

# 曼氏无针乌贼(*Sepiella maindroni*)ALSM 基因的 克隆及序列分析\*

周 超 郭宝英 刘慧慧 吴常文

(浙江海洋学院海洋科学学院 国家海洋设施养殖工程技术研究中心 舟山 316004)

**提要** 采用 RT-PCR 和 RACE 技术分离、克隆得到曼氏无针乌贼肝脏 ALSM (Astacin-like squid metalloprotease)基因。序列分析表明, ALSM cDNA 序列全长 1418bp, 其包括 53bp 的 5'端非翻译区, 1290bp 阅读框以及含有 poly(A)信号 AATAAA 的 75bp 3'端非翻译区, 阅读框编码 429 个氨基酸残基。推导蛋白的相对分子质量为 48916.31Da, 等电点为 7.978。ALSM 氨基酸序列的同源性分析显示, 不同进化地位动物的 ALSM 氨基酸序列都具有较高的同源性。系统发生树分析发现, 曼氏无针乌贼 ALSM 首先与金乌贼、莱氏拟乌贼、长枪乌贼及太平洋褶柔鱼的 ALSM- $\alpha$ 亚型聚类, 再与其 ALSM- $\beta$ 及 ALSM- $\gamma$ 亚型聚类。生物信息学分析表明, ALSM 编码的蛋白为一种亲水性的分泌蛋白。这些结果对进一步研究 ALSM 的结构与功能以及从基因调控角度探索曼氏无针乌贼在养殖条件下生物学特性变化机理具有重要意义。

**关键词** 曼氏无针乌贼, 虾红素, ALSM, 金属蛋白酶, 克隆, 序列分析

**中图分类号** Q789

金属蛋白酶(Metalloproteinase)是其活性中心依赖于金属离子的一类蛋白酶, 大多数金属蛋白酶是  $Zn^{2+}$ 金属蛋白酶。其中, 虾红素金属蛋白酶家族是金属蛋白酶家族中比较重要的成员(Jiang *et al*, 1992)。ALSM (Astacin-like squid metalloprotease)基因最初从太平洋褶柔鱼的外套膜肌肉中被分离发现。其为虾红素金属蛋白酶家族中的一种分泌蛋白(Yokozawa *et al*, 2002)。虾红素金属蛋白酶家族具有较高的保守性, 它们均含有保守的锌结合域(HEXXHXXGFXHEXXRXDRD)和 Met-turn 结构(SXMHY)。某些家族成员还具有非常保守的羧基末端区域(C-terminal domain)(Bode *et al*, 1993)。虾红素金属蛋白酶家族成员具有一系列不同的生物学功能, 包括消化(Vogt *et al*, 1989)、背腹决定(Shimell *et al*, 1991; Blader *et al*, 1997; Piccolo *et al*, 1997)、决定形态发生(Yan *et al*, 1995)、裂解卵壳

(Yasumasu *et al*, 1992; Lee *et al*, 1994)、降解多肽(Butler *et al*, 1987; Barnes *et al*, 1989; Yamaguchi *et al*, 1993)以及免疫和造血(Tsai *et al*, 2004)等。ALSM 主要分布于物种孵化胚胎及成熟个体的消化器官和肌肉组织中(Kanzawa *et al*, 2005)。这表明 ALSM 在生物生长发育及物种成熟个体中都发挥重要作用。

虾红素金属蛋白酶家族成员已在从水螅到人类等不同进化地位的生物物种中逐渐被发现。近年来, 在水产动物中研究较多。樱鲮(*Oncorhynchus masou*)、青鳞(*Oryzias latipes*)、鲤鱼(*Cyprinus carpio*)、斑马鱼(*Danio rerio*)、鳗鲡(*Anguilla anguilla*)、底鳉(*Fundulus heteroclitus*)、大唇朴丽鱼(*Haplochromis chilotes*)、大菱鲂(*Scophthalmus maximus*)等物种中均有虾红素金属蛋白酶家族成员被发现、研究(Inohaya *et al*, 1997; Yasumasu *et al*, 1996; Hung *et al*, 1997; Song *et al*,

\* 国家国际科技合作项目, 2009DFB20290号; 国家科技支撑计划, 2011BAD13B08号; 浙江省海水增养殖重点实验室开放基金, 601.90054号。周 超, E-mail: zc43115@163.com

通讯作者: 吴常文, 教授, E-mail: wucw08@126.com

收稿日期: 2011-10-28, 收修改稿日期: 2011-11-18

2004; Hiroi *et al*, 2004; Kawaguchi *et al*, 2005; Kijimoto *et al*, 2005; 元飞, 2006<sup>1)</sup>)。而在头足类中仅太平洋褶柔鱼(*Todarodes pacificus*)、长枪乌贼(*Loligo bleekeri*)、金乌贼(*Sepia esculenta*)及莱氏拟乌贼(*Sepioteuthis lessoniana*)等几个物种中见报道(Yokozawa *et al*, 2002; Nobuyuki *et al*, 2008)。

目前, 在头足类领域关于基因克隆的研究已逐步开展, 但关于曼氏无针乌贼(*Sepiella maindroni*)基因研究还少见报道, 涉及生命过程的重要酶类——ALSM 更无人涉及。本研究以浙江沿海处于资源严重衰退状态中的曼氏无针乌贼作为研究对象, 通过同源克隆和 RACE 技术获得 ALSM 基因的 cDNA 全长序列, 探索其结构特征, 并进行聚类分析, 这对进一步研究 ALSM 的结构与功能以及从基因调控角度探索曼氏无针乌贼在养殖条件下生物学特性变化机理具有重要意义。

