiang Y A, Hung H Y, Lee C W, *et al.* Shrimp Dscam and its cytoplasmic tail splicing activator serine/arginine (SR)-rich protein B52 were both induced after white spot syndrome virus challenge[J]. *Fish & Shellfish Immunology*, 2013, 34(1): 209-219.
- [20] Lee C W, Chen I T, Chou P H, *et al.* Heterogeneous nuclear ribonucleoprotein hrp36 acts as an alternative splicing repressor in *Litopenaeus vannamei* Dscam[J]. *Developmental & Comparative Immunology*, 2012, 36(1): 10-20.
- [21] Hung H Y, Ng T H, Lin J H, *et al.* Properties of *Litopenaeus vannamei* Dscam (LvDscam) isoforms related to specific pathogen recognition[J]. *Fish & Shellfish Immunology*, 2013, 35(4): 1272-1281.
- [22] Watson F L, Püttmann-Holgado R, Thomas F, *et al.* Extensive diversity of Ig-superfamily proteins in the immune system of insects[J]. *Science*, 2005, 309(5742): 1874-1878.
- [23] Dong Y M, Taylor H E, Dimopoulos G. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the anopheles gambiae innate immune system[J]. *PLoS Biology*, 2006, 4(7): e229.
- [24] El Shikh M E, El Sayed R M, Szakal A K, *et al.* T-independent antibody responses to T-dependent antigens: a novel follicular dendritic cell-dependent activity[J]. *The Journal of Immunology*, 2009, 182(6): 3482-3491.
- [25] Mcheyzer-Williams L J, Mcheyzer-Williams M G. Antigen-specific memory B cell development[J]. *Annual Review of Immunology*, 2005, 23: 487-513.
- [26] Bulet P, Uttenweiler-Joseph S, Moniatte M, *et al.* Differential display of peptides induced during the immune response of *Drosophila*: a matrix-assisted laser desorption ionization time-of-flight mass spectrometry study[J]. *Journal of Protein Chemistry*, 1998, 17(6): 528-529.
- [27] Aoki T, Wang H C, Unajak S, *et al.* Microarray analyses of shrimp immune responses[J]. *Marine Biotechnology*, 2011, 13(4): 629-638.

Cloning and expression analysis of the *Dscam* gene during the inducing immune priming (Quasi-immune) response in *Fenneropenaeus chinensis*

CAO Jiawang^{1,2}, MENG Xianhong^{1,2*}, KONG Jie^{1,2}, SHI Xiaoli^{1,2},
LUAN Sheng^{1,2}, LUO Kun^{1,2}, DONG Lijun^{1,2}, CHEN Baolong^{1,2}

(1. Key Laboratory for Sustainable Utilization of Marine Fisheries Resources, Ministry of Agriculture, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China;

2. Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China)

Abstract: To explore the function of *Dscam* gene of Chinese shrimp *Fenneropenaeus chinensis* (*FcDscam*) during the immune priming response, we cloned the cDNA sequence using the rapid amplification of cDNA ends (RACE), and then analyzed the features of this gene based on bioinformatics software. The full length of *FcDscam* cDNA is 6624 bp, with a 5'-untranslated region (UTR) of 171 bp, a 3'-untranslated region (UTR) of 459 bp, and an opening reading frame (ORF) of 5 994 bp, and it encoded 1 996 amino acids. The protein encoded by *FcDscam* contains 1 signal peptide, 1 cytoplasmic tail, 1 transmembrane domain, 6 of fibronectin type (FN) and 10 of immunoglobulin(Ig). As shown by the homology search and neighbor-joining phylogenetic tree, *FcDscam* gene had the highest homology with that of *Litopenaeus vannamei* (92.4%). To induce the immune response, test animals were first fed with heat-inactivated WSSV bait for six days and infected again 12 days later. Samples were collected at 0, 6 and 12 d of the first-time infection and 12, 24, 48, 72 and 168 h of the second-time infection, respectively, for analysing the mRNA expression level of *FcDscam* by real-time PCR. On the 12th day of the first infection, its expression level in induced infection group began to increase, significantly different from those in the negative control group and the non-induced infection group. After 24 h of the second-time infection, its expression level in the infection group reached a maximum, significantly different from those of the negative control group and the non-induced infection group, and the expression level started to decline at the 48 h. Our research demonstrated the existence of the *Dscam* gene in *F. chinensis*, and its important role in the process of immune sensitization.

Key words: *Fenneropenaeus chinensis*; *Dscam*; immune priming; expression analysis

Corresponding author: MENG Xianhong. E-mail: mengxianhong@ysfri.ac.cn

Funding projects: Central Public-interest Scientific Institution Basal Research Fund, Yellow Sea Fisheries Research Institute (20603022017001); National Natural Science Foundation of China (31372523); The Taishan Scholar Program for Seed Industry; China Agriculture Research System (CARS-48)