文章编号: 1674 − 7054(2022)05 − 0440 − 11

斑节对虾 *PmML1* 和 *PmML2* 基因的鉴定及表达分析

符丹凤[1], 刘洪涛[2], 杨明秋[2], 何玉贵[2]
(1. 海南省水产技术推广站, 海口 570226;
2. 海南省海洋与渔业科学院/海南省热带海水养殖技术重点实验室, 海口 571126)

摘　要: 为进一步了解 ML 家族基因在斑节对虾先天免疫中的作用, 采用 PCR 和 RACE 方法克隆到斑节对虾 (*Penaeus monodon*)2 个 ML 家族基因 *PmML1* 和 *PmML2*, 并对其进行生物信息学分析, 同时通过实时荧光定量 PCR 技术分析 2 个 ML 家族基因在斑节对虾各组织的表达分布。结果显示: *PmML1* 的开放阅读框 (Open Reading Frame, ORF) 为 462 bp, 编码 153 个氨基酸; *PmML2* 的 ORF 为 471 bp, 编码 156 个氨基酸。2 个基因都具有 1 个典型的 ML 家族蛋白结构域。多重比对结果显示, 2 个基因的 ML 结构域与组成 3 对二硫键的半胱氨酸残基位置都相对保守。氨基酸序列对比结果显示, PmML1 和 PmML2 之间的相似性仅为 42.37%。PmML1 与日本囊对虾 (*Marsupenaeus japonicus*) 的 ML 蛋白 MjML4 的相似性最高, 比例高达 91.5%, PmML2 与凡纳滨对虾 (*Litopenaeus vannamei*) 的 ML 蛋白相似性最高, 比例高达 88.46%。系统进化分析结果显示, *PmML1* 和 *PmML2* 存在于 2 个不同的分支, 说明 2 个基因的进化相对独立, 证明本研究中鉴定到的是 2 个不同的基因。实时荧光定量 PCR 结果显示, *PmML1* 在对虾眼柄中的表达量最高, 其次为肠、胃、心、鳃、肌肉和肝胰腺, 在血细胞的表达水平最低; *PmML2* 在眼柄中表达量最高, 其次为肠、鳃、胃、血细胞、心、肝胰腺和肌肉, 在肌肉中的表达水平最低。眼柄是重要的神经内分泌调节器官, 推测 *PmML1* 和 *PmML2* 可能在斑节对虾神经内分泌调控等过程中发挥重要作用。

关键词: 斑节对虾; ML 基因; 克隆; 组织表达
中图分类号: S 917.4　　**文献标志码:** A
引用格式: 符丹凤, 刘洪涛, 杨明秋, 等. 斑节对虾 *PmML1* 和 *PmML2* 基因的鉴定及表达分析 [J]. 热带生物学报, 2022, 13(5): 440−450. DOI: 10.15886/j.cnki.rdswxb.2022.05.003

斑节对虾 (*Penaeus monodon*) 属甲壳纲 (Crustacea) 十足目 (Decapoda) 对虾科 (Penaeidae) 对虾属 (*Penaeus*), 又名黑虎虾、草虾, 是当今世界三大经济对虾养殖品种之一。斑节对虾是对虾属中体型最大的种类, 分布广泛, 主要在印度洋以及太平洋区域[1]。自 20 世纪 80 年代起, 斑节对虾养殖规模逐渐扩大。近年来, 养殖环境污染引起的多种细菌病、病毒病等疾病大规模爆发, 严重制约了斑节对虾养殖产业的健康发展。其中, 细菌病主要是由副溶血弧菌 (*Vibrio parahaemolyticus*)、哈维弧菌 (*V. harveyi*) 等十几种病原菌引发[2]。有研究显示, 细菌脂多糖 (LPS) 是几乎所有革兰氏阴性细菌外表面膜的主要成分, 并且是多种真核生物先天免疫反应的极强刺激剂[3]。对虾的先天性免疫

收稿日期: 2022 − 01 − 20　　**修回日期:** 2022 − 03 − 31
基金项目: 2017 年海南省本级部门预算项目(琼财预 [2017] 233 号); 财政部和农业农村部“国家现代农业产业技术体系项目”(CARS-48); 2021 年海南省创新能力建设计划省重点实验室建设补助经费项目(琼科 [2021] 295 号)
第一作者: 符丹凤(1988−), 女, 工程师. 研究方向: 水产育种与病害. E-mail: fudanfeng2009@126.com
通信作者: 刘洪涛(1986−), 男, 副研究员. 研究方向: 海洋生物遗传育种. E-mail: xmulht@gmail.com; 何玉贵 (1965−), 男, 高级工程师. 研究方向: 海水虾蟹育种与养殖技术. E-mail: yghgood@163.com

是维持虾健康的主要屏障, 深入研究斑节对虾先天免疫防御对控制对虾疾病爆发和健康养殖具有重要意义。ML 蛋白 (MD-2-related lipid-recognition, MD-2 脂质识别蛋白) 家族是一类广泛分布于动植物和真菌等生物体内的多成员的蛋白家族, MD-2 对于细菌脂多糖 (LPS) 介导的信号传导过程和后续的促炎因子分泌过程至关重要[4]。这些蛋白分子较小, 通常都具有 1 个约 150 个氨基酸残基的 ML 结构域, 其中, 至少含有 2 对保守的半胱氨酸[5]。不同的 ML 蛋白具有不同的氨基酸序列, 从而形成具有差异的中央疏水腔, 可结合不同的脂类配体, 产生脂类识别的特异性, 进一步调节不同的生物学功能[6]。ML 蛋白家族成员众多, 因脂质识别特异性导致其在不同生物体的分布位置及行使功能有很大不同[7]。对脊椎动物中 ML 的研究较多, 发现其在先天免疫和脂类代谢上发挥重要作用, 例如人类 MD-2 参与 Toll 样受体 4 (Toll-like receptors 4, TLR4) 识别细菌脂多糖 LPS, 还可以直接结合 LPS[8], 而人类 C 型尼曼匹克蛋白 2 (Niemann-Pick type C2 protein, NPC2) 具有很强的胆固醇结合能力, 能快速转运胆固醇至膜受体从而调节细胞内的脂质平衡[9]。无脊椎动物中, ML 的功能还未分化和明确, 可能行使更为复杂多样的生物学功用, 研究发现, 黑腹果蝇 (*Drosophila melanogaster*) 的 NPC2a 就同时具备了相当于脊椎动物的 MD-2 和 NPC2 两者的功能[10]。此外, ML 家族还有相当多功能不明确的蛋白有待深入研究。近年来, 对无脊椎动物 ML 基因的研究不断深入, 发现 ML 蛋白参与了宿主和病毒的相互作用[11]。在甲壳动物中, 凡纳滨对虾 (*Litopenaeus vannamei*) 的 LvML 蛋白可能参与抗菌和抗病毒的先天免疫过程[12]; 日本囊对虾 (*Marsupenaeus japonicus*) 的 ML 不仅通过识别 WSSV 的脂质成分参与先天免疫应答, 还可能与幼体发育中的变态及蜕皮有一定的关联[13−14]。本研究基于斑节对虾的 2 个 ML 基因的相似片段, 克隆获得了它们的 cDNA 全序列, 分析了其序列特征, 并利用实时荧光定量 PCR 分析目的基因蛋白在各组织的表达分布情况, 为进一步了解 ML 家族基因在斑节对虾先天免疫中的作用提供资料。

