合浦珠母贝组织蛋白酶L2基因的特征与组织表达分析

麻建军[1,2],张殿昌[2],崔淑歌[1,2],姜晶晶[2,3],蒲含林[1],江世贵[2]

(1. 暨南大学 生命科学技术学院,广东 广州 510632; 2. 中国水产科学研究院 南海水产研究所,广东 广州 510300; 3. 上海海洋大学 水产与生命学院,上海 200090)

摘要: 为了研究合浦珠母贝(*Pincatada fucata*)的先天免疫调控机制,利用cDNA文库筛选和重测序技术克隆了1个合浦珠母贝组织蛋白酶L的全长cDNA序列(命名为poCL2)。poCL2 cDNA全长1 094 bp,5′-非编码区(Untranslated region, UTR)长21 bp,3′-UTR长80 bp,开放阅读框(Open reading frame, ORF)为993 bp,编码330个氨基酸组成的多肽链,分子量为37.1 kD,理论等电点为6.9。poCL2蛋白由信号肽(Met^{1}-Ala^{16})、前体域(Arg^{17}-Asp^{113})和成熟域(Leu^{114}-Val^{330})三部分组成,存在一个潜在的糖基化位点(Asn^{97}),6个底物结合位点(Leu^{182}, Met^{183}, Ala^{249}, Leu^{275}, Gly^{278}和Ser^{324}),1个由半胱氨酸(Cys^{138})、组氨酸(His^{277})和天冬酰胺(Asn^{297})残基构成的催化位点和2个组织蛋白酶L签名序列($E^{42}X_3RX_3WX_2NX_3IX_3N^{61}$和$G^{74}X_1NX_1YX_1D^{80}$)。同源性分析表明, poCL2氨基酸序列与其他物种高度保守,相似性在61.5% ~ 73.9%之间,同源性在44.4% ~ 59.8%之间。进化分析显示, poCL2与其他无脊椎动物的组织蛋白酶L聚为一支,和淡水螺(*Radix peregra*)亲缘关系最近。组织表达分析表明poCL2 mRNA在消化腺、外套膜、性腺、闭壳肌、肠、血淋巴和鳃组织中均表达,外套膜表达量最高。应激实验表明,经溶藻弧菌(*Vibrio alginolyticus*)或脂多糖(Lipopolysaccharide, LPS)刺激后, poCL2 mRNA在消化腺中的表达量显著上调,说明poCL2参与了合浦珠母贝的先天免疫应答反应,暗示其在合浦珠母贝先天免疫调控中发挥重要作用。[中国水产科学,2010,17(4): 701-712]

关键词: 合浦珠母贝; 组织蛋白酶L; 先天免疫

中图分类号: Q959; S91 **文献标识码**: A **文章编号**: 1005-8737-(2010)04-0701-12

合浦珠母贝(*Pinctada fucata*)又称马氏珠母贝,隶属软体动物门(Mollusca),瓣鳃纲(Lamellibranchia),异柱目(Anisomyaria),珍珠贝科(Pteriidae),是中国海水珍珠培育的主要种类之一,广泛分布于中国的广东、广西和海南等地,具有较高的经济价值,举世闻名的"南珠"即由这种珍珠贝所产。近年来,种质退化、病害增多以及养殖环境恶化等因素使合浦珠母贝在人工养殖过程中出现大量死亡[1-3]。与其他无脊椎动物一样,合浦珠母贝只有非特异性免疫系统,但是目前对其免疫防御机制缺乏了解,为了解决合浦珠母贝养殖中严重的病害问题,有必要加强合浦珠母贝自身免疫防御机制的研究,克隆鉴定免疫相关基因,研究其免疫调控机制,以期为合浦珠母贝病害免疫防治研究提供新的思路。

组织蛋白酶L是半胱氨酸蛋白酶中木瓜蛋白酶C1家族的主要成员,广泛存在于各种生物有机体中,在溶酶体内的蛋白水解过程中起重要作用[4]。组织蛋白酶L都是以无活性的前体酶原(Preprocathepsin)形式合成,前体酶原由信号肽(Signal peptide)、前体域(Prodomain)和含有催化活性中心的成熟域(Mature domain)组成,在溶酶体的酸性条件下自动水解,或在其他蛋白酶作用下水解,去

收稿日期: 2010-03-28; **修订日期**: 2010-04-30.

基金项目: 国家"863"计划项目(2009AA10Z106); 农业部公益性行业科研专项项目(200903028); 广东省科技兴海项目(A200701002); 中央级公益性科学院基础科研业务费专项资金项目(2009TS23,2010YD03).

作者简介: 麻建军(1983-),男,硕士研究生,研究方向: 基因工程制药学. E-mail: majianjun2010@163.com

通讯作者: 江世贵,研究员. Tel: 020-89108338; E-mail: jiangsg@21cn.com

掉前体域，形成有活性的成熟的组织蛋白酶[5]。组织蛋白酶L作为一种溶酶体蛋白在生物体内参与多种重要生理活动，如抗原呈递[6]、寄生虫感染[7]、炎症发生[8]、MHC（Major histocompatibility complex，组织相容性复合体）降解和T细胞发育[9]。最近，Liu和Sukhova[10]发现，动脉粥样硬化患者血清中组织蛋白酶L存在高表达，推测其与动脉粥样硬化的发病过程相关联。在刀额新对虾（*Metapenaeus ensis*）中，Hu和Leung[11-12]发现组织蛋白酶L在参与消化酶分泌的B细胞中含量非常高，可能参与了食物消化过程。此外，其他研究者发现组织蛋白酶L还参与斑马鱼（*Brachydanio rerio*）、文昌鱼（*Branchiostoma lanceolatum*）卵黄形成和胚胎发育过程[13-14]。

为了深入了解组织蛋白酶L在合浦珠母贝先天免疫调控中的作用，本实验室利用cDNA文库筛选和重测序技术，克隆到2个合浦珠母贝组织蛋白酶L全长cDNA序列，分别命名为poCL1和poCL2，poCL1研究结果已另文投稿。本研究利用荧光定量RT-PCR（Real-time quantitative RT-PCR）方法，研究了poCL2 mRNA的组织表达模式，以及在溶藻弧菌（*V. alginolyticus*）或LPS刺激后poCL2 mRNA的表达调控规律，结果表明，2种刺激条件下poCL2 mRNA在消化腺中的表达量都显著上调，说明poCL2可能参与了合浦珠母贝的先天免疫调控。

1 材料与方法

1.1 实验材料

实验用合浦珠母贝取自中国水产科学研究院南海水产研究所陵水珍珠养殖试验场，贝壳长5.3～6.5 cm，体质量19.6～27.8 g，于25～27 ℃充气海水中暂养1周后进行实验。每天用扁藻（*Tetraselmis suecica*）和金藻（*Isochrysis galbana*）投喂2次。设置正常组、对照组、溶藻弧菌刺激组和LPS刺激组，每个实验组设置3个重复。将未经过任何处理的活力良好的合浦珠母贝作为正常组；将闭壳肌内注射100 μL PBS（pH 6.8）的贝作为对照组；取LPS溶解于PBS（pH 6.8）中，使LPS终质量浓度达到1.0×10^{-3} g/L，闭壳肌内注射100 μL LPS溶液的贝作为LPS刺激组；取过夜培养的溶藻弧菌悬浮于PBS，使$OD_{600}=0.4$，闭壳肌内注射100 μL弧菌悬液的贝作为溶藻弧菌刺激组。