## 1 材料与方法

### 1.1 RNA 的提取

曼氏无针乌贼(*Sepiella maindroni*)样品购于秀山养殖场, 活体解剖, 取 50—100mg 肝脏组织放入液氮速冻后立即放入 TRIZOL(Invitrogen)匀浆, 按 RNA 提取说明书提取总 RNA。通过琼脂糖凝胶电泳、溴乙锭染色检测 RNA 的完整性。最后通过分光光度计测定 RNA 的浓度, 并根据  $OD_{260}/OD_{280}$  判断 RNA 纯度。将高质量的总 RNA 保存于  $-80^{\circ}\text{C}$  备用。

### 1.2 部分序列的克隆

根据 GenBank 中已报道的金乌贼和莱氏拟乌贼 ALSM 基因的 cDNA 序列设计引物。正向引物为 ALSMF, 反向引物为 ALSMR(表 1)。用反转录试剂盒(TaKaRa)进行反转录。25 $\mu\text{l}$  PCR 反应体系为 2.5 $\mu\text{l}$  cDNA, 2.5 $\mu\text{l}$  10 $\times$ 缓冲液, 0.75 $\mu\text{l}$   $\text{Mg}^{2+}$  (25mmol/L), 1 $\mu\text{l}$  dNTPs (10mmol/L), 0.5 $\mu\text{l}$  *Taq* 酶(1U/ $\mu\text{l}$ ), 每条引物各 1 $\mu\text{l}$  (10 $\mu\text{mol/L}$ ), 15.75 $\mu\text{l}$  灭菌无核水。PCR 反应程序为:  $95^{\circ}\text{C}$  预变性 5min;  $94^{\circ}\text{C}$  1min,  $55^{\circ}\text{C}$  1min,  $72^{\circ}\text{C}$  90s, 29 个循环;  $72^{\circ}\text{C}$  延伸 10min, 最后  $-20^{\circ}\text{C}$  保存。PCR 产物用琼脂糖凝胶回收试剂盒(TaKaRa)进行纯化, 纯化后连接到 pMD-18T 载体(TaKaRa)。连接产物转染 DH5 $\alpha$ 感受态细胞(TaKaRa)。PCR 筛选获得的阳性克隆, 选 3 个送到 Invitrogen(上海)公司测序。获得的序列用 BLASTX 进行比对分析, 发现该序列与头足类中已知的 ALSM 基因同源。

### 1.3 快速扩增 cDNA 末端

根据上述已获得的曼氏无针乌贼 ALSM 基因部分序列设计引物, 用 SMART<sup>TM</sup>RACE cDNA 扩增试剂盒(Clontech)进行 RACE 扩增。3'RACE: 第一轮 PCR 用特异性引物 3'157-1 和 UPM 进行扩增。第二轮 PCR 用特异性引物 3'157-2 和 UPM 进行扩增。对获得的 PCR 产物进行纯化、连接、转化、测序及拼接。5'RACE: 第一轮 PCR 用基因特异性引物 GSP-2 和桥连铆钉引物 AAP 对已经加 dC 尾的 cDNA 进行扩增。第二轮 PCR 用基因特异性引物 GSP-3 和

表 1 核苷酸引物序列  
Tab.1 Oligonucleotide primer sequences

引物	序列(5' 3')
ALSMF	GGTCTCCATATGATTACAATTC
ALSMR	TCATCTGTTCCATTTATTCCTT
3'157-1	TTGGTCAGCGTATTGGTATGAGTCAGTTGG
3'157-2	ACTTGGGTACGTTGGTCTGGAGTAACATTC
R321-1 (GSP1)	CAACCACAATAACCAGC
R321-2 (GSP2)	TTCAGGACATCCATAGAATCTC
R321-3 (GSP3)	CTCATACCAATACGCTGACCAA
UPM	Long: CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT Short: CTAATACGACTCACTATAGGGC
AAP	GGCCACGCGTCGACTAGTACGGGHHGGGHHGGGHHG
AUAP	GGCCACGCGTCGACTAGTAC

1) 元 飞, 2006. RACE 技术克隆大菱鲂蛋白磷酸酶 1 催化亚基基因和金属蛋白酶基因及其序列分析. 青岛: 中国海洋大学硕士学位论文, 54—65

通用扩增引物 AUAP 进行巢式 PCR。对获得的 PCR 产物进行纯化、连接、转化、测序及拼接。

#### 1.4 序列生物信息学分析及蛋白质结构预测

对 3 个序列的测序结果进行拼接,并推断曼氏无针乌贼 ALSM 基因全长 cDNA 序列。采用生物信息学软件 DNAMAN 将 ALSM 全长 cDNA 序列翻译成氨基酸序列,利用 DNASTar 软件中的 EdiSeq 程序对其进行开放阅读框(ORF)的预测。在此基础上,将推断的曼氏无针乌贼 ALSM 序列用 BLAST program(<http://www.ncbi.nlm.nih.gov/blast>)进行序列同源性比对。对推导的蛋白序列分析用 ProtParam 软件(<http://web.expasy.org/protparam/>)预测蛋白理化性质。将推测的蛋白序列进行蛋白三级结构在线预测(<http://bmm.cancerresearchuk.org/~3djigsaw/>),并使用 RasMol 软件进行分析。采用 MEGA 4.0 软件构建 Neighbor-joining 系统进化树(Saitou *et al.*, 1987)。

## 2 结果与分析

### 2.1 曼氏无针乌贼肝脏总 RNA 的提取

电泳检测所提取的曼氏无针乌贼肝脏总 RNA,可见 28S 和 18S 两条特异性条带,表明所获得的总 RNA 无明显降解,符合实验要求。

### 2.2 曼氏无针乌贼 ALSM 基因 cDNA 部分序列克隆与分析

以引物 ALSMF 和 ALSMR 为扩增引物,采用 RT-PCR 方法分离曼氏无针乌贼 ALSM 部分 cDNA 序列,1.5%琼脂糖凝胶电泳结果显示在 500—750bp 间有一条特异性目的条带,与预期的片段长度大小相符,见图 1。PCR 产物经 TA 克隆后送出测序,结果表明该序列大小为 530bp。获得的目的序列在 NCBI 上 BLAST 同源性搜索,结果表明:获得的目的序列与金乌贼和莱氏拟乌贼 ALSM 基因型序列同源性分别为:90%、82%,故推测该序列是曼氏无针乌贼 ALSM 基因 cDNA 目的序列。