1　材料与方法

1.1　试验材料　供试的斑节对虾 (*Penaeus monodon*) 来自海南省海洋与渔业科学院琼海科研基地高位池(20 m×20 m×2 m)的养殖群体, 平均单只体长为(13.24±0.35)cm, 平均单只的体质量为(19.45±0.61)g, 暂养于水泥池(9.7 m×1.5 m×1.3 m)中。暂养 7 d 后, 随机选取 6 尾身体完整无损伤的健康雌性个体, 取其血细胞、眼柄、心、鳃、胃、肠、肝胰腺和肌肉, 迅速置于 RNAlater (ThermoFisher ,USA) 中 4 ℃ 过夜后, −20 ℃ 长期保存备用。

RNeasy Mini Kit 购自德国 Qiagen。2×*Taq* PCR StarMix with Loading Dye、2×SuperStar Plus Mix with Loading Dye、StarPrep Gel Extraction Kit 和 EZ-Blunt Zero pTOPO II Cloning Kit 等试剂盒均购自北京康润诚业生物科技有限公司 (GenStar)。SMARTer RACE 5′/3′ Kit、PrimerScript™ RT reagent Kit with gDNA Eraser(Perfect Real Time)、SYBR® Premix Dimer Eraser™(Perfect Real Time) 均购自宝生物工程 (大连) 有限公司。

1.2　引物设计　以斑节对虾 2 个 ML 基因相似片段, 利用引物设计软件 Primer Premier 5.0 和分析软件 Oligo 6.0 设计引物进行 cDNA 全序列的克隆[15]。以斑节对虾延伸因子 1α 基因 (*EF-1α*) 作为内参基因, 基于获得的 *PmML1* 与 *PmML2* 的全长 cDNA 序列设计 qRT-PCR 引物。所有引物 (表 1) 由深圳华大基因有限公司合成。

1.3　*PmML1*、*PmML2* cDNA 克隆　使用 RNeasy Mini Kit 提取斑节对虾肝胰腺组织总 RNA, 取 5 μL RNA 进行 1% 琼脂糖凝胶电泳检测完整性, 并使用 ND-1000 超微量紫外分光光度计检测 RNA 质量。使用 StarScript II First-strand cDNA Synthesis Mix With gDNA Remover 反转录制备 cDNA 模板。

分别利用 1 对引物扩增和验证斑节对虾 *PmML1*、*PmML2* 的序列片段, PCR 反应体系 (50 μL): 无菌双蒸水 (DDW) 22 μL, 12×*Taq* PCR StarMix 25 μL, 正反向引物分别 1 μL, 肝胰腺 cDNA 模板 1 μL。PCR 反应程序: 95 ℃ 5 min; 95 ℃ 30 s, 60 ℃ 30 s, 72 ℃ 2 min, 35 个循环; 72 ℃ 5 min。PCR 产物使用 StarPrep Gel Extraction Kit 回收目的条带, 使用 EZ-Blunt Zero pTOPO II Cloning Kit 连

表 1 斑节对虾 *PmML1*、*PmML2* 克隆与表达所用引物

引物名称	引物序列(5′−3′)	用途
*PmML1*F1	AGCCCACAATCATGCGACAG	*PmML1* cDNA扩增
*PmML1*R1	GAGGTTCCTGAACCTGGCTATAC	*PmML1* cDNA扩增
*PmML2*F1	GTGTAGTATACCATGGCTGCCTC	*PmML2* cDNA扩增
*PmML2*R1	TGAGTCTGCGTTGTCCTTACAAG	*PmML2* cDNA扩增
*PmML1*FF1	TTGACTCGCATGTGAGGTGG	*PmML1* 3′-RACE
*PmML1*RR1	GGGTGACTTCCCCTGCTATG	*PmML1* 5′-RACE
*PmML2*FF1	CTCTGTATGGGCACGAGAGC	*PmML2* 3′-RACE
*PmML2*RR1	CAGGATATCGCCGCCTCATC	*PmML2* 5′-RACE
Q*PmML1*F1	TTGACTCGCATGTGAGGTGG	qRT-PCR目的基因扩增
Q*PmML1*R1	GGGTGACTTCCCCTGCTATG	qRT-PCR目的基因扩增
Q*PmML2*F1	CTCTGTATGGGCACGAGAGC	qRT-PCR目的基因扩增
Q*PmML2*R1	CAGGATATCGCCGCCTCATC	qRT-PCR目的基因扩增
*EF-1α*F1	AAGCCAGGTATGGTTGTCAACTTT	qRT-PCR内参基因扩增
*EF-1α*R1	CGTGGTGCATCTCCACAGACT	qRT-PCR内参基因扩增

接，连接产物加入到 50 μL Trans5α Phage Resistant Chemically Competent Cell 中转化，菌液涂于氨苄青霉素 (Amp^+) 抗性平板上，37 ℃ 培养过夜，随机挑取单菌落 M13 引物进行菌液 PCR 鉴定，挑取阳性克隆送测序获得目的片段序列。基于验证的 *PmML1*、*PmML2* 的序列片段设计的 RACE 巢式引物，根据 SMARTer RACE 5′/3′ Kit 试剂盒的方法说明进行 RACE 实验的 cDNA模板制备及 *PmML1*、*PmML2* 基因的 3′末端序列和 5′末端序列的扩增测序。

1.4 生物信息学分析 斑节对虾 ML 基因的 ORF 区域及氨基酸序列使用 NCBI 的 ORF Finder 在线程序 (https://www.ncbi.nlm.nih.gov/orffinder/) 确定[16]。利用 PSORT II Prediction(https://psort.hgc.jp/form2.html) 进行蛋白质亚细胞定位预测[17]。使用 MAFFT 在线软件 (https://mafft.cbrc.jp/alignment/server/) 进行氨基酸序列的多重比对分析[18]。序列的比对、理化性质和系统发生树的构建等参照文献 [19] 的方法，蛋白的二级结构、功能结构域和糖基化位点的预测等参照文献 [20] 的方法。

1.5 *PmML1* 与 *PmML2* mRNA 的组织分布 提取斑节对虾血细胞、眼柄、鳃、心、胃、肠、肝胰腺与肌肉等 8 种组织总 RNA，制备 cDNA 模板及实时荧光定量 PCR 检测各组织表达分布的方法参照文献 [17] 的方法进行，表达数据采用 $2^{-\Delta\Delta Ct}$ 法[21]，用 Excel2019 软件分析和展示检测结果。