1.2 RNA提取及cDNA的合成

取正常组合浦珠母贝的消化腺、外套膜、性腺、闭壳肌、肠、血淋巴和鳃组织，液氮速冻，−80 ℃保存备用；分别在注射后0 h、2 h、4 h、8 h、12 h、24 h、48 h、72 h取对照组、溶藻弧菌刺激组和LPS刺激组合浦珠母贝的消化腺组织，液氮速冻，−80 ℃保存备用。每5个贝的同种组织等量混合后，于液氮中研磨作为一个样本，按照RNeasy Mini Kit（QIAGEN）使用说明提取总RNA。以RNase-free的DNase I处理后的总RNA为模板，按照PrimeScript™ RT reagent Kit（Perfect Real Time）合成cDNA第一链，反应程序为42 ℃、60 min，70 ℃、15 min，反应结束后，将cDNA反应液用RNase-free H_2O按体积比1∶5稀释，−80 ℃保存以备荧光定量RT-PCR分析使用。

1.3 EST文库构建和poCL2克隆

取溶藻弧菌刺激后的合浦珠母贝软体组织提取总RNA，使用ZAP-cDNA® Synthesis kit（Stratagene）、ZAP-cDNA® Gigapack III Gold Cloning kit（Stratagene）和polyAtract® mRNA isolation system Ⅲ（Promega）等试剂盒构建EST文库。随机挑取单克隆，用引物T3（AATTAACCCTCACTAAAGGG）进行测序，通过Crossmatch程序去载体；通过Phrap程序，进行拼接聚类，得到UniGene数据集，Phrap程序处理后，可以聚到一起的EST序列为Contigs序列，不能聚到一起的EST序列为Singlets序列，两部分序列合起来得到了EST去冗余后的UniGene数据集。分别将获得的Contigs与Singletons在数据库中进行BLASTn和BLASTx分析，鉴定出124个与免疫或应激响应相关的基因，其中1个EST序列与GenBank数据库中其他物种的组织蛋白酶L具有较高的同源性，从文库中挑取该克隆进行重新测序，获得poCL2 cDNA全长序列。

1.4 poCL2序列分析

利用DNATool 6.0软件预测氨基酸序列；信号

肽预测用SingalP 3.0（http://www.cbs.dtu.dk/services/SignalP/）程序[15]；蛋白结构域分析用SMART 4.0（http://smart.embl-heidelberg.de/）程序[16-17]；利用MatGAT软件计算与其他物种组织蛋白酶L序列的相似性和同源性[18]；多序列比对采用Clustal W程序；结合SMART和Blast分析[19]，利用Clustal W将poCL2基因预测的氨基酸序列与人组织蛋白酶L1（P07711）、鼠组织蛋白酶L1（07154）、牛组织蛋白酶S（P25326）和鸡组织蛋白酶L（P09648）的成熟域进行比对和分析，得到poCL2氨基酸序列的前体域和成熟域的切割位点[20]；利用Clustal W程序和MEGA 3.0软件，以邻位相连法（Neighbor-joining）构建系统进化树[21]。

1.5　poCL2 mRNA组织表达分析

根据测序得到的poCL2 cDNA全序列设计一对基因特异性引物poCL2-F（ATTGATGCCAGCCACATGAG）和poCL2-R（AGCAATACCACAGTTGTTACGGCG），用合浦珠母贝β-*actin*基因作为内参照，设计引物β-*actin*-F（GCCGAAAGAGAAATCGTCAG）和β-*actin*-R（TGGCTGGAATAGGGATTCTG），来验证荧光定量RT-PCR反应并校准cDNA模板浓度，以正常组7种组织稀释后的反转录cDNA反应液为模板进行荧光定量RT-PCR扩增。反应体系为20 μL，包含10 μL的2×SYBR Green Real-Time PCR Master Mix（TaKaRa），1 μL模板，0.16 μmol/L引物和8.2 μL的双蒸水，每个样品设置3个重复，以蒸馏水代替模板作为阴性对照。反应参数为95 ℃预变性10 s，然后95 ℃变性10 s，53 ℃退火30 s，72 ℃延伸30 s，共40个循环，最后进行溶解曲线（Melting Curve）分析。采用相对CT法（$2^{-\Delta\Delta CT}$ method）分析poCL2 mRNA的组织表达[22]。

1.6　poCL2 mRNA在溶藻弧菌刺激下的表达特征分析

以对照组、溶藻弧菌刺激组各时间点消化腺cDNA为模板，进行荧光定量RT-PCR扩增。采用相对CT法（$2^{-\Delta\Delta CT}$ method）进行poCL2 mRNA在溶藻弧菌刺激下的表达特征分析。引物、反应体系和反应程序同**1.5**。

1.7　poCL2 mRNA在LPS刺激下的表达特征分析

以对照组、LPS刺激组各时间点消化腺cDNA为模板，进行荧光定量RT-PCR扩增。采用相对CT法（$2^{-\Delta\Delta CT}$ method）进行poCL2 mRNA在LPS刺激下的表达特征分析。引物和反应体系和反应程序同**1.5**。

1.8　统计学分析

运用统计学软件SPSS 10.0进行单因素相关性分析，刺激组与对照组间$P<0.05$则认为差异显著，$P<0.01$则认为差异极显著。

2　结果与分析

2.1　RNA提取

提取的合浦珠母贝总RNA（图1）的A_{260}/A_{280}比值为1.79，琼脂糖凝胶电泳结果显示有28S和18S两条清晰的条带，说明提取的总RNA完整性良好，满足后续实验要求。

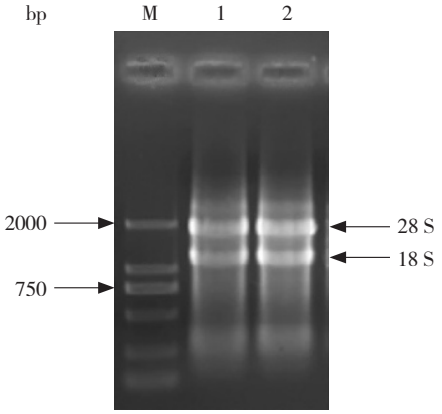

图1　合浦珠母贝总RNA琼脂糖凝胶电泳图
M：DNA分子量标准（DL2000）；1-2：合浦珠母贝总RNA.
Fig. 1　Total RNA of *Pinctada fucata*
M：DNA molecular weight marker（DL2000）；1-2：Total RNA of *Pinctada fucata*.

2.2　EST文库构建和poCL2克隆

测序7 515个克隆，获得了6 741个EST，平均读长456 bp，经聚类分析共得到737个Contigs和2 528个Singletons序列。BLAST分析，发现其中一个420 bp的EST（no. pmpca0_002189），与海胆（*Strongylocentrotus purpuratus*）、蜱（*Dermacentor variabilis*）和短尾负鼠（*Monodelphis domestica*）的组织蛋白酶L基因具有较

高同源性。从文库挑取对应的单克隆重新测通后，去除载体序列得到一个长1 094 bp且3′端包含了polyA的cDNA，经BLAST和SMART分析认定为合浦珠母贝组织蛋白酶L基因全长cDNA序列，命名为poCL2，GenBank注册号GU361539。