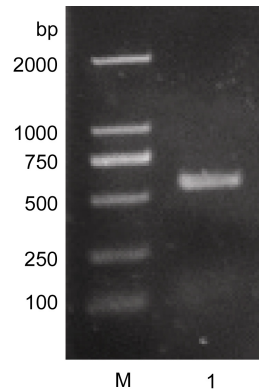


图 1 曼氏无针乌贼 ALSM 基因 cDNA 部分序列 PCR 产物电泳图谱

Fig.1 Electrophoresis map of PCR amplification product of partial ALSM cDNA of *S. maindroni*  
注: M. 2000bp ladder Marker, 1. cDNA 序列 PCR 产物

### 2.3 曼氏无针乌贼 ALSM 基因 5'、3'端 cDNA 序列克隆与分析

以通用引物 AUAP 和特异性引物 GSP-3 进行 5'端扩增,在 750—900bp 中间有一条明显的目的条带,与预期大小相符,PCR 产物经 TA 克隆并测序,获得一个长度为 823bp 的序列(图 2A)。其中包括完整的 5'端序列 780bp。该 5'端序列中含有起始密码子 ATG,以及起始密码子 ATG 前面一段 55bp 的前导序列。同样以通用引物 UPM 和特异性引物 3'157-2 进行 3'端扩增,PCR 产物经 TA 克隆并测序,获得一个长度为 527bp 的序列(图 2B),其中包括终止密码子 TGA。

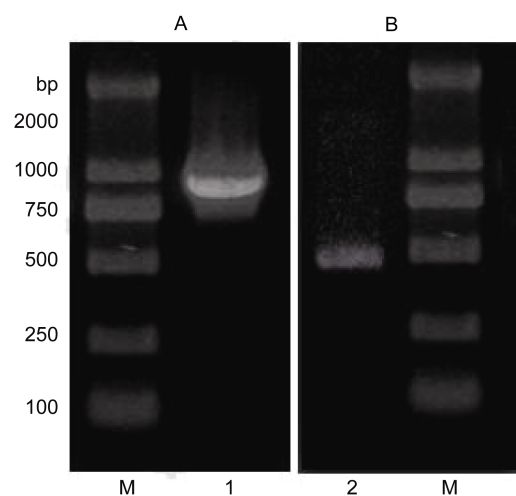


图 2 曼氏无针乌贼 ALSM 基因 5'、3' RACE 扩增结果  
Fig.2 5' and 3' RACE amplification results of *S. maindroni* ALSM gene

A. 5'RACE; B. 3'RACE; M. 2000bp ladder Marker; 1. 5'末端 cDNA 序列 PCR 产物; 2. 3'末端 cDNA 序列 PCR 产物

### 2.4 曼氏无针乌贼 ALSM 生物信息学分析

将 RT-PCR、3'RACE、5'RACE 扩增的序列进行拼接,获得曼氏无针乌贼 ALSM 基因 cDNA 全长序列(图 3)。序列分析显示:曼氏无针乌贼 ALSM 基因 cDNA 序列全长 1418bp,其包括 53bp 的 5'端非翻译区,75bp 的 3'端非翻译区和编码 429 个氨基酸残基的 1290bp 开放阅读框,并在 poly(A)尾上游 19bp 处发现一个多腺苷化信号 AATAAA。对开放阅读框中碱基组成分析发现,G、C 碱基的含量(37.75%)低于 A、T 碱基的含量(62.25%)。

根据曼氏无针乌贼 ALSM 全长 cDNA 序列推断得到其蛋白质序列。序列分析发现:该序列含 429 个氨基酸,其中包含保守的 MAM 域。利用 ProtParam