2 结果与分析

2.1 斑节对虾 2 个 ML 基因序列及氨基酸序列分析 根据测序结果，拼接获得了斑节对虾 2 个 ML 基因的 cDNA 全序列 (图 1)，分别命名为 *PmML1*(GenBank:MZ291448) 与 *PmML2*(GenBank:MZ291449)。*PmML1* 基因由 307 bp 的 5′非编码区 (5′-Untranslated Region,5′-UTR)、462 bp 的开放阅读框 (Open Reading Frame, ORF) 和 911bp 的 3′-UTR 组成，编码 153 个氨基酸；*PmML2* 基因由 95 bp 的 5′-UTR、471 bp 的 ORF 和 574 bp 的 3′-UTR 组成，编码 156 个氨基酸。PmML1 的氨基酸序列中 Arg(R)、Val(V) 和 Ala(A) 含量较高，相对分子质量为 17 604.30，等电点为 6.11；PmML2 的氨基酸组成中 Leu(L) 和 Ser(S) 含量较高，相对分子质量为 17 719.52，等电点为 6.94，说明 2 个蛋白可能

A

```
1       tgaaacacccgggcggcaggagaggggtcaggacctgggtaaggccagtcgaacagcgct
61      cctccgaggggaaccacctacaccgtcgtcatacggttcgatcgttttctttttaaaaag
121     gcattgcagaagggagttgcaacatccagcttgcagtatcgtagagatcacaggagcatc
181     acattgcctgcctgtctgcctgcctgccgtgcctgtgttgcgctacttgccaagggtcga
241     ttgcctgactgaaggattatttaacgcggtctgttgtgaagttttaaaacaaaattagcc
301     cacaatcATGCGACAGTGGCTGGCACTGGTGGCTCTCCTGGCTATTGCAGTTGGACCATC
1              M  R  Q  W  L  A  L  V  A  L  L  A  I  A  V  G  P  S
361     AACATGTTTTCAGCCGCGGAGGATCATGTTTGCTCGCTGTTCGGCAGAAGAAAAGATGCC
19       T  C  F  Q  P  R  R  I  M  F  A  R  C  S  A  E  E  K  M  P
421     CAGAGCTGTGTTCATGTCATGTGCAAACAGACGTGGAGTTTGTGAAGTCAGGATAGGAAG
39       R  A  V  F  M  S  C  A  N  R  R  G  V  C  E  V  R  I  G  S
481     CACACATAATCTACGAGCCATCTTTATACCAGATTATAATTCCACCAACGTTGACTCGCA
59       T  H  N  L  R  A  I  F  I  P  D  Y  N  S  T  N  V  D  S  H
541     TGTGAGGTGGAACTCCTGGTTTGAGGTCCCCCTGCCAGGCCAAGATCGTGATGCGTGTGA
79       V  R  W  N  S  W  F  E  V  P  L  P  G  Q  D  R  D  A  C  D
601     TGGAGAGATTGACTGCCCCGTCATAGCAGGGGAAGTCACCCGTTTCACTTACGTTCTTGA
99       G  E  I  D  C  P  V  I  A  G  E  V  T  R  F  T  Y  V  L  D
661     TATCCAAGAATTCTGGCCTAGGAACGAGTATCCAGTGATTTGGACACTCACAGACAGAGC
119      I  Q  E  F  W  P  R  N  E  Y  P  V  I  W  T  L  T  D  R  A
721     AACAGATGAAACCCTGGTGTGCTTCAAGTTCAAAATCAATATTGTATAGccaggttcagg
139      T  D  E  T  L  V  C  F  K  F  K  I  N  I  V  *
781     aacctcaaaaaagtctaagggatccggtatccagatggaagtactgtgaaggcccagggt
841     aataactttttaattatcaacaagactatgttaagtatatcatagaatgctcagtatca
901     agatatcacccttcccatatagatgtagatgcccactggagttcatatttttgaagactc
961     agctggtagttgcataacacaaaatgcctatgttaacagtgggaaggaatgacaagtcgc
1 021   aagattccacaagcctctcaaggttatcacagctttatacatgttagaggcctccaagg
1 081   gtctgccataatgtatcaaacagacaatagccattacctgaagaaatgtgtttttagtga
1 141   tggtgtaatccgaaagctaagtaagaatgtagtgtcactatcaaatagagccaagtaaat
1 201   taaccgttgattaagatttagggaaggcctactacctaagcttcagtctcattgttacac
1 261   aagctatgtaaacactccaaggagcatagcacatttccttattatatcaagagctgtatt
1 321   cagatttgttgtaaaatcaccattaaataaatgctcgagcactctcttctctacgttatg
1 381   tgtgttaggatgagagctgtcagtctcactgaaagggaccaggtcctgccactgcaaagt
1 441   accttctgcagaaatgtacctcatctgacgaccctccaagccccgcccaactcctcaaaa
1 501   aagaaaaaaaaaacatgtaaaagaaaaaacaaaatattctatatctttcatgatcacttt
1 561   ttcactttttagatttttcactcttggttgtataaagaaaagaattataattttgtca
1 621   aggaatcttttaataaactaaaagaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

B

```
1       gtcgttggggagggcgaaggcgaggcagtgtgagctttgccgtcgtcaggactcgaagtg
61      tgattagcccgcgtgcgtgtagtgtgtagtataccATGGCTGCCTCTCCCTCGAGCTGTC
1                                           M  A  A  S  P  S  S  C
121     TCTTCGTCGCCCTCGCCCTCTGCCTCGTCGGCTCGTCGCGGGCGGTGATGCACGAAATCC
9       L  F  V  A  L  A  L  C  L  V  G  S  S  R  A  V  M  H  E  I
181     CCGTCAGGAGCTGTGATGGAGCAACAACGCCAAGCAGCCTTAAGATGGAATGCTCGCAGT
29      P  V  R  S  C  D  G  A  T  T  P  S  S  L  K  M  E  C  S  Q
241     TCATTCGTGGGACTTGCATACTCGAGAAAGGGCAAACGTATAACCTTACAGCTGAGTTCA
49      F  I  R  G  T  C  I  L  E  K  G  Q  T  Y  N  L  T  A  E  F
301     CACCAGACCGAAACCTCCGAGAGGTGGAGAGCCACGCAGCCTGGAAGACGTGGGTCGAGA
69      T  P  D  R  N  L  R  E  V  E  S  H  A  A  W  K  T  W  V  E
361     TGCCTCTGTATGGGCACGAGAGCCAAGTCTGCAATGGCATGTACCTCACCTGCCCTCTGC
89      M  P  L  Y  G  H  E  S  Q  V  C  N  G  M  Y  L  T  C  P  L
421     GCGCGGGGCAAACTACGAGCTTCTCATACCCATTCCATGTTCATGAGTTCCTGATGAGGC
109     R  A  G  Q  T  T  S  F  S  Y  P  F  H  V  H  E  F  L  M  R
481     GGCGATATCCTGTCATCTGGCGACTGAGGGACGTGGACAGCGAAGAGACGACGCTGTGTT
129     R  R  Y  P  V  I  W  R  L  R  D  V  D  S  E  E  T  T  L  C
541     TCGTTTTCGTTATTAAAATCTTGTAAggacaacgcagactcacacttacacacgcgcaca
149     F  V  F  V  I  K  I  L  *
601     cccactcagcagaagaagctaaggaatgtgaccaaaaaacaaaacaaaaccggcgtggag
661     agtagagtaatttcgaaaatgtgttaagtcggagaagaagacagaaaagagtggaagata
721     aggccagtctacgttttcgaataaaaagaaaaagaacaagaacaacaaagagaaaacgcc
781     agtctgcttctttttgtcgagtcatttcgaagtgtgttaaacatgattacaggaacaagg
841     aattaagaaaatgaagagatcaataaaaatatgaagataaataccggaatatgaccgact
901     agtctactacataaaatttaagcattgaagagatggcgataatattgtgaaaacgagaag
961     aaaagtgtttctcaaaacaaatgaagacagagaagcgaatgtacaaatgcacatcgaaca
1 021   cattatttattaaccagaccggtgaagaatttaccgccaaagcacggtatacgcttgaaa
1 081   ttaaacttgtatttaataaaactaaaagataaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