2.3 poCL2序列分析

合浦珠母贝组织蛋白酶L2基因cDNA全长1 094 bp，5′-UTR长21 bp，3′-UTR长80 bp，开放阅读框（ORF）为993 bp，编码330个氨基酸组成的多肽链，分子量为37.1 kD，理论等电点为6.9。poCL2蛋白由信号肽（Met^{1}-Ala^{16}）、前体域（Arg^{17}-Asp^{113}）和成熟域（Leu^{114}-Val^{330}）三部分组成，存在1个潜在的糖基化位点（Asn^{97}），6个底物结合位点（Leu^{182}，Met^{183}，Ala^{249}，Leu^{275}，Gly^{278}和Ser^{324}），1个由半胱氨酸（Cys^{138}）、组氨酸（His^{277}）和天冬酰胺（Asn^{297}）残基构成的催化位点和2个组织蛋白酶L签名序列（$E^{42}X_3RX_3WX_2NX_3IX_3N^{61}$和$G^{74}X_1NX_1YX_1D^{80}$）。poCL2基因cDNA全序列及预测的氨基酸序列如图2所示。

```
CAGGCAACGAGGGAAGTCAAGATGTTTCGTTTCGCCATTGTAGCCGCCTTAGTGGCCGTG     60
                    M  F  R  F  A  I  V  A  A  L  V  A  V        13
AGTTTCGCTCGTGTTCCACGTGTTGGGCTGGACAATGAGTGGAATATATTCAAGAAACAA    120
S  F  A  R  V  P  R  V  G  L  D  N  E  W  N  I  F  K  K  Q      33
TACAACAAACTCTACCAAAACGAAGAGGAGGCCAGAAGGCGATTGGTATGGGAGAGCAAT    180
Y  N  K  L  Y  Q  N  E  E  E  A  R  R  L  V  W  E  S  N        53
TTAGACTTCATTACCCTGCACAATCTGGCTGCTGACCGCGGAGAGCACACCTTCTGGGTG    240
L  D  F  I  T  L  H  N  L  A  A  D  R  G  E  H  T  F  W  V     73
GGAATGAACGAATATGGAGATATGACAAACGAGGAGTTCACAAAGACAATGAACGGATAC    300
G  M  N  E  Y  G  D  M  T  N  E  E  F  T  K  T  M  N  G  Y     93
AGAATGAGAAACAAGACCAGCAATGCTCCTGTGTTCATGCCACCAAACAACATGGGTGAC    360
R  M  R  N  K  T  S  N  A  P  V  F  M  P  P  N  N  M  G  D ↑  113
TTACCCGATACAGTTGATTGGAGGCCGAAAGGATACGTCACACCAATCAAAAACCAGGGT    420
L  P  D  T  V  D  W  R  P  K  G  Y  V  T  P  I  K  N  Q  G     133
CAGTGCGGATCCTGCTGGTCATTCTCTGCCACTGGTTCCTTGGAAGGTCAGACTTTTAAG    480
Q  C  G  S  C  W  S  F  S  A  T  G  S  L  E  G  Q  T  F  K     153
AAGACAGGCAAACTTGTGTCACTCTCAGAACAGAATCTCGTGGACTGCTCCAAGAAACAA    540
K  T  G  K  L  V  S  L  S  E  Q  N  L  V  D  C  S  K  K  Q     173
GGAAACCATGGTTGTGAGGGAGGTCTTATGGACGATGCTTTCACCTACATTAAAGCCAAC    600
G  N  H  G  C  E  G  G  L  M  D  D  A  F  T  Y  I  K  A  N     193
AATGGAATTGACACAGAAGCTTCCTACCCATACAAGGCTAGGGACGGAAAGTGCGAGTTC    660
N  G  I  D  T  E  A  S  Y  P  Y  K  A  R  D  G  K  C  E  F     213
AAATCTGCCGATGTCGGTGCTACAGATACTGGATTTGTTGACATCAAGACTAAGGACGAG    720
K  S  A  D  V  G  A  T  D  T  G  F  V  D  I  K  T  K  D  E     233
GAAGCCCTTAAACAAGCCGTAGCCACCGTGGGTCCCATCAGCGTTGCCATTGATGCCAGC    780
E  A  L  K  Q  A  V  A  T  V  G  P  I  S  V  A  I  D  A  S     253
CACATGAGCTTCCAGCTTTACAGGACTGGGGTATACCACGACTGGTTCTGTAGCCAGACC    840
H  M  S  F  Q  L  Y  R  T  G  V  Y  H  D  W  F  C  S  Q  T     273
AAGTTGGACCATGGTGTATTGGCTGTAGGCTACGGTACTGAGGACTCAAAGGACTACTGG    900
K  L  D  H  G  V  L  A  V  G  Y  G  T  E  D  S  K  D  Y  W     293
CTGGTTAAGAACAGTTGGGGTGAGAGCTGGGGACAGAAGGGATACATCCAGATGTCCAGG    960
L  V  K  N  S  W  G  E  S  W  G  Q  K  G  Y  I  Q  M  S  R     313
AATCGCCGTAACAACTGTGGTATTGCTACATCCGCCAGCTATCCCACTGTGTAAAATGGA   1020
N  R  R  N  N  C  G  I  A  T  S  A  S  Y  P  T  V  *          330
CTTCTTCATCTAATGAAAATAAACTCTTTTATGATTTATTGAATTATTTTTGTATCGAAC   1080
TAAAAAATAAAAAA                                                 1094
```

图2 poCL2 cDNA全序列及预测的氨基酸序列

信号肽以单下划线标出；组织蛋白酶L签名序列$E^{42}X_3RX_3F/WX_2NX_3IX_3N^{61}$和$G^{74}X_1NX_1FX_1D^{80}$以阴影表示；潜在的糖基化位点（$Asn^{97}$）以加粗表示；前体域切点以箭头表示；保守的催化位点（Cys^{138}，His^{277}和Asn^{297}）用加粗、方框标出；6个底物结合位点（Leu^{182}，Met^{183}，Ala^{249}，Leu^{275}，Gly^{278}和Ser^{324}）以双下划线标出；终止密码子以星号表示；典型的polyA加尾信号AATAAA用方框标出.

Fig. 2 Full cDNA sequence and predicted amino acid sequence of poCL2 gene

The signal peptide is underlined. The cathepsin L signature sequences $E^{42}X_3RX_3F/WX_2NX_3IX_3N^{61}$ and $G^{74}X_1NX_1FX_1D^{80}$ are shown in shadow. The potential N-glycosylation site (Asn^{97}) is bold. The vertical arrow indicates putative cleavage site for prodomain. The conserved catalytic triad residues (Cys^{138}, His^{277} and Asn^{297}) are bold, boxed. The conserved substrate binding sites (Leu^{182}, Met^{183}, Ala^{249}, Leu^{275}, Gly^{278} and Ser^{324}) are double underlined. The asterisk marks the stop codon at the end of the open reading frame. The classical polyadenylation signal is boxed.

