

DOI: 10.19663/j.issn2095-9869.20190122002

<http://www.yykxjz.cn/>

王倩, 张雪, 陈亚东, 沙珍霞. 半滑舌鳎甘露糖结合凝集素相关丝氨酸蛋白酶 1 基因的克隆及表达分析. 渔业科学进展, 2020, 41(3): 49–59

Wang Q, Zhang X, Chen YD, Sha ZX. Cloning and expression analysis of the MASP1 gene from the half-smooth tongue sole (*Cynoglossus semilaevis*). Progress in Fishery Sciences, 2020, 41(3): 49–59

半滑舌鳎甘露糖结合凝集素相关丝氨酸蛋白酶 1 基因的克隆及表达分析^{*}

王倩¹ 张雪¹ 陈亚东^{2,3} 沙珍霞^{1,2①}

(1. 青岛大学生命科学学院 青岛 266071;

2. 青岛海洋科学与技术试点国家实验室海洋渔业科学与食物产出过程功能实验室 青岛 266071;

3. 中国水产科学研究院黄海水产研究所 农业农村部海洋渔业可持续发展重点实验室 青岛 266071)

摘要 甘露糖结合凝集素相关丝氨酸蛋白酶 1 (Mannan-binding lectin associated serine protease 1, MASP1) 是补体凝集素途径中起重要作用的激活蛋白。本研究以半滑舌鳎 (*Cynoglossus semilaevis*) 为研究对象, 应用 RACE 技术和实时荧光定量 qRT-PCR 技术对半滑舌鳎 MASP1 基因 (*Cynoglossus semilaevis MASP1*, *CsMASP1*) 进行了克隆和表达模式分析。结果显示, *CsMASP1* 基因 cDNA 序列全长为 2507 bp, 5'非编码区长 82 bp, 3'非编码区长 142 bp, 开放阅读框长 2283 bp, 共编码 760 个氨基酸; *CsMASP1* 基因包含 13 个外显子和 12 个内含子, 与多数已知鱼类的 *MASP1* 基因结构一致; SMART 分析显示, *CsMASP1* 包含 6 个结构域, 与哺乳动物、鸟类、其他鱼类的结构域一致; 同源比对发现, *CsMASP1* 和鱼类的相似度较高, 与金目鲈 (*Lates calcarifer*) 的相似性最高, 为 76%。*CsMASP1* 基因在 11 种健康组织(血液、脑、鳃、性腺、心脏、头肾、肠、肝、皮肤、脾和后肾) 中均有表达, 其中, 在肝、脾和头肾中表达量较高; 鳗弧菌 (*Vibrio anguillarum*) 感染后, *CsMASP1* 基因在 6 种免疫组织(血液、鳃、头肾、肠、肝和脾) 中呈现不同的表达模式, 在 6 种免疫组织中呈现明显的上调表达, 肝的表达峰值出现在感染后 24 h; 脾和鳃的表达峰值出现在感染后 6 h; 肠、头肾和血液的表达峰值均出现在感染后 12 h; 随着病原菌感染时间增加, 基因表达量逐渐降低并恢复至正常水平。研究表明, *CsMASP1* 基因参与了半滑舌鳎的免疫应答过程, 本结果为半滑舌鳎的免疫机理研究奠定了基础。

关键词 半滑舌鳎; *MASP1* 基因; 基因克隆; 基因表达; 免疫应答

中图分类号 S971.4 **文献标识码** A **文章编号** 2095-9869(2020)03-0049-11

甘露糖结合凝集素相关丝氨酸蛋白酶 1 (MASP1) 属于丝氨酸蛋白酶 (Serine proteases, SP) 超家族成员, 包括 MASP1 (Takahashi *et al*, 2013)、MASP2 (Swierzko *et al*, 2014) 和 MASP3 (Dahl *et al*, 2001)。MASP1 是补体凝集素途径中起重要作用的激活蛋白, MASP1 可

被甘露糖结合凝集素 (Mannan-binding lectin, MBL) 激活以响应病原体感染, MBL-MASP 复合物共同激活凝集素途径 (Sunyer *et al*, 1998; Schwaebel *et al*, 2002)。

目前, 关于 MASP1 的研究相对较少, 2000 年首次在人中克隆鉴定出 *MASP1* (Chen *et al*, 2000), 随后

* 国家自然科学基金项目(31572644)资助 [This work was supported by National Natural Science Foundation of China (31572644)]. 王倩, E-mail: qianwangqd@163.com

① 通讯作者: 沙珍霞, 教授, E-mail: shazhenxia@163.com

收稿日期: 2019-01-22, 收修改稿日期: 2019-03-09

构建了人的 MASP1 N 端片段原核表达载体并在大肠杆菌中进行了表达(Cai et al, 2008)。迄今仅在文昌鱼(*Branchiostoma belcheri*)(Endo et al, 2003)、七鳃鳗(*Lampetra fluviatilis*)(Endo et al, 2003)、草鱼(*Ctenopharyngodon idella*)(Dang et al, 2017)和菊海鞘(*Botryllus schlosseri*)(Nicola et al, 2017)等水生生物中开展了 MASP1 的克隆表达及免疫应答相关研究。硬骨鱼 MASP1 的研究还极为有限。

半滑舌鳎(*Cynoglossus semilaevis*)是我国重要的海水养殖鱼类。近年来,随着水产养殖行业的快速发展,病害问题特别是弧菌病对半滑舌鳎养殖业造成严重的经济损失。有关半滑舌鳎免疫机制的研究和有效的病害防控手段还有待深入研究(Chen et al, 2015)。本研究通过 PCR 技术克隆了 *CsMASP1* 基因的 cDNA,并分析了其基因结构及其进化关系,还研究了 *CsMASP1* 基因在半滑舌鳎 11 种健康组织中的表达特征,及鳗弧菌感染后该基因在 6 种免疫组织中的时空表达规律,为进一步深入了解 *CsMASP1* 基因在半滑舌鳎免疫防御中的作用奠定基础。