图 1 斑节对虾 *PmML1*(A)、*PmML2*(B)基因序列及推导的氨基酸序列

左侧数字分别是核苷酸 (上) 及氨基酸 (下) 的编号; 信号肽用加粗下划线标示; 半胱氨酸位点用方框加粗标示。

是差异较大的小分子蛋白。PmML1 的疏水平均值 (GRAVY) 为 -0.035, 为亲水性蛋白; 而 PmML2 的疏水平均值 (GRAVY) 为 0.024, 属疏水性蛋白。2 个蛋白的疏水性差异暗示其可能在不同的部位发挥生物学功能。此外, PmML1 在 71 位发现 1 个“NSTN”基序的 N–糖基化位点; PmML2 在 38 和 113 位各有 1 个 O–糖基化位点, 在 63 位有 1 个“NLTA”基序的 N–糖基化位点。

预测显示, PmML1 蛋白 N 端存在 1 个具有 20 个氨基酸残基的信号肽; 而 PmML2 蛋白 N 端存在 1 个具有 23 个氨基酸残基的信号肽 (图 1)。蛋白亚细胞定位预测结果显示, PmML1 蛋白定位于细胞外 (包括细胞壁)、细胞质、线粒体、液泡和高尔基体的比例分别为 33.3%、22.2%、22.2%、11.1% 和 11.1%; 而 PmML2 蛋白定位于细胞外 (包括细胞壁)、内质网和液泡的比例分别为

66.7%、22.2% 和 11.1%。2 个基因的定位差异表明他们可能在细胞内外执行不同的生物学功能。预测 PmML1 的半胱氨酸残基可形成 3 对二硫键，位点依次为 31 ~ 145, 45 ~ 103 和 52 ~ 97; PmML2 半胱氨酸残基可形成 3 对二硫键，位点依次为 33 ~ 99, 46 ~ 54 和 106 ~ 148。这些二硫键除了用于维持 PmML1 和 PmML2 分子内构象以单体发挥作用外，可能还会形成二聚体或寡聚体、多聚体来执行某些生物学功能。蛋白二级结构分析显示，PmML1 蛋白中 β–片层结构占比 31.37%; PmML2 蛋白中 β–片层结构占比为 31.41% (图 2)。蛋白结构域分析显示，PmML1 蛋白第 52 ~ 150 号氨基酸残基之间有 1 个典型的 ML super family 结构域 (cl00274); PmML2 在第 54 ~ 153 号氨基酸残基之间具有 1 个 ML 结构域 (smart00737)(图 2)。

2.2　PmML1、PmML2 多重比对及系统进化树分析

氨基酸序列比对分析发现，PmML1 和 PmML2 之间的相似性仅为 42.37%，除保守区外两者的差异氨基酸会影响中央疏水腔结构，进而识别特异性脂类，使得 2 个基因可能在斑节对虾生理活动中行使不同的功能。PmML1 与 PmML2 氨基酸序列与凡纳滨对虾、日本囊对虾等近缘动物不同 ML 多重比对表明其 ML 蛋白结构域及半胱氨酸位置相对保守 (图 3)。PmML1 与日本囊对虾的 ML 蛋白 MjML4 相似性最高(91.5%); PmML2 与凡纳滨对虾 ML 蛋白 (XP027236046.1) 相似性最高(82.8%)。系统进化分析结果显示，*PmML1* 和 *PmML2* 分别存在于 2 个不同的进化分支，暗示 2 个基因在进化过程中相对独立，证明本研究中鉴定到的是 2 个不同的基因 (图 4)。