2.4　同源性分析

利用GenBank数据库，经BLAST和MatGat软件分析发现，poCL2预测的氨基酸序列与其他物种的组织蛋白酶L氨基酸序列具有较高的相似性和同源性，相似性在61.5%～73.9%之间，同源性在44.4%～59.8%之间，从表1可以看出poCL2预测的氨基酸序列与甲壳动物凡纳滨对虾（*Litopenaeus vannamei*）的相似性（73.9%）和同源性（59.8%）都最高。

结合BLAST和SMART分析结果，使用Clustal W软件，将poCL2预测的氨基酸序列与成熟域氨基酸序列已经明确的人组织蛋白酶L1（P07711）、鼠组织蛋白酶L1（07154）、牛组织蛋白酶S（P25326）和鸡组织蛋白酶L（P09648）进行比对和分析，得到poCL2氨基酸序列的前体域和成熟域的切割位点在Asp^{113}和Leu^{114}残基之间。切割位点及序列联配如图3所示。

采用Clustal W和MEGA 3.0软件，以邻位相连法构建系统进化树（图4）。从系统树上可以看出，脊椎动物和无脊椎动物的组织蛋白酶L分别聚成两大支，在无脊椎动物这一支中，昆虫、甲壳动物、软体动物和棘皮动物分别聚成一小支，合浦珠母贝组织蛋白酶L2与淡水螺（*Radix peregra*）组织蛋白酶L聚成一支。

表1　poCL2氨基酸序列与其他已知物种组织蛋白酶L的氨基酸序列同源性分析

Tab. 1　Homology analysis of poCL2 amino acid sequence with other known cathepsin L amino acid sequences

物种 Species	注册号 Accession number	相似性/% Similarity	同源性/% Identity	氨基酸数目 Amino acids
合浦珠母贝 poCL2	ADC52431	100	100	330
凡纳滨对虾 *Litopenaeus vannamei*CL	CAA68066	73.9	59.8	328
尖吻鲈 *Lates calcarifer*CL	ABV59078	72.7	54.3	337
刀额新对虾 *Metapenaeus ensis*CL	AAM96000	72.4	59.6	322
玉米象 *Sitophilus zeamais*CL	BAA24442	72.2	56.6	338
马*Equus caballus*CL	XP_001494409	72.2	55.8	334
肉蝇 *Sarcophaga peregrina*CL	Q26636	72.0	55.7	339
大西洋庸鲽 *Atlantic halibut*CL	ABJ99858	72.0	55.2	336
卤虫 *Artapenaeus franciscana*CL	AAV63977	71.9	52.8	338
淡水螺 *Radix peregra*CL	ABK90856	71.8	57.3	324
斑马鱼 *Danio rerio*CL1	CAN88536	71.8	54.7	337
埃及斑蚊 *Aedes aegypti*CL	XP_001655999	71.7	56.8	339
鼠 *Rattus norvegicus*CL1	P07154	71.6	54.9	334
非洲爪蟾 *Xenopus tropicalis*CL2	CAJ83143	71.3	53.7	335
人 *Homo sapiens*CL1	NP_001903	71.2	55.1	333
光滑爪蟾 *Xenopus laevis*CL	AAH60335	71.0	54.0	335
牛 *Bos taurus*CL2	NP_776457	71.0	55.9	334
葱蝇 *Delia radicum*CL	AAL16954	70.6	53.8	337
鲑鱼 *Salmo salar*CL1	NP_001140018	70.4	54.3	338
海参 *Stichopus japonicus*CL	ABW98676	70.2	56.5	332
黑腹果蝇 *Drosophila melanogaster*CL	NP_725347	69.2	52.3	341
海胆 *Strongylocentrotus purpuratus*CL	XP_783218	68.9	53.7	334
家蚕 *Bombyx mori*CL	NP_001037464	68.6	53.2	341
鸡 *Gallus gallus*CL	XP_418273	65.5	46.8	336
挪威海螯虾 *Nephrops norvegicus*CL	CAA56914	64.8	51.2	324
北方长额虾 *Pandalus borealis*CL	BAC65418	64.5	48.9	318
珍珠鸡 *Taeniopygia guttata*CL	XP_002194340	61.5	44.4	330

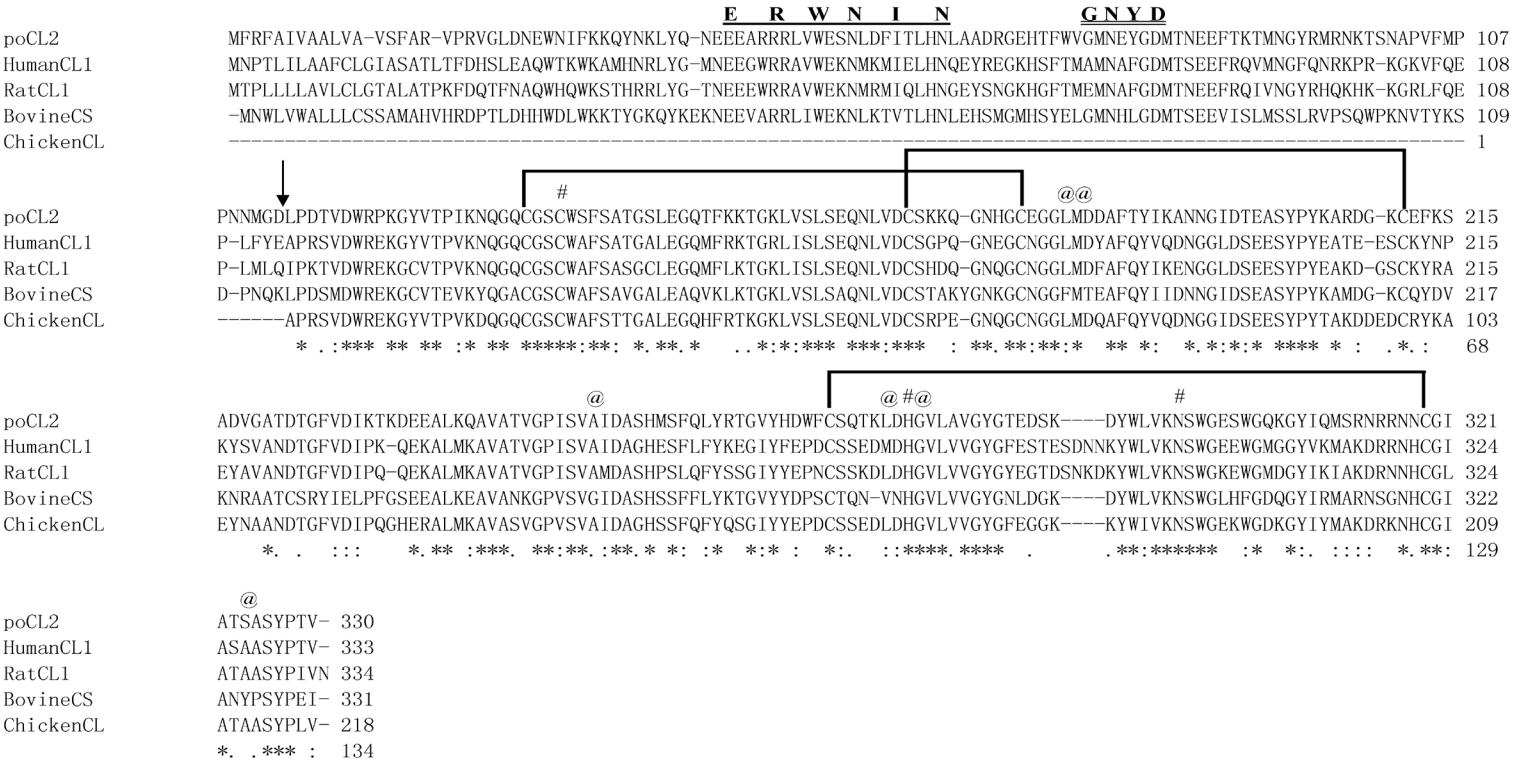

图3 poCL2预测的氨基酸与其他物种组织蛋白酶多序列联配

序列联配使用氨基酸序列包括人组织蛋白酶L1(P07711)、鼠组织蛋白酶L1(07154)、牛组织蛋白酶S(P25326)和鸡组织蛋白酶L(P09648);组织蛋白酶签名序列标签$E^{42}X_3RX_3F/WX_2NX_3IX_3N^{61}$和$G^{74}X_1NX_1FX_1D^{80}$分别用单下划线和双下划线标在序列上方;前体域和成熟域切割位点用箭头标示;推测的底物结合位点和催化位点分别以#和@标在序列上方;成熟域中保守的6个半胱氨酸形成3对二硫键(Cys^{135}–Cys^{178}, Cys^{169}–Cys^{211}和 Cys^{270}–Cys^{319})以线段相连.