1 材料与方法

1.1 实验材料

半滑舌鳎购自山东海阳黄海水产有限公司,鱼龄为 2 龄左右,平均体重为(400±5) g,平均体长为(40.3±0.1) cm。健康实验鱼在 24℃ 水族箱中暂养 7 d,每天换新鲜灭菌海水,尽量消除环境胁迫。鳗弧菌为本实验室保存菌种,经 16S rRNA 测序进行菌种鉴定。

1.2 样品处理和采集

随机选取 3 条健康的半滑舌鳎,分别采集每条鱼的血液、脑、鳃、性腺、心脏、头肾、肠、肝、皮肤、

脾和后肾共 11 种组织,并立即投入液氮中冷冻,然后转移至-80℃冰箱保存,用于总 RNA 提取。鳗弧菌感染参照(Sha et al, 2017)实验方法略作修改,采用腹腔注射方法对实验组进行鳗弧菌感染,感染剂量(半致死剂量)为 3.18×10^5 CFU/g 鱼体,对照组注射相同体积的无菌 PBS 溶液。在注射后 0、6、12、24、48 和 72 h 共 6 个时间点,每个时间点取 3 条鱼,分别收集实验组和对照组的血液、鳃、头肾、肠、肝脏和脾脏 6 种组织,保存方式同上所述。

1.3 RNA 提取和 cDNA 制备

采用液氮研磨收集的组织,按照 RNA 提取试剂盒(天根,北京)的说明书进行组织总 RNA 提取。提取的 RNA 用琼脂糖凝胶电泳(1.00%)检测其完整性。用 Agilent 2100 biological analyzer (Applied Biosystems, 美国)检测其纯度和浓度。按照 PrimeScript II 1st Strand cDNA Synthesis Kit (TaKaRa, 大连)试剂盒反转录成 cDNA, -20℃ 冰箱保存备用。

1.4 *CsMASP1* 基因的 cDNA 序列的扩增

通过对半滑舌鳎转录组和基因组数据进行生物信息学分析,筛选了 *CsMASP1* 基因的部分 cDNA 序列,并对该序列设计特异性引物 P1 和 P2(表 1),以半滑舌鳎 11 个组织混合 cDNA 为模板,进行 PCR 扩增。PCR 反应体系(20 μl): Trans Taq® HiFi Polymerase (5 U/μl) 0.2 μl, 10×Trans Taq HiFi Buffer II 2 μl, 2.5 mmol/L dNTPs 1.6 μl, P1 1 μl, P2 1 μl, cDNA 模板 2 μl, ddH₂O 12.2 μl。反应程序: 95℃ 5 min; 95℃ 30 s, 58℃ 30 s, 72℃ 30 s, 38 个循环; 72℃ 10 min; 4℃ 保存。1.00% 琼脂糖凝胶电泳检测后,用 DNA 胶回收试剂盒(诺唯赞,南京)纯化 PCR 扩增特异性产物,连接到 pEASY-T1(全式金,北京)载体上,重组载体转化到 *TransT1* 感受态细胞(全式金,北京)中,挑取阳

表 1 半滑舌鳎 *MASP1* 基因克隆和表达分析所用引物
Tab.1 PCR primers used for cloning and expression analysis of *MASP1*

引物 Primer	序列 Sequence (5'~3')	用途 Application
P1	TTCTTCTTCATCTACTTT	PCR 扩增
P2	GATATTTTTCAACTAACAG	PCR 扩增
P3	GTAGTGAGCCATGAAACCTGAATGTC	5'RACE 扩增
P4	GTGTGTGAAGTACAACCTGACTCGG	5'RACE 扩增
P5	TAGAGAAACACCTCCAGATTTC	3'RACE 扩增
P6	CTGTGCTGGCTTCTTGAGGG	3'RACE 扩增
P7	AACGTGGCCTGCCGTGAGG	qRT-PCR 扩增
P8	TGGAGCCAAGCACCAAC	qRT-PCR 扩增
Cs18s-F	GGTCTGTGATGCCCTAGATGTC	qRT-PCR 内参引物
Cs18s-R	AGTGGGGTTCAGCGGGTTAC	qRT-PCR 内参引物

性克隆进行菌液 PCR 验证后测序(上海派森诺生物科技有限公司)。在克隆获得的序列 5'端和 3'端设计 RACE (Rapid-Amplification of cDNA Ends)引物 P3 和 P5(表 1)和巢式特异性引物 P4 和 P6(表 1), 使用 SMARTer™ RACE cDNA Amplification Kit(TaKaRa, 大连)获得该基因的 5'和 3'端序列, 将克隆的序列进行拼接获得 *CsMASP1* 基因的 cDNA 全长序列。

1.5 *CsMASP1* 基因序列分析

将 *CsMASP1* 的全长 cDNA 序列提交到 NCBI 网站(<http://blast.ncbi.nlm.nih.gov/blast.cgi>), 利用 BLAST 程序进行核苷酸同源性比对, 利用 ORF Finder (<http://blast.ncbi.nlm.nih.gov/gorf/gorf.html>)查找开放阅读框, 并推导其相应的氨基酸序列; 采用 SMART(Simple Modular Architecture Research Tool)在线软件分析 *CsMASP1* 蛋白的结构域, 利用 ExPASy 软件(<http://www.expasy.org>)预测其相对分子质量和等电点; 用 SignalP 4.0 Server (<http://www.cbs.dtu.dk/services/SignalP/>)软件分析 *CsMASP1* 蛋白的信号肽。应用 Splign(<http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/>)在线软件确定 *CsMASP1* 基因的内含子与外显子的数目, 分析不同物种 *MASP1* 的基因结构以及其所在染色体的位置。