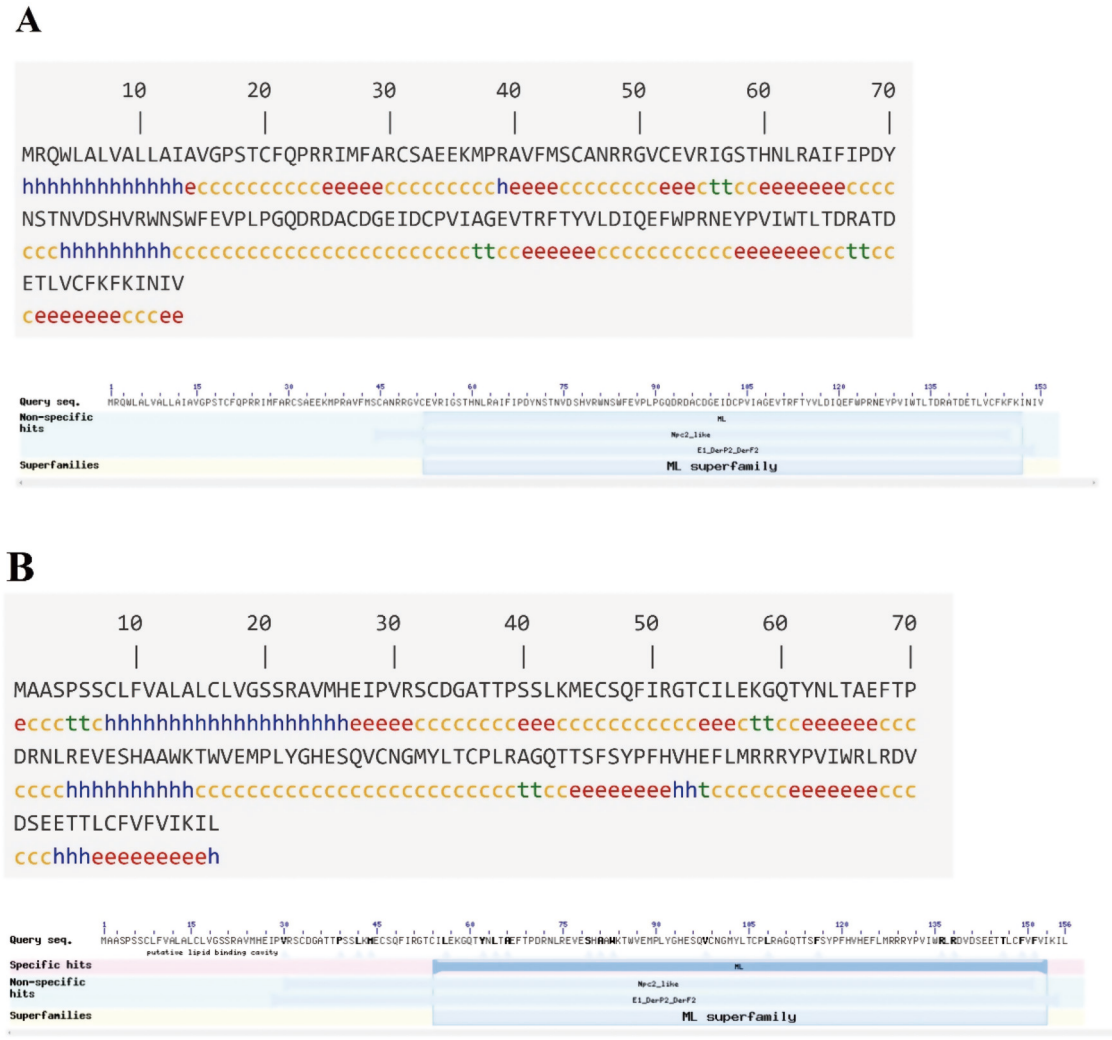

图 2　PmML1(A)、PmML2(B)的 2 级结构(上)与蛋白质结构域(下)预测

h: α–螺旋; t: β–转角; e: 延伸链; c: 无规则卷曲。

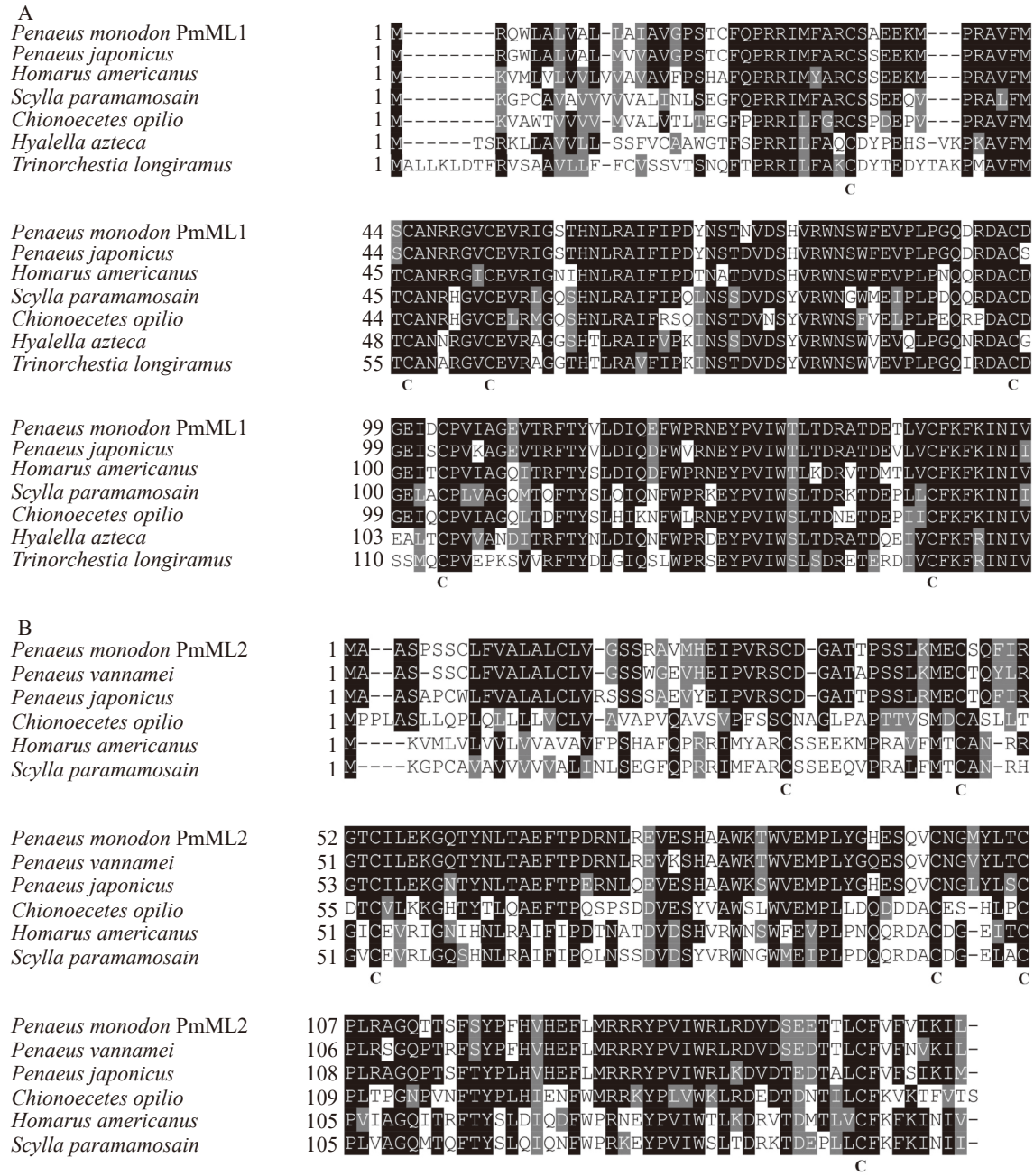

图 3 PmML1 与 PmML2 氨基酸序列的多重对比

A 为 PmML1 氨基酸序列多重对比,参与对比的序列信息为日本囊对虾 (QDX01885.1);美洲螯龙虾 (KAG7167740.1); 拟穴青蟹 (QBJ05649); 雪蟹 (KAG0727136); 端足虫 (XP018024550.1); 沙蚕 (KAF2364063.1)。B 为 PmML2 氨基酸序列多重对比,参与对比的序列信息为凡纳滨对虾 (XP027236046.1); 日本囊对虾 (QDX01882.1); 雪蟹 (KAG0727135.1); 美洲螯龙虾 (KAG7157740.1); 拟穴青蟹 (QBJ05649.1)。