Fig. 3 Multiple alignment of the predicted amino acid sequence of poCL2 with other vertebrate cathepsin amino acid sequences

The poCL2 amino acid sequence aligned with other eukaryote cathepsin amino acid sequences are from HumanCL1 (*Homo sapiens*, P07711), RatCL1 (*Rattus norvegicus*, P07154), BovineCS (*Bos taurus*, P25326), and ChickenCL (*Gallus gallus*, P09648). The signature sequences $E^{42}X_3RX_3F/WX_2NX_3IX_3N^{61}$ and $G^{74}X_1NX_1FX_1D^{80}$ motifs are shown by single and double overlines, respectively. Open arrow indicate putative cleavage site for prodomain of the poCL2. The potential substrate binding sites are indicated as at pound signs (#), signs (@), respectively, above the sequences. The six conserved cysteines are linked in the mature domain (Cys^{135}–Cys^{178}, Cys^{169}–Cys^{211}, and Cys^{270}–Cys^{319}).

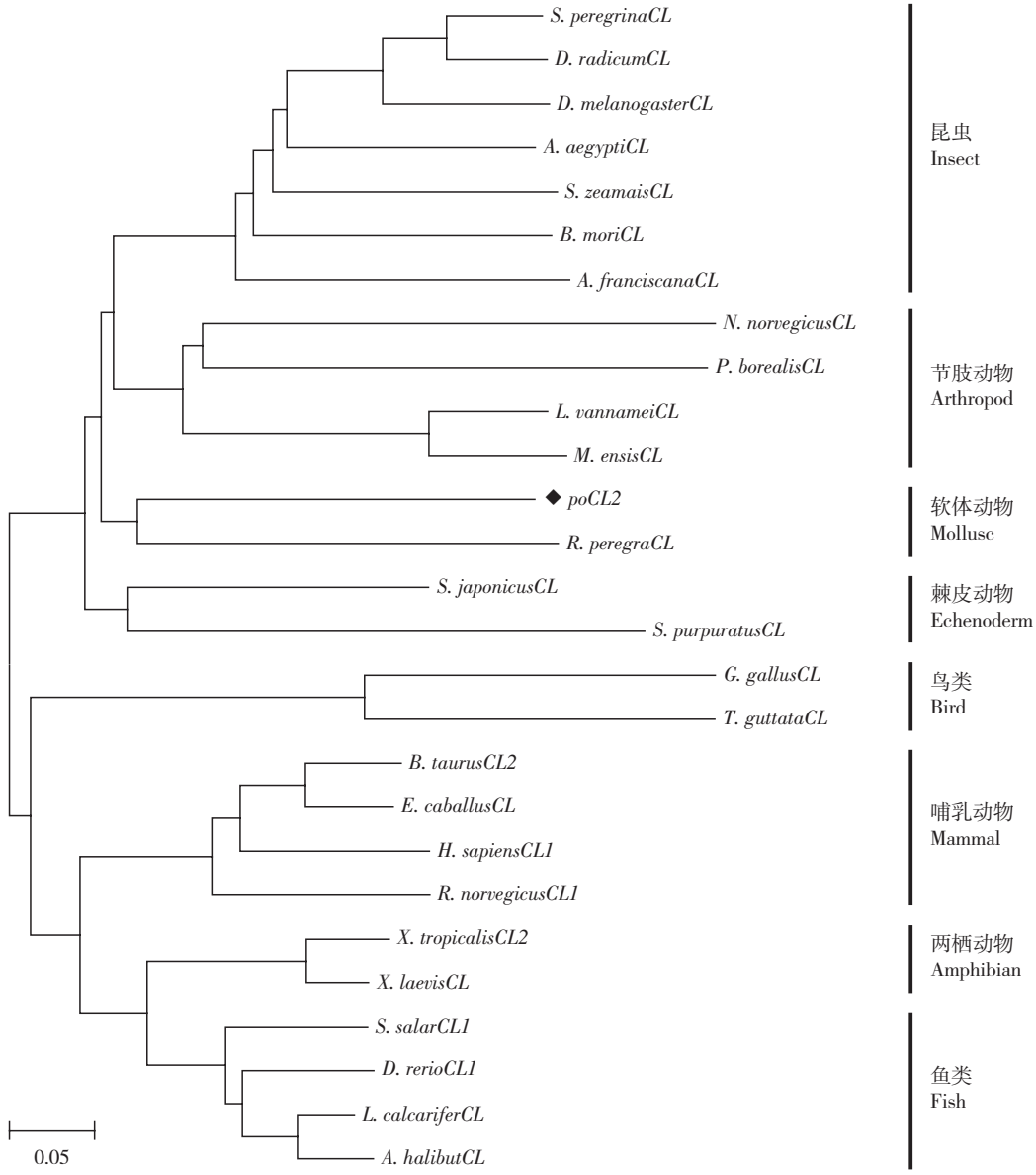

图4　利用Clustal W程序和MEGA 3.0软件构建的poCL2系统进化树
系统进化树使用序列见表1.

Fig. 4　Phylogenetic tree of the poCL2 amino acid sequence in different groups by program Clustal W and MEG A3.0
The cathepsin L amino acid sequences used in this analysis are listed in Table 1.

2.5　poCL2 mRNA组织表达特征分析

以poCL2-F和poCL2-R为引物，空白组7种不同组织cDNA为模板，*β–actin*为内参基因，利用荧光定量RT-PCR方法研究合浦珠母贝poCL2 mRNA的组织表达特征，结果如图5所示，合浦珠母贝组织蛋白酶L2 mRNA在消化腺、外套膜、肠、性腺、闭壳肌、鳃和血淋巴均有表达，其中外套膜中表达量最高，消化腺次之，血淋巴中表达最低。

2.6　poCL2 mRNA在溶藻弧菌刺激下的表达特征分析

以poCL2-F和poCL2-R为引物，利用荧光定量RT-PCR方法研究合浦珠母贝poCL2 mRNA在对照组和溶藻弧菌刺激后不同时间点相对于*β–actin*基因的表达特征。结果如图6所示，poCL2 mRNA在溶藻弧菌刺激2 h后表达逐渐上调，在溶藻弧菌刺激8 h后表达量达到最大值，是对照组的1.73倍；随后