-53 CTCTTGGCTTGGTACCTCGCAAGAACTGIGATATATTTGTTAATCCTTATT  
 1 ATGGCTGGAGTTGTTACTTGGATTCTGGTTTTGTGCTTTGGATATATAACGGCTTTGACCTTGGATTGGAGTCTGCCTCTGTGCCTGGA  
 M A G V V T W I L V L C F G Y I T A L T L D L E S A S V P G  
 91 TTTGACCCACGTTATCCTGAAGAAGTACCAGGCTTTTTCAAGGTGACATTGAATGAGCAGCAATCGTTTTATCAGAAGTGGTTGGG  
 F D P R Y P E E V P G L F Q G D I E L S S N R F I R S V V G  
 181 GATCTGAAAAACGATGGCCGATAAGAAAAATCCAATTGTGATTCTATTTTGACAAAATATATCAGAAATACAATCATTGAATCA  
 D L K K R W P I R K I P I V I H S Y F D K Y I R N T I I E S  
 271 CTAAGGGAAATGAAGAAATGCACGTGTGGATGGAAAAGACTGCATAACATTTGTCAACCAGACAGATGAAGAGATCTACATTTATTT  
 L R E I E E N A R V D G K D C I T F V N Q T D E E I Y I Y F  
 361 TCCACTGGTAAAGGATGCGGTAGCCGTATTGGATACACAGGATCACGTCAGGGAGTCAAATTTGGGTTATGGATGTCGCACGAAAGGAATA  
 S T G K G C R S R I G Y T G S R Q G V K L G Y G C R T K G I  
 451 ATAATTCATGAAGTTTTACACAGTCTGGATTCTATCATGAACAAAGTCGACCTGACAGAGACAAGTATGTTAAAAATATTATGGATAAT  
 I I H E V L H S L G F Y H E Q S R P D R D K Y V K I I M D N  
 HEXXHXXGFXHEXXRXDRD Frame 1  
 541 GTTAAAGAAGCGGACAAAGAAAATTTGTGAAACTTTTACCGCCAGTGATCAACACACAGAACCTTCTTATGACTACAACCTCAATAATG  
 V K E A D K G N F V K L L P P V I N T Q N L P Y D Y N S I M  
 631 CATTATGAGCGTTCTACTTTTGGCCATCGACAGAACAAAACCAACAATGTACCATTGAAAAAGATGCCAATATTGGTCACGGTATTGGT  
 H Y E R S T F A I D R T K P T I V P L K K D A N I G Q R I G  
 Met-turn Frame 2  
 721 ATGAGTCAGTTGGACATTATACAATTACAGAGATTCTATGGATGTCCTGAAAGGAAATTCGTACCTCTTGTACAAAATAATCCTGTTTCT  
 M S Q L D I I Q L Q R F Y G C P E R K F V P L V T N N P V S  
 811 GTCAACTGTACATTTGATGCTGGTTATTTGGTTGGAAACATCTATTTATGGAACCACTGCAAAAGATAAATACTGGGTACGTTGGTCT  
 V N C T F D A G Y C G W K H L F M E P P A K I N T W V R W S  
 901 GGAGTAACATTCACAGATGGTACTGGACCAAAAAATGATCACACTATTGAAACATTTGAAGTCATTATCTTATACATCTGCTTCTGGC  
 G V T F T D G T G P K N D H T I G T F E G H Y L Y T S A S G  
 991 CGATTTTTATCTATTGCAAAAATTCACACAGGCTCATTTGCTGGTGACTATTGCTGACGTTTTGGTATCATATGTACGGCAAAGAA  
 R F L S I A K I Q T P E L I A G D Y C L T F W Y H M Y G K E  
 MAM domain Frame 3  
 1081 ATGGGCACTCTGAGGGTGAATTTAATTTGAGGAGAAAAAGGGAAGAATTTATGATTATCTCTGGAGACCAAGGAAATAAATGAAACAG  
 M G T L R V N L I E E K K K G E L M I I S G D Q G N K W K Q  
 1171 ATGAAAATCTCCATAACAGCGACTTCACATTCAAAGGTTGAATTTGAGAGTATCATGGGTGATGATGAAAGAAGTATATCTGCATTGAT  
 M K I S I T A T S H S K V E F E S I M G H D E R S D I C I D  
 1261 GATGTCCTATTTTGGAGTGGGAAATGCTGA 1290  
 D V L F L S G K C \*  
 -75 GATACAAATGTAAGGTAATAAATAAATAAATATATTCCACCATCAAAAAAAAAAAAAAAAAAAAAA

图3 曼氏无针乌贼 ALSM 基因及推导出的氨基酸序列

Fig.3 Nucleotide sequence and deduced amino-acid sequence of *S. maindroni* ALSM gene  
 注: Frame 1: “HEXXHXXGFXHEXXRXDRD” 序列; Frame 2: Met-turn 结构; Frame 3: MAM domain. 起始密码子和终止密码子用下划线标出, 星号标记的为终止密码子。多聚腺苷酸加尾信号用双下划线标出

软件预测曼氏无针乌贼 ALSM 基因编码蛋白的理化性质, 推测该蛋白的相对分子质量为 48916.31Da, 等电点为 7.978。包含 53 个碱性氨基酸(K, R), 50 个酸性氨基酸(D, E), 137 个疏水性氨基酸(A, I, L, F, W, V), 113 个极性(亲水性)氨基酸(N, C, Q, S, T, Y), 在 pH 7.0 时有 4.465 电荷。用 SignalP 软件对曼氏无针乌贼 ALSM 氨基酸序列进行信号肽分析, 分析结果表明第 18—19 个氨基酸间存在信号肽剪切位点, 说明该蛋白含有信号肽, 进而推测该蛋白为一种分泌蛋白。亲/疏水性分析结果显示, ALSM 蛋白的亲水性区域占大部分区域, 因此推测该蛋白是亲水性蛋白。

2.5 曼氏无针乌贼 ALSM 同源性分析与系统进化树构建

通过 NCBI Blast 软件对曼氏无针乌贼 ALSM 推测的氨基酸序列与金乌贼、莱氏拟乌贼、长枪乌贼及

太平洋褶柔鱼 ALSM 基因的各亚氨基酸序列进行在线比对。结果表明, 该序列与金乌贼 ALSM- 氨基酸序列的同源性最高为 82%, 与莱氏拟乌贼、长枪乌贼及太平洋褶柔鱼 ALSM- 氨基酸序列的同源性依次为 74%、71%、70%; 与长枪乌贼、金乌贼及莱氏拟乌贼 ALSM- 氨基酸序列的同源性均为 54%; 与太平洋褶柔鱼 ALSM- 氨基酸序列的同源性为 55%。

目前, 虾红素金属蛋白酶家族成员在鱼类中多见研究, 本研究对斑马鱼、点带石斑鱼、青鳞、红鳍东方鲀、太平洋牡蛎、金乌贼、莱氏拟乌贼、太平洋褶柔鱼、长枪乌贼及推断的曼氏无针乌贼金属蛋白酶家族氨基酸序列使用 Neighbor-Joining 方法构建系统进化树(图 4)。结果表明, 系统进化树分为两个分支, 头足类聚为一支, 太平洋牡蛎与其它几种鱼类聚为一支。头足

类分支又根据 ALSM 亚型分为两个小分支, ALSM- 亚型单独聚为一支, ALSM- 和 ALSM- 亚型聚为一支。在 ALSM- 亚型小分支中, 曼氏无针乌贼 ALSM 序列与金乌贼 ALSM- 亚型密切聚为一支, 之后依次与太平洋褶柔鱼、莱氏拟乌贼及长枪乌贼 ALSM- 亚型聚类。