2.3 *PmML1*、*PmML2* 的 mRNA 组织表达 qRT-PCR 检测 *PmML1* 和 *PmML2* 在斑节对虾肝胰腺、鳃丝、胃、肠、心、肌肉、眼柄和血细胞等 8 种组织中的表达。从图 5 可知,*PmML1* 在 8 种组织中均有表达,以肝胰腺中表达量为对照,其在眼柄中表达量最高,其次为肠、胃、心、鳃、肌肉和肝胰腺,而在血细胞的表达水平最低。*PmML2* 在 8 种组织中也均有表达,以肝胰腺中表达量为对照,其在眼柄中表达量最高,其次为肠、鳃、胃、血细胞、心、肝胰腺和肌肉,在肌肉中的表达水平最低。*PmML2* 在眼柄的高表达,提示其可能参与了该部位重要器官 X 器官窦腺复合体调控的多种重要生理活动。

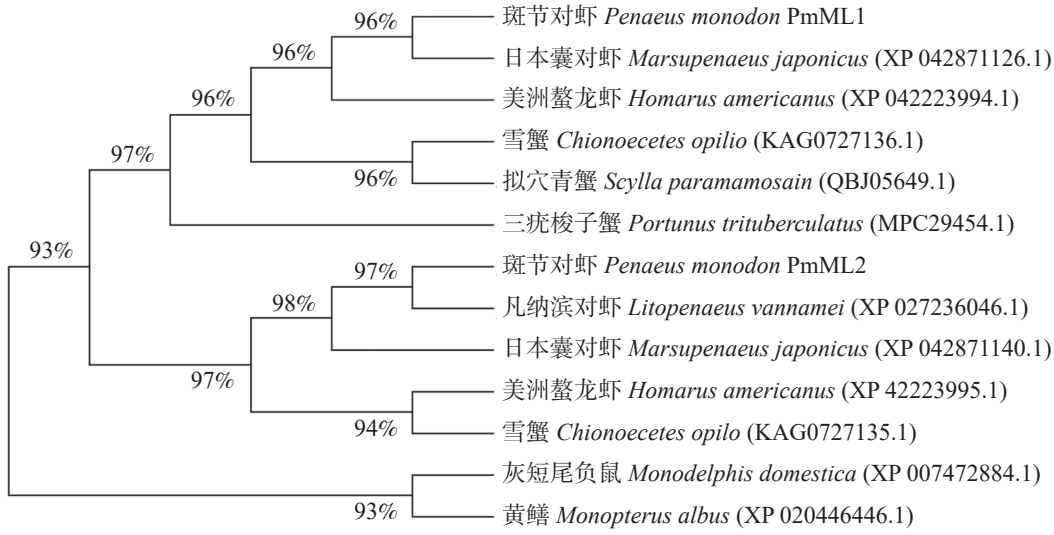

图 4　基于 PmML1 与 PmML2 氨基酸序列构建的 ML 系统进化树

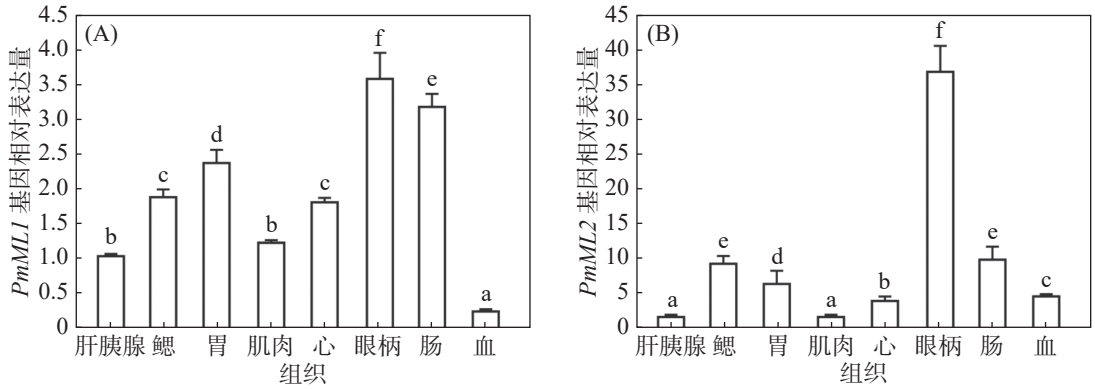

图 5　*PmML1*(A)和 *PmML2*(B) mRNA 的组织表达分布

3　讨　论

ML(MD-2-related Lipid-recognition) 基因家族是由一类具有 ML 单结构域的蛋白质所组成，包括 MD-1、MD-2、GM2 激活蛋白、Niemann-Pick 型 C2 (NPC2) 蛋白和多种广泛分布于动植物及真菌的未知功能蛋白[6]。到目前为止，已有多个物种的 ML 家族基因被鉴定。例如，在果蝇(*Drosophila melanogaster*)基因组编码 8 个 ML 家族基因，在冈比亚按蚊 (*Anopheles gambiae*) 基因组编码 13 个 ML 家族基因，在埃及伊蚊 (*Aedes aegypti*) 基因组中编码 24 个 ML 家族基因[22]。本试验利用斑节对虾 2 个 ML 相似片段，成功获得斑节对虾 2 个 ML 基因 (*PmML1*、*PmML2*) 的 cDNA 全序列，在斑节对虾中尚属首次。PmML1由 153 个氨基酸组成，由 1 个 20 个氨基酸残基的信号肽和 3 对半胱氨酸残基的 ML 结构域组成；PmML2 由 156 个氨基酸组成，由 1 个 23 个氨基酸残基的信号肽和 3 对半胱氨酸残基的 ML 结构域组成，2 个基因都符合 ML 蛋白多由 140 ~ 160 个氨基酸残基组成的小分子量以及都有 1 个 N 端信号肽序列和至少 2 对保守的半胱氨酸残基的结构特征[7]。因此，本研究获得的 2 个基因都是 ML 基因家族的新成员。

ML 蛋白的 ML 结构域由多个 β 片层结构组成，通过形成 2 个大的 β 折叠构成 1 个 Ig-like 结构的中央疏水腔[23]。该中央疏水腔通过包被脂类配体的疏水部分促进其溶解性并被膜上受体和酶所识别[7]。斑节对虾 PmML1 和 PmML2 的 2 级结构预测显示，其 β 片层结构分别占比 31.37% 和 31.41%；同源性和多重对比分析显示，在配体结合部位两者存在较高的序列变化，特别是疏水性氨基酸残基的差异。说明这两个蛋白形成的中央疏水腔的结构存在较大差异，能识别不同的特异性脂类，在斑节对虾生理活动中可能行使不同的功能。这一结果在许多研究中也得到证实，例如

MD-2 结合 TLR4 的复合体除识别 LPS 外，还能识别紫杉醇、人衣原体 HSP60、纤维粘连蛋白的功能结构域等物质[24]；而 NPC2 蛋白晶体结构分析则显示其内部疏水腔能结合 1 个胆固醇分子[25]。