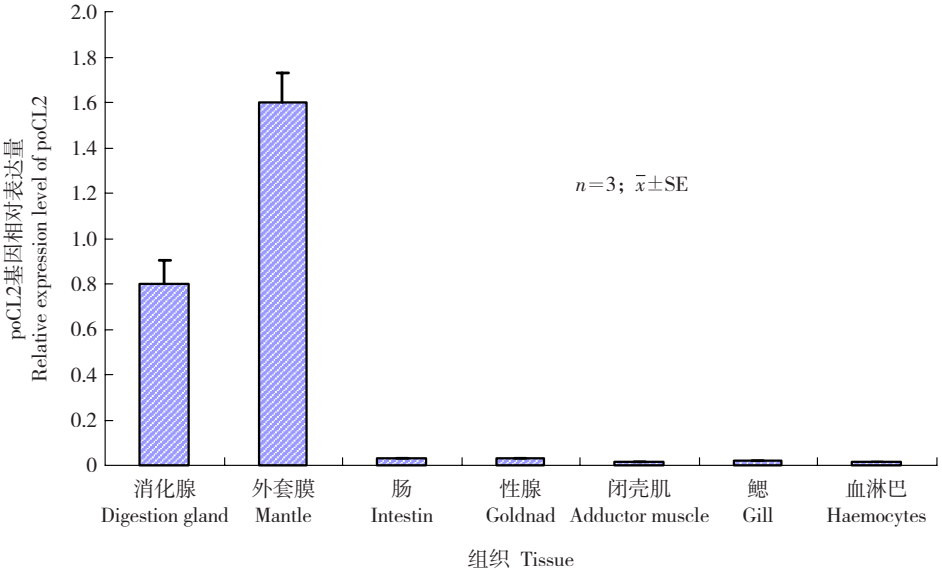

图5 合浦珠母贝 poCL2 mRNA 组织表达模式分析

Fig. 5 Expression level of the poCL2 mRNA in different tissues

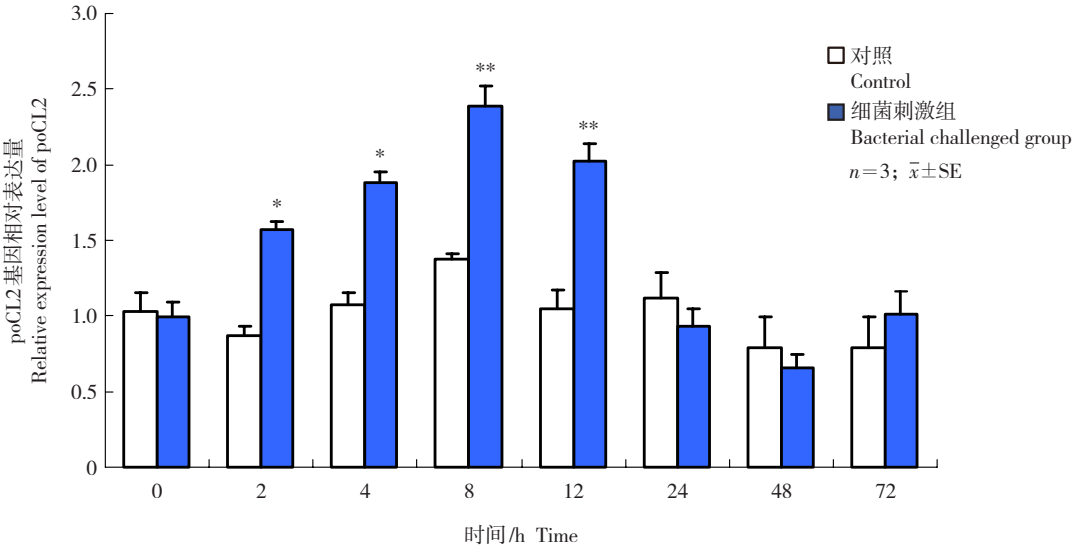

图6 合浦珠母贝 poCL2 mRNA 在溶藻弧菌刺激下的表达特征分析

"*" 表示刺激组与对照组差异显著($P<0.05$);"**" 表示差异极显著($P<0.01$).

Fig. 6 Expression level of the poCL2 mRNA in digestive gland after *V. alginolyticus* stimulation

"*" represents $P<0.05$; "**" represents $P<0.01$.

poCL2 mRNA的表达水平逐渐下降,在溶藻弧菌刺激后72 h, poCL2 mRNA的表达基本恢复到了初始水平。统计分析显示,在溶藻弧菌刺激后的2 h、4 h、8 h、12 h,溶藻弧菌刺激组poCL2 mRNA表达水平均显著高于对照组。

2.7 poCL2 mRNA在LPS刺激下的表达特征分析

同样以poCL2-F和poCL2-R为引物,利用荧光定量RT-PCR方法研究合浦珠母贝poCL2 mRNA在对照组和LPS刺激后不同时间点相对于β–*actin*基因的表达特征。结果如图7所示, poCL2 mRNA在LPS刺激后的前4 h,刺激组和对照组相比表达量没有显

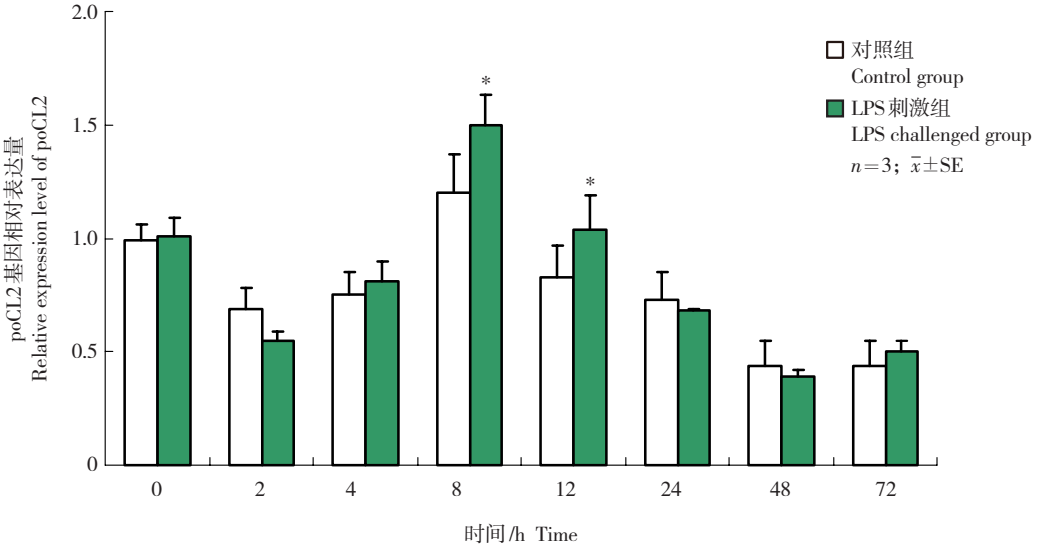

图7　合浦珠母贝poCL2 mRNA在LPS刺激下的表达特征分析
“*”表示刺激组与对照组差异显著（$P<0.05$）；“**”表示差异极显著（$P<0.01$）.
Fig. 7　Expression level of the poCL2 mRNA in digestive gland after LPS stimulation
“*” represents $P<0.05$, “**” represents $P<0.01$.

著性变化；在刺激后的8 h和12 h，poCL2 mRNA上调表达，随后表达水平呈下降趋势。刺激8 h后表达量达到最大值，是对照组的1.25倍。统计分析显示，在LPS刺激后的8 h和12 h，poCL2 mRNA表达水平均显著高于对照组（$P<0.05$）。

3　讨论

本实验以合浦珠母贝为研究对象，在双壳贝类中克隆得到组织蛋白酶L全长cDNA序列。其前体域含有2个典型的组织蛋白酶L保守签名序列标签ERFNIN[23]和GNYD[24]。成熟域中的6个保守的半胱氨酸残基可以形成3对保守的二硫键，它们对于维持poCL2成熟域的空间构象是必须的[4–5]。糖基化位点预测显示，合浦珠母贝poCL2氨基酸序列只有1个潜在糖基化位点，其他物种如斑点叉尾鮰（*Ictalurus punctatus*）[25]、刚地弓形虫（*Toxoplasma gondiis*）[26]和锥虫（*Trypanosoma carassii*）[27]的组织蛋白酶L分别有1到3个糖基化位点，而大片吸虫（*Fasciola gigantica*）[28]的组织蛋白酶L没有糖基化位点，这暗示了不同物种中组织蛋白酶L以前体酶原形式合成后的转运机制存在差别。进化分析表明，合浦珠母贝组织蛋白酶L2位于无脊椎动物这一支，与软体动物淡水螺（*R . peregra*）的亲缘关系最近，这与传统分类学方法得到的结果基本一致。