1.6 MASP1 同源比对系统进化树构建

从 GenBank 中查询不同物种 *MASP1* 基因的氨基酸序列, 采用 DNAMAN 软件进行不同物种的 *MASP1* 蛋白的氨基酸序列多重比对, 用 MEGA 6.0 软件中的 Neighbor-joining 法构建系统进化树。

1.7 *CsMASP1* 基因表达的实时定量分析

根据 *CsMASP1* 基因设计扩增引物 P7 和 P8(表 1), 以半滑舌鳎 18S rRNA 基因为内参设计特异引物 Cs18s-F 和 Cs18s-R(表 1)。基因实时定量表达参照已报道的方法进行(Lu *et al.*, 2013; Yu *et al.*, 2017)。使用 ChamQ™SYBR® Color qPCR Master Mix 试剂(诺唯赞, 南京), 按照说明书于 Applied Biosystems 7500 Real Time PCR 仪上进行该基因的实时定量分析(qRT-PCR)。反应体系为 20 μl: 10 μl 2×ChamQ SYBR Color qPCR Master Mix, 1 μl 模板 cDNA, 0.4 μl 引物 P7, 0.4 μl 引物 P8, 0.4 μl 50×ROX Reference Dye 2, 7.8 μl ddH₂O。反应程序: 预变性 95°C 3 min; 95°C 10 s, 60°C 30 s, 40 个循环; 熔解曲线 95°C 15 s; 60°C 1 min; 95°C 15 s。每个样品设置 3 个技术重复, 采用 $2^{-\Delta Ct}$ 方法计算目的基因的相对表达量, *Ct* 值取 3 个平行样品的平均值。qRT-PCR 检测结果使用 SPSS

19.0 统计软件进行单因素方差分析(One-way ANOVA)的 Duncan 法对多组样本平均值进行两两比较分析, *P*<0.05 时, 认为存在显著性差异。

2 结果

2.1 *CsMASP1* 基因序列分析和结构特征

克隆获得的 *CsMASP1* 基因 cDNA 序全长为 2507 bp (登录号: XM_008309208.2), 5'非编码区(Untranslated Region, UTR, 5' UTR)长 82 bp, 3' UTR 长 142 bp, 开放阅读框(Open reading frame, ORF)长 2283 bp(图 1)。推测 *CsMASP1* 编码氨基酸为 760 个, 预测其分子量为 84.95 kDa, 理论等电点(pI)为 5.33。SignalP 4.0 Server 分析显示, 此氨基酸序列无信号肽。SMART 软件分析显示, *CsMASP1* 蛋白含 6 个结构域: 分别是 2 个类 Clr/Cls 蛋白(like Clr/Cls Protein domain, CUB)结构域(47~168 aa, 215~327 aa), 1 个表皮细胞生长因子样(Calcium-binding epidermal growth factor-like domain, EGF-CA)结构域(169~212 aa), 2 个补体控制蛋白区(Complement control protein domain, CCP)结构域(331~392 aa, 397~461 aa)和 1 个丝氨酸蛋白激酶区(Serine protease domain, SP)Tryp-SPc 结构域(478~752 aa)(图 2)。

利用 Splign 软件分析了 7 个物种的 *MASP1* 基因结构并进行了比较(图 3), 它们是半滑舌鳎(XM_008309208.2)、剑尾鱼(*Xiphophorus maculatus*)(XM_023351221.1)、大黄鱼(*Larimichthys crocea*)(XM_027280128.1)、斑马鱼(*Danio rerio*)(XM_001341900.6)、马(*Equus caballus*) (XM_001499629.5)、人(*Homo sapiens*)(NM_001879.5)和中国仓鼠(*Cricetulus griseus*)(XM_003495604.2)。结果显示, *CsMASP1* 含 13 个外显子, 位于 4 号染色体上; 剑尾鱼 *MASP1* 含 13 个外显子, 位于 18 号染色体上; 大黄鱼 *MASP1* 含 13 个外显子, 位于 7 号染色体上; 斑马鱼 *MASP1* 含 12 个外显子, 位于 15 号染色体上; 马 *MASP1* 含 16 个外显子, 位于 19 号染色体上; 人 *MASP1* 含 16 个外显子, 位于 3 号染色体上; 中国仓鼠 *MASP1* 含 15 个外显子, 位于 4 号染色体上。半滑舌鳎 *MASP1* 基因结构与多数已知鱼类(剑尾鱼和大黄鱼)一致, 与哺乳类(人、马和中国仓鼠)在外显子数目上有差异, 说明该基因在鱼类保守。