2.6 曼氏无针乌贼 ALSM 氨基酸序列分析与蛋白质结构预测

进一步分析曼氏无针乌贼 ALSM 全长 cDNA 序列推断得到氨基酸序列(图 3), 结果发现在第 452 位氨基酸中含有一个“HEVLHSLGFYHEQSRPDRD”序列, 而“HEXXHXXGXXHXXRXDRD”为金属蛋白酶的特有序列(Bode et al, 1993), 其中“E”为虾红素家族区别其它家族所特有的残基(Jiang et al, 1992)。金属蛋白酶另一序列为“SXMXY”的保守

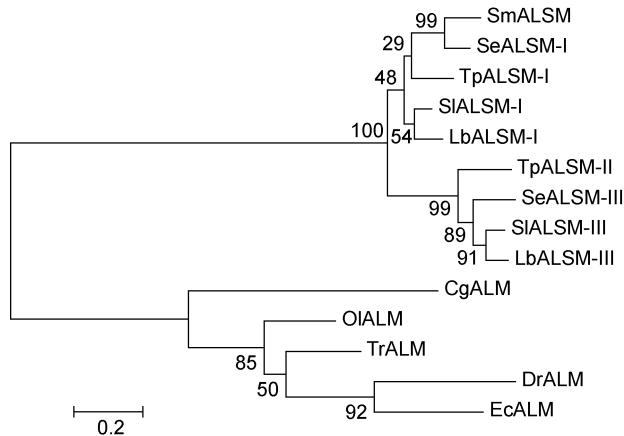


图 4 基于虾红素金属蛋白酶家族氨基酸序列 (ALM/ALSM)构建的 NJ 系统进化树

Fig.4 Phylogenetic analysis of astacin metalloprotease family amino acid sequences using the Neighbor-Joining method  
Cg: 太平洋牡蛎(*Crassostrea gigas*); Dr: 斑马鱼(*Danio rerio*); Ec: 点带石斑鱼(*Epinephelus coioides*); Ol: 青鳉(*Oryzias latipes*); Tr: 红鳍东方鲀(*Takifugu rubripes*); Se: 金乌贼(*Sepia esculenta*); Sl: 莱氏拟乌贼(*Sepioteuthis lessoniana*); Tp: 太平洋褶柔鱼(*Todarodes pacificus*); Lb: 长枪乌贼(*Loligo bleekeri*); Sm: 曼氏无针乌贼(*S. maindroni*)

Met-turn 结构(Bode *et al*, 1993), 亦同时在曼氏无针乌贼 ALSM 基因推导的氨基酸序列中发现。

利用 DNAMAN 软件对曼氏无针乌贼 ALSM 氨基酸序列与金乌贼、莱氏拟乌贼、长枪乌贼及太平洋褶柔鱼等物种 ALSM 各亚氨基酸序中的 MAM 域进行同源性比对, 结果发现曼氏无针乌贼 ALSM 氨基酸 MAM 域存在多个虾红素金属蛋白酶家族的保守位点。虾红素金属蛋白酶家族的 MAM 域中具有高度保守的 4 个 Cys 残基(Bond *et al*, 1995), 而曼氏无针乌贼 ALSM 与参与比较的其它 8 个序列同时含有这 4 个保守 Cys 位点中的 3 个(图 5 中由\*标记)。综上所述, 可以认为曼氏无针乌贼 ALSM 是虾红素金属蛋白酶家族中的一员。根据曼氏无针乌贼 ALSM 氨基酸序列预测的蛋白质三级结构如图 6 所示。

### 3 讨论

世界各地学者已对从水螅到人类等不同进化阶段生物物种中发现的虾红素金属蛋白酶家族成员的序列、结构、基因功能、蛋白质三维结构以及进化关系等进行了广泛而深入的研究。虾红素金属蛋白酶家族成员是生物发育及成熟过程中起重要作用的基因, 因而备受人们关注。ALSM 基因在生物的生长发育及其

成体的生命活动中都发挥重要作用。本研究首次克隆得到曼氏无针乌贼完整的 ALSM 基因, 全序列为 1418bp, 阅读框为 1290bp, 编码长度为 429 个氨基酸, 其中包含保守的“HEVLHSLGIFYHEQSRPDRD”序列、Met-turn 结构及 MAM 域。利用生物信息学分析基因和蛋白质序列模式对鉴别蛋白的功能具有重要的意义。通过运用多种生物学软件对得到的曼氏无针乌贼 ALSM 基因进行分析, 结果表明, 该蛋白是一种亲水性分泌蛋白。

根据兔骨骼肌 MyHC 的特异性水解位点把 ALSM 分为三个亚型: ALSM-I : Ala1159-Thr1160, ALSM-II : Glu1379-Thr1380 及 ALSM-III : Glu1098-Asp1099。目前, 已经从长枪乌贼、金乌贼及莱氏拟乌贼中获得 ALSM 的亚型: ALSM-I 和 ALSM-III, 从太平洋褶柔鱼获得的 ALSM 的亚型: ALSM-I 和 ALSM-II。通过蛋白质同源性检索发现, 曼氏无针乌贼 ALSM 与金乌贼、莱氏拟乌贼、长枪乌贼及太平洋褶柔鱼 ALSM- 亚型的同源性均较高, 与金乌贼、莱氏拟乌贼及长枪乌贼 ALSM- 的同源性亦达到 54%, 同时与太平洋褶柔鱼 ALSM- 的同源性亦为 55%。由此推测曼氏无针乌贼 ALSM 可能属于 ALSM-

亚型。通常认为, 蛋白质的空间结构及其功能比核酸的一级序列更保守, 一般认为, 如果蛋白质序列间的相似性大于 30%, 则它们可能是同源的(郝柏林等, 2000)。分析结果表明, ALSM 基因在进化上是高度保守的, 保守位点的存在对于其功能的维持起着至关重要的作用。系统进化树分析显示, 曼氏无针乌贼 ALSM 氨基酸序列首先与金乌贼、莱氏拟乌贼、长枪乌贼、太平洋褶柔鱼聚类, 再与太平洋牡蛎及其它各种鱼类聚在一起, 两分支距离较远。聚类的远近表明结构的相似程度, 而这些结构的相似性程度又进一步显示功能的相似性。