Hu和Leung[29]的研究表明，组织蛋白酶L在刀额新对虾（*M. ensis*）的血淋巴、胃、肠、精巢、卵巢和肌肉中均有表达，其中血淋巴的表达量最高，是其他被测组织的10倍以上[29]。在斑马鱼中，Tingaud-Sequeira和Cerdà克隆了3种组织蛋白酶L（Ctsl-a，b，c），其中Ctsl-a在肝脏、肾脏、鳃、肠、精巢和卵巢中均有表达，在肝脏、鳃和卵巢中表达量最高[30]。组织表达模式分析表明poCL2 mRNA在消化腺、外套膜、肠、性腺、闭壳肌、鳃和血淋巴均有表达，外套膜中表达量最高，消化腺次之，血淋巴中表达最低，其中外套膜中poCL2 mRNA的表达量是消化腺表达量的2倍，是其他被测组织的48倍以上。这些结果说明尽管合浦珠母贝组织蛋白酶L2在不同组织中是以组成型的方式表达，但是不同物种的组织蛋白酶L的组织表达模式是有差别的，可能参与不同的功能活动。

De Gregorio对黑腹果蝇（*Drosophila melanogaster*）在细菌刺激条件下免疫基因的表达变化进行了研究，发现果蝇组织蛋白酶L表达量在细菌刺激后

显著上调,认为组织蛋白酶L是一种参与机体免疫反应的重要基因[31]。另外,Nair和Del Valle对海胆(*Strongylocentrotus purpuratus*)的研究表明,组织蛋白酶L的表达也受到LPS刺激的影响,LPS能显著上调其表达量,推测组织蛋白酶L是一种重要的免疫相关基因[32]。Venier研究了紫贻贝(*Mytilus galloprovincialis*)在含有重金属的应激养殖环境下免疫相关基因的表达情况,发现组织蛋白酶L基因的表达量显著高于对照组,认为其可能参与了免疫应激反应[33]。贝类的消化腺能分泌众多参与消化和免疫防御的蛋白酶和蛋白酶抑制剂,是一种非常重要的免疫器官[34]。弧菌是海洋贝类重要的致病菌,LPS是革兰氏阴性菌的细胞壁成份,能刺激多种细胞因子的表达[35]。为了进一步探究合浦珠母贝组织蛋白酶L2可能的生物学功能,本实验分别进行了poCL2 mRNA在溶藻弧菌或LPS刺激条件下的表达特征分析研究,结果表明poCL2 mRNA在2种刺激条件下均显著上调表达,在刺激后8 h表达量都达到最高值,分别是对照组的1.73倍和1.25倍,说明poCL2 mRNA的表达能被溶藻弧菌或LPS诱导,但是溶藻弧菌诱导效果的获得,是否是由细胞壁LPS成分直接作用的引起,还需要进一步深入研究才能证实。

以上实验结果的获得,说明poCL2可能参与了合浦珠母贝的免疫应答反应,暗示其在合浦珠母贝先天免疫调控中发挥重要作用。

参考文献:

[1] Wu X, Pan J. Studies on rickettsia-like organism disease of the tropical marine pearl oyster I: the fine structure and morphogenesis of *Pinctada maxima* pathogen rickettsia-like organism [J]. J Invertebr Pathol, 1999, 73: 162–172.

[2] Lau K W, Ren J, Wai N L, et al. *Marinomonas ostreistagni* sp. nov., isolated from a pearl oyster culture pond in Sanya, Hainan Province, China [J]. Int J Syst Evol Microbiol, 2006, 56: 2271–2275.

[3] Richard K P, Jackie A C. Environmental contaminants influencing immune function in marine bivalve mollusks [J]. Fish Shellfish Immunol, 1995, 8: 581–595.

[4] Barrett A J, Rawlings N D. Evolutionary lines of cysteine peptidases [J]. Biol Chem, 2001, 382: 727–733.

[5] Berti P J, Storer A C. Alignment/phylogeny of the papain superfamily of cysteine proteases [J]. J Mol Biol, 1995, 246: 273–283.

[6] Nakagawa T, Roth W, Wong P, et al. Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus [J]. Science, 1998, 280: 450–453.

[7] Lecaille F, Kaleta J, Broemme D. Human and parasitic papain-like cysteine proteases: Their role in physiology and pathology and recent developments in inhibitor design [J]. Chem Rev, 2002, 102, 4459–4488.

[8] Villadangos J A, Ploegh H L. Proteolysis in MHC class II antigen presentation: who's in charge [J]. Immunity, 2000, 12: 233–239.

[9] Dahl S W, Halkier T, Lauritzen C, et al. Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing [J]. Biochemistry, 2001, 40: 1671–1678.

[10] Liu J, Sukhova G K, Yang J T, et al. Cathepsin L expression and regulation in human abdominal aortic aneurysm, atherosclerosis, and vascular cells [J]. Atherosclerosis, 2006, 184: 302–311.

[11] Hu K J, Leung P C. Shrimp cathepsin L encoded by an intronless gene has predominant expression in hepatopancreas, and occurs in the nucleus of oocyte [J]. Comp Biochem Physiol B Biochem Mol Biol, 2004, 137: 21–33.

[12] Hu K J, Leung P C. Food digestion by cathepsin L and digestion-related rapid cell differentiation in shrimp hepatopancreas [J]. Comp Biochem Physiol B Biochem Mol Biol, 2007, 146: 69–80.

[13] Tingaud-Sequeira A, Cerdà J. Phylogenetic relationships and gene expression pattern of three different cathepsin L (Ctsl) isoforms in zebrafish: Ctsla is the putative yolk processing enzyme [J]. Gene, 2007, 386: 98–106.

[14] Wang Y, Zhao B, Ding F, et al. Gut-specific expression of cathepsin L and B in amphioxus *Branchiostoma belcheri tsingtauense larvae* [J]. Eur J Cell Biol, 2008, 87: 185–193.

[15] Bendtsen J D, Nielsen H, Von-Heijne G, et al. Improved prediction of signal peptides: SignalP 3.0 [J]. J Mol Biol, 2004, 340: 783–795.

[16] Schultz J, Milpetz F, Bork P, et al. SMART, a simple modular architecture research tool: identification of signaling domains [J]. Proc Natl Acad Sci USA, 1998, 95: 5857–5864.

[17] Campanella J J, Bitincka L, Smalley J. MatGat: An application that generates similarity/identity matrices using protein or DNA

sequences [J]. BMC Bioinformatics, 2003, 4: 1–4.

[18] Letunic I, Copley R P, Pils B, et al. SMART 5: domains in the context of genomes and networks [J]. Nucleic Acids Res, 2006, 34: 257–260.

[19] Altschul S F, Madden T L, Schaffer A A, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs [J]. Nucleic Acids Res, 1997, 25: 3389–3402.

[20] Kongkerd N, Uparanukraw P, Morakote N, et al. Identification and characterization of a cathepsin L-like cysteine protease from *Gnathostoma spinigerum* [J]. Mol Biochem Parasitol, 2008, 160: 129–137.

[21] Kumar S, Tamura K, Nei M. MEGA3: An integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. Brief Bioinform, 2004, 5: 150–163.

[22] Zhang D C, Jiang J J, Jiang S G, et al. Molecular characterization and expression analysis of a putative LPS-induced TNF-α factor (LITAF) from pearl oyster *Pinctada fucata* [J]. Fish Shellfish Immunol, 2009, 27: 391–396.