2.2 半滑舌鳎与其他物种 *MASP1* 同源比对和系统进化分析

通过 BLAST 比对, 将 *CsMASP1* 氨基酸序列

aacacacacacactctgagcagagtccggagaatggagagacggagacaactggtagagcagcagacaaaatggtag	82
ATGAAAGAAGAAGAAGTGAAGAGAATGAAAGAATCATCATCAGTGATTCTCATCAACTGTGTTGAAGTCAGTCTGCTTCACTGTGAGAAGAAAAGGATGATCAGGTGCCTCCCTTCCTG	202
M K K E E V K R M K E S S S V I L H Q L S L K S V F H P E K K R M I R C L L F L	40
CTTCTCCTGACGTTGGCAGTGCAGCTCAGGTTCTGTCTTGTACAATGATGGCAGCATCAAATGCCAAATTCCAGAGCCATATCCAATGAGACAGACAATGCGCTGGAATGTC	322
L L L T L A T D A Q V L S L S Q M Y G S I K S P N F P E P Y P N E T E L R W N V	80
AGCATTCTGAAGGATTCAGCTCAAGTTGACTTCACACACTTGACCTAGAGCCGCTCACCTGTGAGTACGACTTGTCAAGGTGGAGGCAGAGAGGGAGGCTGGCATGGT	442
S I P E G F R V K L Y F T H F D L E P S Y L C E Y D F V K V E A E R E E L A L F	120
TGCGGGAAAGGAGGAGGAGGAGACACTGAAGCTGTGCGGCTCAGCAGGTCATCACCTCACCCAGAAAACCCCTCAGCCTGCTCTCTCTGACTCTCCAATGAGCAGAGACATTCAAGGT	562
C G K E E S D T E A V P A Q Q V I T S P R N S L S L L F S S D F S N E Q R H S G	160
TTCATGGCTCACTACAGTGCAGAGGATGTTGATGAATGCACTGATTCCAGTGTGAGATCTGTTGTGATCACTCTGTACAACCTTCAATTGGAGGTTACTGTGCTCTGCTGCTAT	682
F M A H Y S A E D V D E C S D S S D E D L F C D H F C H N F I G G Y Y C S C R Y	200
GGTACCTGCTGCATACAGACAATCGAACCTGCAGAGTGGAGTCAGGATGCTGAGGATCTGTTGATTCAGCTGAGTGTGATTTCCGGCTCCGTATCCAAAAGCTCT	802
G Y L L H T D N R T C R V E C S D A V F R E R S G I L S S V D F P A P Y P K S S	240
GAECTGTTGATCGTATTGAGGTGGAGGCTGGATTAAGATCCGCCTCCAGTTGACCTGACCTCAACGTCGAGGATCATCTGACCTGAGGTGCTCTACGACTACATCAAGATCACA	922
D C L Y R I E V E P G F K I R L Q F D P S F N V E D H P D V R C P Y D Y I K I T	280
GGGGCAGCACTGACCAGGCCCTTTGTGGAGATAATCCCAGGAGTCATTGAGGTGGACAGTAACGTGGCTCTGCTCTCCACAGTGACAACACTCAGGAGAAAACCTGGATGG	1042
G G S T E Q G P F C G D K S P G V I E V D S N V A S V V F H S D N S G E N L G W	320
AGGTTGACGTACACAGCTATAGGAAGTAAATGCGTAGTACCTGAAATCCAGCCCCACGCTCTGATGAATCTGTCACATCAGAATATTCTTTAAAGACCAACTTGTCTCACCTGTGAG	1162
R L T Y T A I G S K C V V P E I Q P H A L M N P V Q S E Y S F K D H I V F T C E	360
CCTGGATTTGCACTGCAAGGATGCCAGTATCTGGACCATTTCCAGATCGACTGTGTTCTAATGGATCTGACTAGTCGTCCTCTGACATGTCACATGGTGGATTGTGGAGTC	1282
P G F R L L K D G Q Y L D H F Q I D C G S N G S W T S R P P Q C H M V D C G S P	400
AAGGGTGGGACCTGGCAGAGGTGGTTTGGAACCCAGACAACACCCTCGATTGGATCGTCAGCTGTAGAGGGGACATGATGGAGGCCAAGCAGCAACACGTCT	1402
K V V D L A E V V F G N H D N T T R F G S S V T Y V C R G D M M E P S T S N T S	440
TACACCTGTGCCAGAAATGGAGAATGGACAAATGAGACAGGAGTCAAACTGCCCGCTGCTCCAGCTGGACGTCCTCGCTGACCTTCCCCCTCAGATGAAGCGTATCGT	1522
Y T C G Q N G E W T N V E A G V K L P R C L P A C G R P S R T F P P Q M K R I V	480
GGAGGGCGGACTCGCAGAGGCCCTCCCATGGCAGGCTCGTCAGTGTGAGGACAGTCAGAGTTCCAGAGAACGACACTGGTTGGTCTGGAGCGCTGCTGAGTCTGG	1642
G G R T A E P G L F P W Q V L L S V E D Q S R V P E D H W F G S G A L L S E F W	520
GTCTCACGGCAGCCCCAGTTCTCAGGTCAGAGGAGAGATGCTAGTGTGGTCTGTGGCACAGAACACGTGAAGGTCTCTGGTCTCCATGACCCAACAACAAACATAAGGCT	1762
V L T A A H V L R S Q R R D A S V V P V A P E H V K V F L G L H D A N N K H K A	560
GACAGTCGCTCAGTGGACAAGGTGGCTCCATCCACACTCCAACCAACAACATAATAATGACATGCCCTCATAGGTGACTGAGGAGCTTAAACAGTTACATCCGACCT	1882
D S R S V D K V V L H P F Q P N N Y N N D I A L I R L T E A V E F N S Y I R P	600
GTGTGTTGCCCTCCCCATGATGAGCTAACTCTCGTCACTCCACCTAACTCTGGAGTTGGCTGGTGGGGATCTCAACCCCTAACCTCCATTTCTCTCATCCCCA	2002
V C L P S P N D E P N S S S L L P N S L G V V A G W G I S N P N S T F S S S S P	640
CCGACCTCCGTACCCACGGGAACCTTGACCCCGACTGACCTCTGACATCCTACAGTACGTAAGCTACCGGTGGTTCTCAGGGAGTGTGGCCAGCTACTCCTCGCTCCGTC	2122
P T S V T T A N F D P G L T S D I L Q Y V K L P V V S Q G E C R A S Y S S R S V	680
AGTACAAACATCACAGACAACATGTTCTGCTGGCTCTTGAAGGGGGGGGACACCTGCTGGTGACAGCGGAGGGCGTTGTGACAGAGGACCACTCAGCCGGAGGTGGGT	2242
S Y N I T D N M F C A G F F E G G R D T C L G D S G G G A F V T E D Q L T R R W V	720
GTGTCAGGGCTTGTGTCATGGGGAGGCCAGAAGAGTGTGGACTGGCGGGGTATGGTGTACACTCGAGTGGCATATGTGAGTGGATAGAGAAACACTCCAGATTTCCCC	2362
V S G L V S W G G P E E C G S R R V Y G V Y T R V A S Y V E W I E K H L Q I F P	760
TGAtggtagactctgttaacacgcggaaaaagggtgtccaaactgtgagatttgacagttacaatgtataatgtataacacatcacacgaaatggaaaaaaa	2482
*	760
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	2507