ML 蛋白的糖基化位点和半胱氨酸残基对 ML 蛋白正确行使功能发挥至关重要的作用。本研究中，PmML1 在 71 位点有 1 个“NSTN”基序的 N-糖基化位点；PmML2 在 38 和 113 位点各有 1 个 O–糖基化位点，在 63 位点有 1 个“NLTA”基序的 N–糖基化位点。相关研究已证实，在糖基化位点发生突变后，MD-2 仍可表达于细胞表面，并保持与 TLR4 结合，但辅助 TLR4 介导 NF-kb 的活化能力减弱，不能促进 LPS 诱导 IL-8 分泌，也不能活化 JNK 途径[26－28]。在 NPC2 蛋白还发现不同的糖基化形式将决定其能否正确靶向溶酶体从而调节胆固醇的代谢水平[25,29]。本研究中，PmML1 氨基酸序列中有 7 个半胱氨酸残基，除 1 个位于信号肽序列外，剩余 6 个可形成 3 对二硫键，PmML2 氨基酸序列中有 8 个半胱氨酸残基，但其中有 2 个位于信号肽序列，剩余 6 个可形成 3 对二硫键。这些半胱氨酸残基及其形成的二硫键，将对 PmML1 和 PmML2 蛋白的分子构象及生物学功能发挥具有重要影响。半胱氨酸位点突变实验证实，MD-2 分子的寡聚化不依赖任何单一的半胱氨酸残基，但 95 位和 105 位的半胱氨酸残基突变将导致 MD-2 丧失与 TLR4 结合并赋予 TLR4 对 LPS 反应的能力[28,30]。二硫键形成的共价结构，能够为许多蛋白质提供基本的稳定性并符合明确定义的几何构象[31]。MD-2 蛋白半胱氨酸残基构成的二硫键可以形成分子内二硫键，也可以由游离的二硫键在分子间连接形成稳定的二聚体或寡聚体，多聚体，但单体 MD-2 比多聚体 MD-2 更易结合 TLR4，能够有效促进 TLR4 对 LPS 的识别[32]。因此，笔者推测，MD-2 的二聚体、寡聚体、多聚体可能是 MD-2 单体避免被代谢掉的一种储存形式，但具体的分子机制有待进一步的研究。

不同的 ML 家族蛋白在不同物种的分布具有很大的差异。本试验利用实时荧光定量 PCR 技术检测发现，*PmML1* 和 *PmML2* 基因在斑节对虾鳃丝、肝胰腺等 8 种组织中的相对表达量，发现 *PmML1* 和 *PmML2* 在 8 种组织中均有表达，*PmML1* 在眼柄表达量最高，而 *PmML2* 在眼柄中表达量最高，且与其他组织有明显差异。总的来说，*PmML1* 和 *PmML2* 是在机体内普遍存在的，但有组织偏好性，在眼柄中表达量最高，而在免疫组织中表达量相对较低。这在日本囊对虾和凡纳滨对虾等近缘物种的研究中也得到证实。例如，在日本囊对虾的 ML 家族基因中，*MjML1*、*MjML2* 和 *MjML4* 在各个检测组织中均有分布，而 *MjML3* 和 *MjML5* 则仅在肝胰腺中表达[14]；在凡纳滨对虾中，LvML 则仅在肝胰腺组织中表达[12]。位于对虾眼柄的主要功能器官为 X 器官窦腺复合体 (X-organ sinus gland complex, XO-SG complex)，是对虾神经内分泌调节系统之一，在结构和功能上类似脊椎动物的下丘脑-神经垂体系统，它参与调控甲壳动物血糖水平、生长蜕皮和性腺发育等多种重要的生理活动[33]。斑节对虾中 *PmML1* 和 *PmML2* 在眼柄的高表达，暗示这 2 个基因可能主要在神经内分泌系统发挥调控作用。*PmML1* 和 *PmML2* 在多个免疫组织表达量相对较低，暗示 PmML1、PmML2 2 种蛋白参与的生物学过程可能不是和免疫反应直接相关的。此外，徐伊桦等[13]对日本囊对虾 MjML 的研究发现，其 ML 蛋白可能参与了对虾的变态及蜕皮等重要生理过程。本研究的 2 个 ML 蛋白或许具有类似的生物学功能。综上所述，可以认为 *PmML1* 和 *PmML2* 是 ML 基因家族的新成员，可能在斑节对虾先天免疫及神经内分泌调控等方面发挥重要作用。

参考文献：

[1] HOLTHUIS L B. FAO Species Catalogue: Shrimps and prawns of the world. An annotated catalogue of the species of interest to fisheries[M]. Rome: FAO Fisheries Synopsis, 1980.

[2] 俞开康, 战文斌, 周丽, 等. 我国沿海养殖对虾的疾病及研究现状[J]. 中山大学学报 (自然科学版), 2000, 39(S1): 1－5.

[3] ALEXANDER C, RIETSCHEL E T. Bacterial lipopolysaccharides and innate immunity [J]. Journal of Endotoxin Research, 2001, 7(3): 167－202.

[4] LIAO J X, YIN Z X, HUANG X D, et al. Cloning and characterization of a shrimp ML superfamily protein [J]. Fish Shellfish Immunol, 2011, 30(2): 713－719.

[5] 梅雄娥, 李致宏, 于滨, 等. 家蚕 ML 家族基因的鉴定及病原诱导表达特征[J]. 蚕业科学, 2019, 45(1): 32－39.

[6] INOHARA N, NUÑEZ G. ML—a conserved domain involved in innate immunity and lipid metabolism [J]. Trends in Biochemical Sciences, 2002, 27(5): 219－221.