[23] Karrer K M, Peiffer S L, Ditomas M E. Two distinct gene subfamilies within the family of cysteine protease genes [J]. Proc Natl Acad Sci USA, 1993, 90: 3063–3067.

[24] Vernet T, Berti P J, De-Montigny C, et al. Processing of the papain precursor. The ionization state of a conserved amino acid motif within the Pro-region participates in the regulation of intramolecular processing [J]. J Biol Chem , 1995, 270: 1038–1046.

[25] Yeh H Y, Klesius P H. Channel catfish, *Ictalurus punctatus*, cysteine proteinases: Cloning, characterisation and expression of cathepsin H and L [J]. Fish Shellfish Immunol, 2009, 26: 332–338.

[26] Huang R, Que X C, Hirata K, et al. The cathepsin L of *Toxoplasma gondii* (TgCPL) and its endogenous macromolecular inhibitor, toxostatin [J]. Mol Biochem Parasitol, 2009, 164: 86–94.

[27] Ruszczyk A, Forlenza M, Savelkoul H F, et al. Molecular cloning and functional charactrisation of a cathepsin L-like proteinase from the fish kinetoplastid parasite *Trypanosoma carassii* [J]. Fish Shellfish Immunol, 2008, 24: 205–214.

[28] Yamasaki H, Mineki R, Murayama K, et al. Characterisation and expression of the *Fasciola gigantic* cathepsin L gene [J]. Int J Parasitol, 2002, 32: 1031–1042.

[29] Hu K J, Leung P C. Shrimp cathepsin L encoded by an intronless gene has predominant expression in hepatopancreas, and occurs in the nucleus of oocyte [J]. Comp Biochem Physiol B Biochem Mol Biol, 2004, 137: 21–33.

[30] Tingaud-Sequeira A, Cerdà J. Phylogenetic relationships and gene expression pattern of three different cathepsin L (Ctsl) isoforms in zebrafish: Ctsla is the putative yolk processing enzyme [J]. Gene, 2007, 386: 98–106.

[31] Tisca P G, Mosca F. Defense mechanisms in farmed marine mollusks [J]. Vet Res Commun, 2004, 28: 57–62.

[32] De-Gregorio E, Spellman P T, Rubin G M, Lemaitre B. Genomewide analysis of the *Drosophila melanogaster* immune response by using oligonucleotide microarrays [J]. Proc Natl Acad Sci USA, 2001, 98: 12590–12595.

[33] Nair S V, Del V H, Gross P S, et al. Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin. Identification of unexpected immune diversity in an invertebrate [J]. Physiol Genomics, 2005, 22: 33–47.

[34] Venier P, De-Pitta C, Pallavicini A, et al. Development of mussel mRNA profiling: Can gene expression trends reveal coastal water pollution [J]. Mutat Res, 2006, 602: 121–134.

[35] Chappell V L, Le L X, LaGrone L, et al. STAT proteins play a role in tumor necrosis factor alpha gene expression [J]. Shock , 2000, 14: 400–403.

Molecular characterization and expression analysis of cathepsin *L2* cysteine protease from pearl oyster *Pinctada fucata*

MA Jianjun[1,2], ZHANG Dianchang[2], CUI Shuge[1,2], JIANG Jingjing[2,3], PU Hanlin[1], JIANG Shigui[2]

(1. School of Life Science and Technology, Jinan University, Guangzhou 510632, China; 2. Division of Aquaculture and Biotechnology, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510300, China; 3. School of Fisheries and Life Science, Shanghai Ocean University, Shanghai 200090, China)

Abstract: Pearl oyster *Pinctada fucata* is the most popular farming shellfish for seawater pearl production in Guangdong, Guangxi and Hainan province of China. In recent years, as other marine animals, higher frequencies of disease epidemics and the emergence of new diseases have been reported in artificial cultivation of Pearl oyster. Many researchers considered the reasons for high mortality were ocean pollution, disease outbreaks and stock degeneration. In order to control disease and enhance the yields and quality of seawater pearl, it is necessary to further research the innate immune mechanisms of pearl oyster. Cathepsin L is a member of cysteine protease family and involved in various biological functions, which is distributed widely in living organisms. In this study, we identified one cDNA encoding a cathepsin L cysteine proteases from EST library of pearl oyster *Pinctada fucata* (designated as poCL2). The poCL2 cDNA was 1 094 bp long and consisted of a 5′-untranslated region (UTR) of 21 bp, a 3′-UTR of 80 bp with a polyadenylation signal (AATAAA), and an open reading frame (ORF) of 993 bp encoding a polypeptide of 330 amino acids, which contained a typical signal peptide sequence (Met^{1}–Ala^{16}), a prodomain (Arg^{17}–Asp^{113}), and a mature domain (Leu^{114}–Val^{330}). The preprocathepsin contained a potential N-glycosylation site (Asn^{97}), six substrate binding sites (Leu^{182}, Met^{183}, Ala^{249}, Leu^{275}, Gly^{278}, and Ser^{324}), three catalytic sites (Cys^{113}, His^{278}, and Asn^{298}). The conserved $E^{42}X_3RX_3WX_2NX_3IX_3N^{61}$ motif is discovered in the prodomain which may be play important function in the inhibition of proteolytic activity. The $G^{74}X_1NX_1YX_1D^{80}$, which may be related to pH-dependent intramolecular processing, is also found in the prodomain of the poCL2. Position of prodomain cleavage site and conserved cysteine residues are based on the N-terminal sequence information from HumanCL1 (*Homo sapiens*, P07711), RatCL1 (*Rattus norvegicus*, P07154), BovineCS (*Bos taurus*, P25326), and ChickenCL (*Gallus gallus*, P09648), and the prodomain cleavage site was between Asp^{113} and Leu^{114}.The number of N-glycosylation site is not conserved, there is one, two and three potential N-glycosylation sites in *Ictalurus punctatus*, *Toxoplasma gondiis* and *Trypanosoma carassii* cathepsin L, respectively, which revealed that the intracellular transport mechanism of cathepsin L proteases is different in living organisms. Homology analysis of poCL2 by MatGAT software revealed that the poCL2 shared 61.5–73.9% similarity and 44.4–59.8% identity to other known cathepsin L sequences. The phylogenetic tree showed that the poCL2 clustered with the invertebrate cathepsin L cysteine proteases and was closely related to *Radix peregra* cathepsin L, and a clear clade division was observed between vertebrates and invertebrates cathepsin L proteins. The previous study demonstrated that the expression pattern of cathepsin L was different in different species and exhibited specificity for some tissues in different species. The mRNA expression of the poCL2 in normal group could be detected in digestive gland, gonad, haemocytes, gills, mantle, adduct muscle and intestines, and with the higher level in mantle, the results indicated that the expression pattern of cathepsin L gene in various tissues was somewhat different, and had diverse functions. The digestive gland of mollusk was thought to be an important immune organ, which could secrete various enzymes to hydrolyze microorganisms and involve in digestive and defense functions. So we selected the digestive gland to research the temporal expression pattern of poCL2 after *Vibrio alginolyticus* and LPS challenge. The results showed that the expression level of the poCL2 was up-regulated in digestive gland not only in *V. alginolyticus* group but also in LPS group, and the highest expression was observed at 8 h after stimulation and was 1.73-fold, 1.25-fold higher than the control group, respectively. These results suggested that the poCL2 was involved in the innate immune response of pearl oyster and might play an important function in immune regulation. [Journal of Fishery Sciences of China, 2010, 17(4): 701–712]

Key words: *Pinctada fucata*; cathepsin L; innate immunity

Corresponding author: JIANG Shigui. E-mail: jiangsg@21cn.com