图 1 半滑舌鳎 *MASP1* cDNA 序列和推测的氨基酸序列Fig.1 Nucleotide and deduced amino acid sequences of *Cynoglossus semilaevis MASP1*

推导的氨基酸序列显示在核苷酸序列下方, 用大写字母表示。方框所示为起始密码子(ATG)、星号(*)所示为终止密码子(TGA); 实线(____)所示为 CUB 结构域; 圆点线(....)所示为 EGF-CA 结构域; 短划线(—)所示为 CCP 结构域; 双划线(==)所示为 Tryp-SPc 结构域; 波浪线(~~)所示为 poly A 结构

Translated amino acid sequence was shown under the nucleotide sequence as uppercase. The initiation codon and termination codon were marked by a box and an asterisk, respectively; CUB domain was shown in solid line(____); EGF-CA domain was shown in dot line(....); CCP domain was shown in dashed line(—); Tryp-SPc domain was shown in double solid line(==); poly A was shown in wavy line(~~~)

(XP_008307430.1) 与金目鲈 (*Lates calcarifer*) (XP_018560455.1)、牙鲆 (*Paralichthys olivaceus*) (XP_019966501.1)、黄鳍 (*Monopterus albus*) (XP_020477488.1)、伯氏朴丽鱼 (*Haplochromis burtoni*) (XP_014184888.1)、非洲爪蟾 (*Xenopus laevis*) (AAI70221.1)、海龟 (*Chelonia mydas*) (EMP37646.1)、红喉潜鸟 (*Gavia stellata*) (XP_009815055.1)、皇帝企鹅 (*Aptenodytes forsteri*) (KFM09600.1)、白鹭 (*Egretta garzetta*) (XP_009633907.1)、西伯利亚虎 (*Panthera tigris altaica*) (XP_015393589.1)、羊驼 (*Vicugna pacos*) (XP_006201039.1)、马 (XP_005601936.1)、中国仓鼠 (XP_007606939.1)、人 (XP_016862358.1) 的 MASP1 氨基酸序列进行同源比对, 相似性范围为 76%~42%。其中, CsMASP1 与金目鲈的相似性最高, 为 76%; 而与中国仓鼠的相似性最低, 为 42%。将半滑舌鳎和其他 13 个物种 MASP1

的氨基酸序列进行多重比对, 可以看出不同物种 MASP1 存在高度保守的氨基酸残基, 如图 4 所示, 保守氨基酸出现在丝氨酸蛋白酶活性中心处: 组氨酸 His⁵²⁶、天冬氨酸 Asp⁵⁸² 和丝氨酸 Ser⁷⁰⁵, 这 3 个氨基酸的保守定位对于维持丝氨酸蛋白酶的活性是十分重要的。

利用 MEGA 6.0 软件对上述各个物种的 MASP1 氨基酸序列进行比对构建了系统进化树(图 5)。结果显示, 半滑舌鳎与伯氏朴丽鱼、金目鲈和黄鳍等鱼类的 MASP1 聚为一支。两栖类海龟和非洲爪蟾、鸟类白鹭、皇帝企鹅和红喉潜鸟, 哺乳类马、西伯利亚虎、羊驼、人和中国仓鼠等动物的 MASP1 聚为一支。但两栖类、鸟类和哺乳类又按亲缘关系远近聚成相应的亚类。MASP1 在进化上与伯氏朴丽鱼的亲缘关系最近, 与哺乳类、爬行类和鸟类等动物的 MASP1 的亲

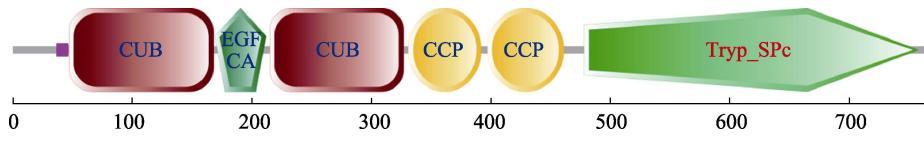


图 2 半滑舌鳎 MASP1 蛋白结构域预测
Fig.2 Domain structure prediction of *Cynoglossus semilaevis* MASP1

SMART 软件在线预测 CsMASP1 蛋白的结构域, CsMASP1 包含 6 个结构域, 从 C 端到 N 端分别是 CUB 结构域(47~168 aa), EGF-CA 结构域(169~212aa), CUB 结构域(215~327 aa), CCP 结构域(331~392 aa), CCP 结构域(397~461 aa), Tryp-SPc 结构域(478~752 aa)

SMART software predicts the domain of CsMASP1 protein online. CsMASP1 contains six domains:
CUB domain (47~168 aa), EGF-CA domain (169~212 aa), CUB domain (215~327 aa), CCP domain (331~392 aa),
CCP domain (397~461 aa), Tryp-SPc domain (478~752 aa)



图 3 不同物种的 MASP1 基因结构比较
Fig.3 Comparison of MASP1 gene structures of different species