目前, 已研究物种的 ALSM 氨基酸序列均含有一个保守的 MAM 域, MAM 域中具有高度保守的 4 个 Cys 残基。MAM 域存在于许多细胞的表面蛋白中并被认为涉及细胞间粘附、蛋白相互作用和信号转导(Brady-Kalnay *et al*, 1994; Cismasiu *et al*, 2004)。MAM 域中 Cys 残基的作用可能是与结构性  $Zn^{2+}$  配对有关(Vallee *et al*, 1990), 并对分子交互作用的稳定性有重要作用(Marchand *et al*, 1996), 而 Cys 的保守性可能与二硫键的存在有关。曼氏无针乌贼 ALSM 的 MAM 域中发现 3 个保守的 Cys 残基, 可能有助于底物的结合。虾红素金属蛋白酶家族是活性中心依赖于

SmALSM	IMHYERSTFA	IDRTKPTIVP	LKKDANIGQR	IGMSQLDIIQ	LQRFYGCPEP
SeALSM-I	IMHYDRFSFA	IDRTKPTIVP	LKKDADIGQR	IGMSQLDIVQ	LQRFYGCPEP
SlALSM-I	IMHYEESFA	IDRYKPTIVP	LKKNVNIGQR	IGMSQLDIVQ	LQRFYGCPEP
LbALSM-I	IMHYEESFA	IDRFKPTIVP	LKKNVNIGQR	IGMSQLDIVQ	LQRFYGCPEP
TpALSM-I	LMHYDRYFFA	IDRSKPTLVP	LKKNVDIGQR	IGMSQLDIVQ	LQRFYGCPEP
SlALSM-III	LTHYGPTTFS	VDGQKPTIIP	LKEGVTIGQR	QGMSQLDIIQ	LQRFYGCPEP
LbALSM-III	LTHYGPTTFS	VDHIKPTIIP	LKEGVTIGQR	QGMSQLDIVQ	LQRLYKCKER
SeALSM-III	LTHYGRSFS	IDGKKPTIIP	LKEGVTIGQR	KGMSQLDIIQ	VQRFYGCPEP
TpALSM-II	LTHYGVNTFA	IDRSKPTIIP	LKKGAAGQR	KGMSQLDIIQ	LQRLYKCKER
	*	*			
SmALSM	KFVPLVTNNP	VSVNCTFDAG	YCGWKHLFME	PPAKINTWVR	WSGVTFDTDG-
SeALSM-I	KLVPVITNTP	VSVNCTFEAG	YCGWKHLFME	PPAKINTWIR	WSGGTHTDTG-
SlALSM-I	KLVPVITKTD	VSINCTFDAG	YCGWTHIFMD	PPVKQNTWIR	WYGETDTEG-
LbALSM-I	KLVPVITEDT	VSINCTFDAG	YCGWTHIFMD	PPVKKNTWIR	WSGITETEG-
TpALSM-I	KLV-TIPPKT	VSVNCTFDAD	YCGWTHVFE	PPMKKNTWIR	WSGETNPDRE
SlALSM-III	KIIPQTPGL	VTPNCTFDS	LCGWTLVDVL	PPLKNTWVR	WSGETSNYH-
LbALSM-III	KLIKQTPGM	VTPNCTFDS	LCGWTLIDIL	PPLKNTWTR	WSGETVNYH-
SeALSM-III	KLIKIQPSGM	VTPNCTFDS	LCGWTHVEIL	PPLMNTWVR	WRGETVNYH-
TpALSM-II	KVIKPEPAGS	VSPSCTFDS	LCGWTNIEVL	PPLKNTWTR	WSGESENWH-
			*		
SmALSM	TGPKNDHTIG	TFEGHYLYTS	ASGRFLSIK	IQTPELIA-G	DYCLTFWYHM
SeALSM-I	TGPKTDHTIR	TFEGHYLYSD	ASDRFLSIK	IQTPEFISPG	DHCLIFWYHM
SlALSM-I	TGPRTDHTFG	TFEGHYLYTE	ASARFLTIK	IQTPELMP-G	DHCLIFWYHM
LbALSM-I	TGPRTDHTFG	TFEGHYLYSTE	ASLRFGSVK	IQTPELKP-G	DHCLKFWYHM
TpALSM-I	TGPKTDHTFG	TFEGHYLYTE	ASDRHLAVK	IQTPELTA-G	DHCLTFWYHM
SlALSM-III	SGPVDHTFK	TFEGFYIYVN	AFKNYSVAK	IQSPEIAP-G	KYCLSFWYHM
LbALSM-III	SGPVDHTFN	TFEGFYIYVN	AFKNYSVAK	IQSPEIAP-G	DYCLSFWYHM
SeALSM-III	SGPVDHTIK	TFEGFYIYVN	AFRNYSVAK	IQSPEIAP-G	TYCLTFWYHM
TpALSM-II	SGPTDHTVK	TFEGFYIYVN	AFKNYSVAK	IQSPEITP-G	DHCFTFWYHM
SmALSM	YGKEMGTLRV	NLIEEKKKGE	LMIISGDQGN	KWKQMKISIT	ATSH-SKVEF
SeALSM-I	YGKEMGSLRV	NLVEGKQKRD	LMIISGDQGN	KWQMKISIT	ATSN-STIEF
SlALSM-I	YGQDMGSRV	NLLDGKRTNT	LMIISGDQGN	KWKEMKISLT	APSK-SKIEF
LbALSM-I	YGKDMGSRV	NLLYGKQVQS	LMFIQGNQGN	KWKEMKISLT	APSDPSKIEF
TpALSM-I	FGNDMGSLRV	KLLDGKRTKN	LMTISGDQGN	KWKQMRIFLN	APPK-SMIEF
SlALSM-III	HGRDMGSLYV	NLVKNGIARK	LTSKGDQGN	KWKQMRIFLP	APDG-GKIEF
LbALSM-III	HGRDMGSLFV	NLVKNGMAWK	LRGLKGDQGN	KWKQMKVVFVQ	APNG-GKIEF
SeALSM-III	YGRDMGSLFV	NLVKNGLTKW	IASLKGQGN	VWKQKQVXLQ	APAG-GKIEF
TpALSM-II	YGRDIGSLNV	NLVNKGLSWT	LKRLVGEQGN	EWKQMKVHIQ	APDG-GKIEF
SmALSM	ESIMGHDRS	DICIDVFLFL	SGKC		
SeALSM-I	ESVIGHSKRS	DICIDVFLFL	NRKC		
SlALSM-I	VSIIGYHEKS	DIGIDVLLL	DGKC		
LbALSM-I	VSITGYVKS	DIAIDVLLL	DGKC		