[7] 徐伊桦. 日本囊对虾 MjML 基因的克隆与表达研究[D]. 厦门: 厦门大学, 2017.
[8] MANČEK M, PRISTOVŠEK P, JERALA R. Identification of LPS-binding peptide fragment of MD-2, a toll-receptor accessory protein [J]. Biochemical and Biophysical Research Communications, 2002, 292(4): 880 − 885.
[9] CHERUKU S R, XU Z, DUTIA R, et al. Mechanism of cholesterol transfer from the Niemann-Pick type C2 protein to model membranes supports a role in lysosomal cholesterol transport [J]. Journal of Biological Chemistry, 2006, 281(42): 31594 − 31604.
[10] SHI X Z, ZHONG X, YU X Q. *Drosophila melanogaster* NPC2 proteins bind bacterial cell wall components and may function in immune signal pathways [J]. Insect biochemistry and molecular biology, 2012, 42(8): 545 − 556.
[11] JUPATANAKUL N, SIM S, DIMOPOULOS G. *Aedes aegypti* ML and Niemann-Pick type C family members are agonists of dengue virus infection [J]. Developmental & Comparative Immunology, 2014, 43(1): 1 − 9.
[12] LIAO J X, YIN Z X, HUANG X D, et al. Cloning and characterization of a shrimp ML superfamily protein [J]. Fish & shellfish immunology, 2011, 30(2): 713 − 719.
[13] 徐伊桦, 苏永全, 毛勇, 等. 日本囊对虾 ML 基因的克隆与表达分析[J]. 热带生物学报, 2015, 6(4): 365 − 373.
[14] GAO J, WANG J X, WANG X W. MD-2 homologue recognizes the white spot syndrome virus lipid component and induces antiviral molecule expression in shrimp [J]. The Journal of Immunology, 2019, 203(5): 1131 − 1141.
[15] 张新宇, 高燕宁. PCR 引物设计及软件使用技巧[J]. 生物信息学, 2004(4): 15 − 18.
[16] ROMBEL I T, SYKES K F, RAYNER S, et al. ORF-FINDER: a vector for high-throughput gene identification [J]. Gene, 2002, 282(1/2): 33 − 41.
[17] HORTON P, NAKAI K. Better prediction of protein cellular localization sites with the k nearest neighbors classifier[C]// TERRY G, PETER D K, KEVIN K, et al. Proceedings of the 5th International Conference on Intelligent Systems for Molecular Biology. Greece: AAAI Press, 1997: 147-152.
[18] KATOH K, ROZEWICKI J, YAMADA K D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization [J]. Briefings in Bioinformatics, 2019, 20(4): 1160 − 1166.
[19] 刘洪涛, 王军, 毛勇, 等. 日本囊对虾 Na-K-2Cl 共同转运蛋白基因的克隆与组织表达分析[J]. 厦门大学学报 (自然科学版), 2016, 55(1): 46 − 54.
[20] 何金娇, 刘洋洋, 毛雪飞, 等. SARS-CoV-2 S 蛋白结构与功能的生物信息学分析[J]. 基因组学与应用生物学, 2021(S1): 2346 − 2355.
[21] LIVAK K J, SCHMITTGEN T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_t}$ method [J]. Methods, 2001, 25(4): 402 − 408.
[22] WATERHOUSE R M, KRIVENTSEVA E V, MEISTER S, et al. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes [J]. Science, 2007, 316(5832): 1738 − 1743.
[23] PARK B S, SONG D H, KIM H M, et al. The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex [J]. Nature, 2009, 458(7242): 1191 − 1195.
[24] MIYAKE K. Innate recognition of lipopolysaccharide by CD14 and toll-like receptor 4-MD-2: unique roles for MD-2 [J]. International Immunopharmacology, 2003, 3(1): 119 − 128.
[25] ROSENBAUM A I, MAXFIELD F R. Niemann-Pick type C disease: molecular mechanisms and potential therapeutic approaches [J]. Journal of Neurochemistry, 2011, 116(5): 789 − 795.
[26] DA SILVA CORREIA J, ULEVITCH R J. MD-2 and TLR4 N-linked glycosylations are important for a functional lipopolysaccharide receptor [J]. Journal of Biological Chemistry, 2002, 277(3): 1845 − 1854.
[27] OHNISHI T, MUROI M, TANAMOTO K. MD-2 is necessary for the toll-like receptor 4 protein to undergo glycosylation essential for its translocation to the cell surface [J]. Clinical and Vaccine Immunology, 2003, 10(3): 405 − 410.
[28] SCHROMM A B, LIEN E, HENNEKE P, et al. Molecular genetic analysis of an endotoxin nonresponder mutant cell line: a point mutation in a conserved region of MD-2 abolishes endotoxin-induced signaling [J]. The Journal of Experimental Medicine, 2001, 194(1): 79 − 88.
[29] CHIKH K, VEY S, SIMONOT C, et al. Niemann-Pick type C disease: importance of N-glycosylation sites for function and cellular location of the NPC2 protein [J]. Molecular Genetics and Metabolism, 2004, 83(3): 220 − 230.
[30] MULLEN G E D, KENNEDY M N, VISINTIN A, et al. The role of disulfide bonds in the assembly and function of MD-2 [J]. Proceedings of the National Academy of Sciences, 2003, 100(7): 3919 − 3924.
[31] ZHOU H, WANG W, LUO Y. Contributions of disulfide bonds in a nested pattern to the structure, stability, and biological functions of endostatin [J]. Journal of Biological Chemistry, 2005, 280(12): 11303 − 11312.
[32] VISINTIN A, MAZZONI A, SPITZER J A, et al. Secreted MD-2 is a large polymeric protein that efficiently confers lipopolysaccharide sensitivity to Toll-like receptor 4 [J]. Proceedings of the National Academy of Sciences, 2001, 98(21): 12156 − 12161.
[33] 郭明申, 康现江. 中国对虾眼柄神经分泌结构的超微结构研究[J]. 台湾海峡, 2002, 21(3): 279 − 283.

Identification and expression analysis of *PmML1* and *PmML2* in *Penaeus monodon*

FU Danfeng[1], LIU Hongtao[2], YANG Mingqiu[2], HE Yugui[2]
（1. Hainan Provincial Fishery Technical Extension Station, Haikou, Hainan 570226; 2. Hainan Provincial Key Laboratory of Tropical Maricultural Technologies, Hainan Academy of Ocean and Fisheries Sciences, Haikou, Hainan 571126, China）

Abstract: MD-2-related Lipid-recognition proteins (ML) are a class of proteins that play an important role in the innate immunity and lipid recognition and metabolism. Two ML family genes *PmML1* and *PmML2* were cloned from shrimps (*Penaeus monodon*) by using PCR and RACE methods for bioinformatics analysis, and their expression profiles in various tissues of the shrimps were analyzed by using the quantitative real-time PCR (qRT-PCR). The results showed that the open reading frame (ORF) length of *PmML1* was 462 bp, encoding 153 amino acids and that the ORF length of PmML2 was 471 bp, encoding 156 amino acids. There is a typical ML family protein structural domain in these two genes. Multiple alignments revealed that the ML domain and the cysteine position in the *PmML1* and *PmML2* were relatively conserved. Alignment of the amino acid sequences showed that the similarity between PmML1 and PmML2 was only 42.37%. Moreover, the similarity between PmML1 and the MjML4 of *Marsupenaeus japonicus* was the highest (91.5%), and the similarity between PmML2 and the ML protein of *Litopenaeus vannamei* was the highest (88.46%). Phylogenetic analysis showed that *PmML1* and *PmML2* were found in two different branches, indicating that the evolution of two genes is relatively independent, which proved that two genes were identified different. Furthermore, the qRT-PCR results showed that *PmML1* expressed highest in the eye stalk, followed by in the intestine, stomach, heart, gill, muscle and hepatopancreas, and the lowest expression level was in the blood cells. *PmML2* expressed highest in the eye stalk, followed by in the intestine, gill, stomach, blood cells, heart, hepatopancreas and muscle, and the lowest expression level was in the muscle. As the eyestalk is an important neuroendocrine regulatory organ, it is speculated that *PmML1* and *PmML2* may play a role in the neuroendocrine regulation in *P. monodon.* This result might provide basic data for further study of ML family genes and their biological functions in shrimps.

Keywords: *Penaeus monodon*; ML (MD-2-related Lipid-recognition); gene cloning; tissue expression

(责任编委:罗启香　责任编辑:潘学峰)