7 个物种的 MASP1 基因结构图, 各个物种的 MASP1 基因序列号与所在染色体基因组序列号分别为: 人 NM_001879.5, NG_029440.1; 马 XM_001499629.5, NC_009162.3; 中国仓鼠 XM_003495604.2, NW_003613596.1; 剑尾鱼 XM_023351221.1, NC_036460.1; 大黄鱼 XM_027280128.1, NC_040017.1; 半滑舌鳎 XM_008309208.2, NC_024310.1; 斑马鱼 XM_001341900.6, NC_007126.7

The sequence numbers of the MASP1 gene and the genome numbers of the chromosomes in seven species were as follows: *Homo sapiens* NM_001879.5, NG_029440.1; *Equus caballus* XM_001499629.5, NC_009162.3; *Cricetulus griseus* XM_003495604.2, NW_003613596.1; *Xiphophorus maculatus* XM_023351221.1, NC_036460.1; *Larimichthys crocea* XM_027280128.1, NC_040017.1; *Cynoglossus semilaevis* XM_008309208.2, NC_024310.1; *Danio rerio* XM_001341900.6, NC_007126.7

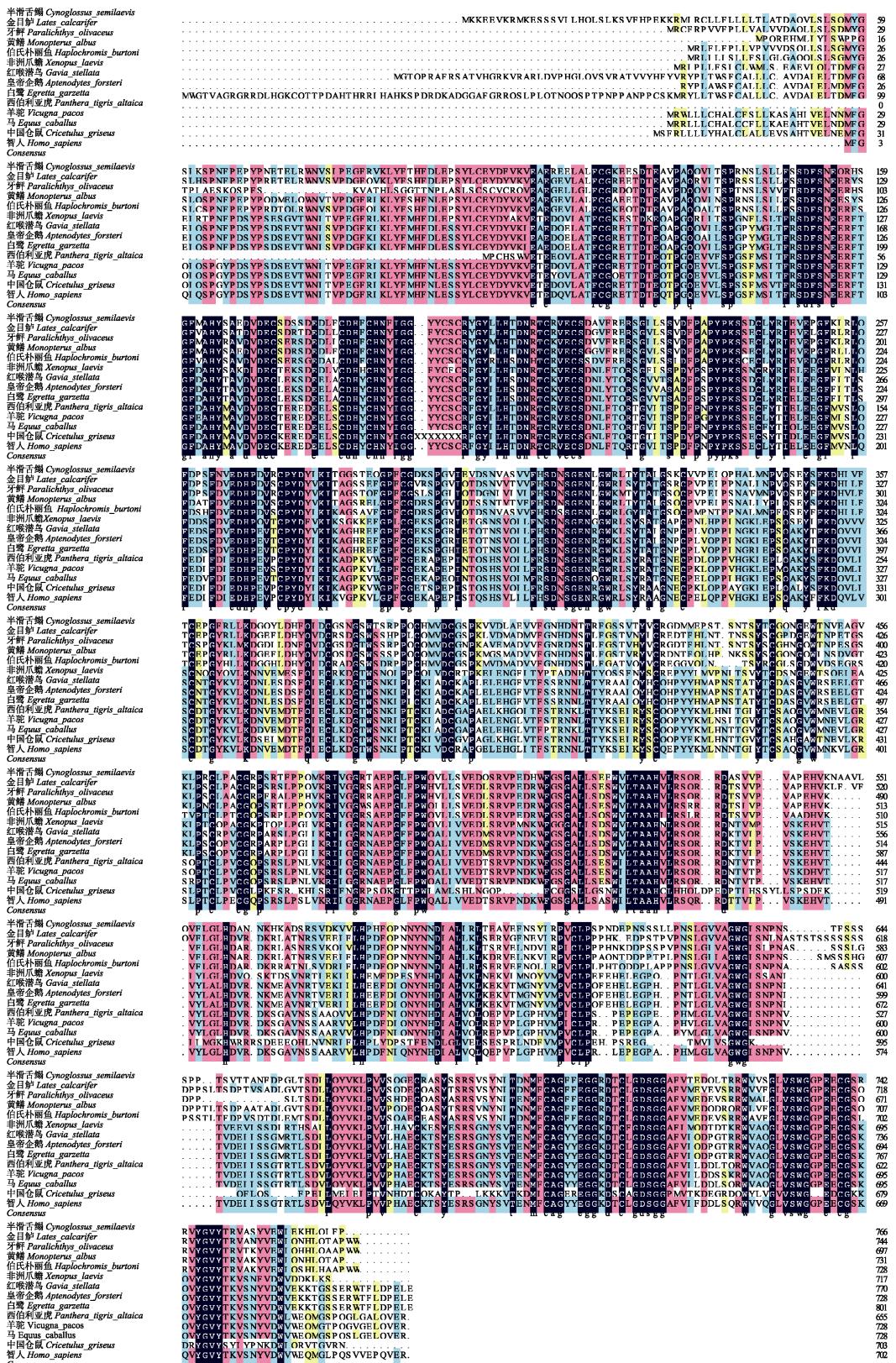


图 4 半滑舌鳎和其他物种 MASP1 氨基酸序列的多重比对

Fig.4 Multiple alignment of the deduced amino acids of MASP1 among half-smooth tongue sole and other different species

MASP1 氨基酸序列比对，丝氨酸蛋白酶结构域中保守氨基酸用黑框标出：组氨酸 His⁵²⁶、天冬氨酸 Asp⁵⁸² 和丝氨酸 Ser⁷⁰⁵

Amino acid sequence alignment of MASP1, conservative amino acids in serine protease domain were marked by the black box: His⁵²⁶, Asp⁵⁸² and Ser⁷⁰⁵