图5 ALSM 基因推导氨基酸序列中 MAM 域的同源性比对

Fig.5 Multiple sequence alignment of the MAM domain of ALSM deduced amino-acid sequence

注: \*表示保守的 Cys 残基。Se: 金乌贼(*Sepia esculenta*); Sl: 莱氏拟乌贼(*Sepioteuthis lessoniana*); Tp: 太平洋褶柔鱼(*Todarodes pacificus*); Lb: 长枪乌贼(*Loligo bleekeri*); Sm: 曼氏无针乌贼(*S. maindroni*)。

Zn<sup>2+</sup>的一类蛋白酶, 因此推测曼氏无针乌贼 ALSM 氨基酸序列中的“HEVLHSLGFYHE”序列是催化区的

修复工作逐步展开, 自然海域曼氏无针乌贼资源出现了明显的恢复迹象(吴常文等, 2010)。但曼氏无针

图6 曼氏无针乌贼 ALSM 的三级结构预测

Fig.6 Predicted three dimension structure of ALSM in *S. maindroni*

活性中心, 其中含有的 3 个组氨酸残基(以大写字母 H 代表), 可能是 Zn<sup>2+</sup>结合部位, 而谷氨酸残基(以大写字母 E 代表)则为活化位点。通过 DNAMAN 软件对曼氏无针乌贼、金乌贼、莱氏拟乌贼、长枪乌贼及太平洋褶柔鱼 ALSM 氨基酸序列中 MAM 域进行多序列比对, 结果显示, 该区域具有较高的同源性和进化上的保守性。这种结构上的保守性也表明 ALSM 的生物学功能在生物进化过程中的重要性。

自 20 世纪 70 年代末以来, 由于过度捕捞曼氏无针乌贼和对其产卵场的破坏, 舟山渔场曼氏无针乌贼资源已濒临绝迹。随着曼氏无针乌贼人工育苗技术的突破(常抗美等, 2009), 曼氏无针乌贼的增殖放流以及产卵场的保护和

乌贼在养殖条件下由一年生变为一年繁育两茬, 并且个体变小(张建设等, 2011)。本研究克隆得到了曼氏无针乌贼 ALSM 基因序列, 并对其序列及结构特征进行了分析, 对进一步深入研究 ALSM 的结构与功能以及从基因调控角度探索曼氏无针乌贼在养殖条件下生物学特性变化机理具有重要意义。

### 参 考 文 献

- 吴常文, 董智勇, 迟长凤等, 2010. 曼氏无针乌贼(*Sepiella maindroni*)繁殖习性及其产卵场修复的研究. 海洋与湖沼, 41(1): 39—46
- 张建设, 夏灵敏, 迟长凤等, 2011. 人工养殖曼氏无针乌贼(*Sepiella maindroni*)繁殖生物学特性研究. 海洋与湖沼, 42(1): 55—59
- 郝柏林, 张淑誉, 2000. 生物信息学手册. 上海: 上海科学技术出版社, 10
- 常抗美, 吴常文, 吕振明等, 2009. 曼氏无针乌贼胚胎发育与人工育苗技术的研究. 浙江海洋学院学报(自然科学版), 28(3): 257—263
- Barnes K, Ingram J, Kenny A J, 1989. Proteins of the kidney microvillar membrane. Structural and immunochemical properties of rat endopeptidase-2 and its immunohistochemical localization in tissues of rat and mouse. Biochem J, 264: 335—346
- Blader P, Rastegar S, Fischer N *et al*, 1997. Cleavage of the BMP-4 antagonist chordin by zebrafish tolloid. Science, 278(5345): 1937—1940
- Bode W, Gomis-Ruth F X, Stockler W, 1993. Astacins, serraplysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. FEBS Lett, 331: 134—140
- Bond J S, Beynon R J, 1995. The astacin family of metalloendopeptidases. Protein Sci, 4: 1247—1261
- Brady-Kalnay S M, Tonks N K, 1994. Identification of the hemophilic binding site of the receptor protein tyrosine phosphatase PTPmu. J Biol Chem, 269: 28472—28477
- Butler P E, McKay M J, Bond J S, 1987. Characterization of meprin, a membrane-bound metalloendopeptidase from mouse kidney. Biochem J, 241: 229—235
- Cismasiu V B, Denes S A, Reilander H *et al*, 2004. The MAM (meprin/A5-protein/PTPmu) domain is a homophilic binding site promoting the lateral dimerization of receptor-like protein-tyrosine phosphatase mu. J Biol Chem, 279: 26922—26931
- Hiroi J, Maruyama K, Kawazu K *et al*, 2004. Structure and developmental expression of hatching enzyme genes of the Japanese eel *Anguilla japonica*: an aspect of the evolution of fish hatching enzyme gene. Dev Genes Evol, 214(4): 176—184
- Hung C H, Huang H R, Huang C J *et al*, 1997. Purification and cloning of carp nephrosin, a secreted zinc endopeptidase of the astacin family. J Biol Chem, 272(21): 13772—13778
- Inohaya K, Yasumasu S, Araki K *et al*, 1997. Species-dependent migration of fish hatching gland cells that express astacin-like proteases in common. Dev Growth Differ, 39(2): 191—197
- Jiang W, Bond J S, 1992. Families of metalloendopeptidases and their relationships. FEBS Lett, 312(2—3): 110—114
- Kanzawa N, Tatewaki S, Watanabe R *et al*, 2005. Expression and tissue distribution of astacin-like squid metalloprotease (ALSM). Comp Biochem Physiol B, 142: 153—163
- Kawaguchi M, Yasumasu S, Shimizu A *et al*, 2005. Purification and gene cloning of *Fundulus heteroclitus* hatching enzyme. A hatching enzyme system composed of high choriolytic enzyme and low choriolytic enzyme is conserved between two different teleosts, *Fundulus heteroclitus* and medaka *Oryzias latipes*. FEBS J, 272: 4315—4326
- Kijimoto T, Watanabe M, Fujimura K *et al*, 2005. cimp1, a novel astacin family metalloproteinase gene from east African cichlids, is differentially expressed between species during growth. Molecular Biology and Evolution, 22(8): 1649—1660
- Lee K S, Yasumasu S, Nomura K *et al*, 1994. HCE, a constituent of the hatching enzymes of *Oryzias latipes* embryos, releases unique proline-rich polypeptides from its natural substrate, the hardened chorion. FEBS Lett, 339: 281—284
- Marchand P, Volkmann M, Bond J S, 1996. Cysteine mutations in the MAM domain result in monomeric meprin and alter stability and activity of the proteinase. J Biol Chem, 271: 24236—24241
- Nobuyuki Kanzawa, Takuya Ogawa, Masanori Asakura *et al*, 2008. Comparative expression and tissue distribution analyses of astacin-like squid metalloprotease in squid and cuttlefish. Zoological Society, 25: 14—21
- Piccolo S, Agius E, Lu B *et al*, 1997. Cleavage of Chordin by Xolloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. Cell, 91: 407—416
- Saitou N, Nei M, 1987. The Neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology Evolution, 4: 406—425
- Shimell M J, Ferguson E L, Childs S R *et al*, 1991. The *Drosophila* dorsal-ventral patterning gene tolloid is related to human bone morphogenetic protein 1. Cell, 67: 469—481
- Song H D, Sun X J, Deng M *et al*, 2004. Hematopoietic gene expression profile in zebrafish kidney marrow. Proc Natl Acad Sci USA, 101(46): 16240—16245
- Tsai Pei-Ling, Chen C-H, Huang C-J *et al*, 2004. Purification and Cloning of an Endogenous Protein Inhibitor of Carp Nephrosin, an Astacin Metalloproteinase. J Biol Chem,