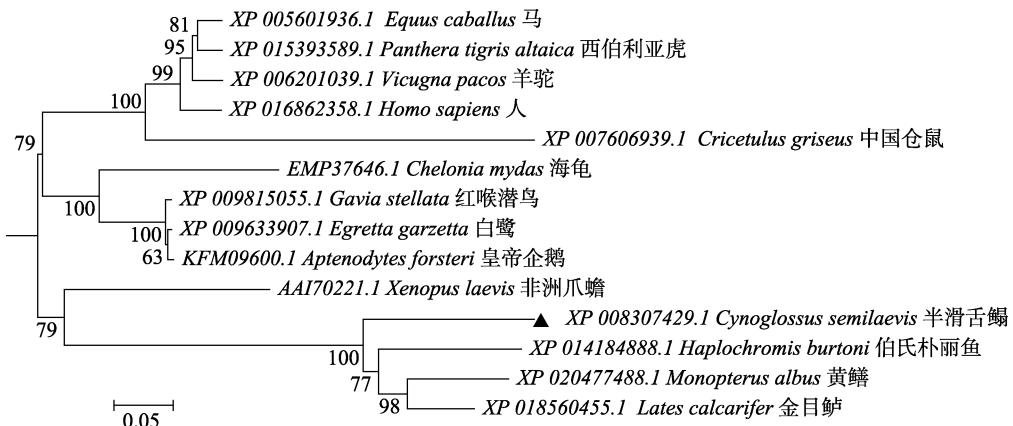


图 5 MEGA 6.0 软件构建的基于 MASP1 氨基酸序列的系统进化树
Fig.5 Phylogenetic tree based on MASP1 amino acid sequence by MEGA 6.0

缘关系较远。从进化的结果看, 几个关键节点的支持率较高, 均在 90%以上。例如: 金目鲈和黄鳝, 半滑舌鳎和伯氏朴丽鱼这几个物种聚类为一个分支, 支持率分别为 98% 和 100%。

2.3 CsMASP1 基因在健康组织中的表达分析

利用 qRT-PCR 分析 *CsMASP1* 基因在 11 种健康组织中的表达特征。结果显示, *CsMASP1* 基因在血液、脑、鳃、性腺、心脏、头肾、肠、肝、皮肤、脾和后肾中均有表达; 其中, 肝中相对表达量最高(185.4), 其次为脾(138.1)和头肾(76.2), 在脑、心脏、血液、肠、后肾、性腺、皮肤和鳃中表达量较低且表达量依次降低(图 6)。

2.4 鳗弧菌感染后 *CsMASP1* 基因在免疫组织中的表达分析

鳗弧菌感染半滑舌鳎后可诱导 *CsMASP1* 在免疫组织中的表达变化, 在感染 0、6、12、24、48、72 h 后, *CsMASP1* 在 PBS 对照组和感染组半滑舌鳎的血液、鳃、头肾、肠、肝和脾 6 种免疫组织中的表达水平如图 7 所示。*CsMASP1* 基因表达总体趋势表现为先升高后降低。在脾和鳃中, *CsMASP1* 基因的表达呈现相同的上调趋势, 其表达峰值出现在鳗弧菌感染后 6 h, 在脾中为对照组表达量的 15.4 倍; 在鳃中则是对照组的 4.6 倍; 在肝中 *CsMASP1* 基因表达上调, 最大表达量出现在感染后 24 h, 是对照组的 12.4 倍; 在肠中 *CsMASP1* 基因表达上调, 最大表达量出现在感染后 12 h, 是对照组的 7.7 倍, 72 h 时依然保持较高的表达水平; 在头肾和血液中, *CsMASP1* 基因的表达峰值均出现在鳗弧菌感染后 12 h, 在头肾中为对照

组表达量的 9.7 倍; 在血液中则是对照组的 6.8 倍; 随后基因表达量逐渐下降至略低于正常水平值。

3 讨论

3.1 *CsMASP1* 基因 cDNA 的克隆及序列分析

甘露聚糖结合凝集素相关丝氨酸蛋白酶是机体

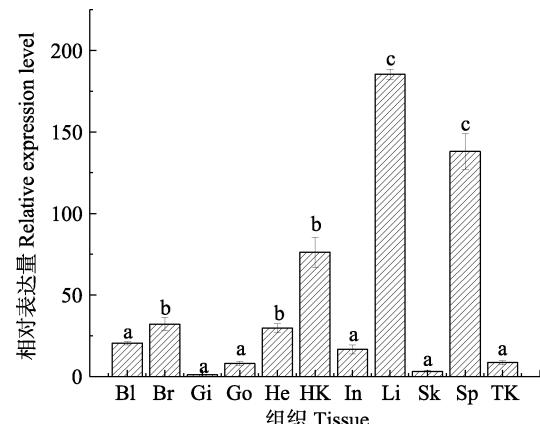


图 6 *CsMASP1* 基因在半滑舌鳎健康组织中的相对表达量

Fig.6 The relative expression of *CsMASP1* gene in normal tissues of *C. semilaevis*

Bl: 血液; Br: 脑; Gi: 鳃; Go: 性腺; He: 心脏; HK: 头肾; In: 肠; Li: 肝; Sk: 皮肤; Sp: 脾; TK: 后肾。字母“a, b, c”代表 SPSS 多重分析的不同分组, 有相同字母表示差异不显著($P>0.05$), 无相同字母表示差异显著($P<0.05$); 下同

Bl: Blood; Br: Brain; Gi: Gills; Go: Gonad; He: Heart; Hk: Head kidney; In: Intestine; Li: Liver; Sk: Skin; Sp: Spleen; Tk: Trunk kidney. The letters of ‘a, b, and c’ indicated the Duncan grouping in SPSS 19.0 software. The same letters indicate no significant difference ($P>0.05$), the different letters indicate significant difference ($P<0.05$); the same applies below