- 279(12): 11146—11155
- Vallee B L, Auld D S, 1990. Zinc coordination, function, and structure of Zinc enzymes and other proteins. *Biochemistry*, 29: 5647—5659
- Vogt G, Stocker W, Storch V *et al*, 1989. Biosynthesis of Astacus protease, a digestive enzyme from crayfish. *Histochemistry*, 91: 373—381
- Yamaguchi T, Kido H, Kitazawa R *et al*, 1993. A membrane-bound metalloendopeptidase from rat kidney: its immunological characterization. *J Biochem (Tokyo)*, 113: 299—303
- Yan L, Pollock G H, Nagase H *et al*, 1995. A  $25.7 \times 10^3$  M(r) hydra metalloproteinase (HMP1), a member of the astacin family, localizes to the extracellular matrix of *Hydra vulgaris* in a head-specific manner and has a developmental function. *Development*, 121: 1591—1602
- Yasumasu S, Shimada H, Inohaya K *et al*, 1996. Different exon-intron organizations of the genes for two astacin-like proteases, high choriolytic enzyme (choriolysin H) and low choriolytic enzyme (choriolysin L), the constituents of the fish hatching enzyme. *Eur J Biochem*, 237(3): 752—758
- Yasumasu S, Yamada K, Akasaka K *et al*, 1992. Isolation of cDNAs for LCE and HCE, two constituent proteases of the hatching enzyme of *Oryzias latipes*, and concurrent expression of their mRNAs during development. *Dev Biol*, 153: 250—258
- Yokozawa Y, Tamai H, Tatewaki S *et al*, 2002. Cloning and biochemical characterization of astacinlike squid metalloprotease. *J Biochem (Tokyo)*, 132(5): 751—758

## MOLECULAR CLONING AND SEQUENCE ANALYSIS OF ALSM cDNA FROM *SEPIELLA MAINDRONI*

ZHOU Chao, GUO Bao-Ying, LIU Hui-Hui, WU Chang-Wen

(Marine Science College of Zhejiang Ocean University, National Engineering Research Center of Marine Facilities Aquaculture, Zhoushan, 316004)

**Abstract** The ALSM (Astacin-like squid metalloprotease) gene of *Sepiella maindroni* was cloned from liver by RT-PCR and RACE technique. Sequence analysis revealed a 1418bp full-length cDNA sequence containing 53bp 5'-untranslated region, 75bp 3'-untranslated region and 1290bp open reading frame (ORF) and this sequence was encoded by 429 amino acids. The molecular weight of deduced protein was 48916.31Da and its pI was 7.978. The deduced amino acid sequence aligned with those of ALSM genes from different species showed high degree of sequence homology. The phylogenetic analysis tree showed that ALSMs are classified into two clades: ALSM-I forms one clade, and ALSM-II and -III form the other. The ALSM protein of *S. maindroni* firstly clustered with ALSM-I of golden cuttlefish, bigfin reef squid, spear squid and Japanese common squid, then with ALSM-II and -III of them. The bioinformatics analysis revealed that the predicted protein was a hydrophilic secretion protein. The results are very important for further study of on structures and functions of ALSM and exploration of the mechanism of biological characteristics changes of *S. maindroni* in breeding conditions.

**Key words** *Sepiella maindroni*, Astacin, ALSM, Metalloprotease, Cloning, Sequence analysis