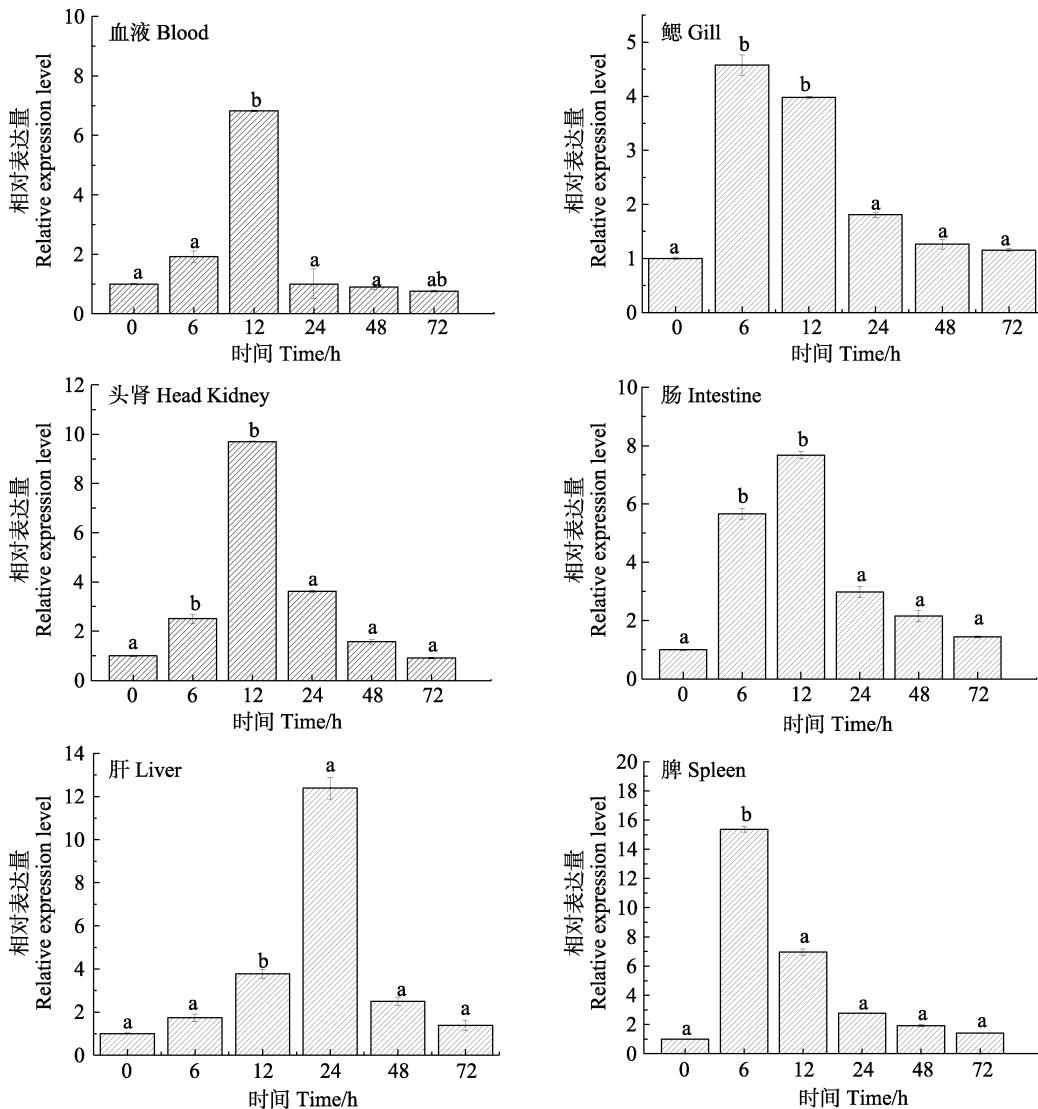


图 7 鳗弧菌感染半滑舌鳎后不同时间点 6 种组织(血液、鳃、头肾、肠、肝和脾)中 *CsMASP1* 基因的相对表达水平

Fig.7 Relative expression of *CsMASP1* gene in six tissues (blood, gill, head kidney, intestine, liver and spleen) of *C. semilaevis* at different time points after *Vibrio anguillarum* infection

重要的天然免疫防御分子,通过激活补体和调理吞噬 2 种方式来清除病原体(Chen *et al.*, 2000)。MASP1 是形成 C4bC2a 转化酶的必要因子,是凝集素激活途径中的特异性启动酶(Oroszlan *et al.*, 2016)。MASP 最初是作为与 Complement component 1s(Cls)功能相似可以裂解 C4 和 C2 的一种新型丝氨酸蛋白酶被发现的(Matsushita *et al.*, 1992)。鉴于 MASP 与 Cls 相似,推测 MBL-MASP1-MASP2 与 C1q-C1r2-C1s2 相似,因此, MASP1 和 MASP2 可能分别与类 Complement component 1r (Chr)和类 C1s 丝氨酸蛋白酶有相似的功能(Matsushita *et al.*, 1998、2000)。虽然在硬骨鱼的免疫系统中已经鉴定出了 MASP1,但其对细菌感染的免疫功能尚不清楚(Endo *et al.*, 2003; Dang *et al.*,

2017)。本研究成功克隆了 *CsMASP1* 基因,预测其编码蛋白的分子量为 84.95 kDa,与已报道的草鱼 MASP1 (Dang *et al.*, 2017)的分子量和人 MASP1 (Matsushita *et al.*, 1998)的分子量类似。*CsMASP1* 编码的氨基酸序列经 SignalP 软件预测无信号肽,而草鱼 MASP1 蛋白序列中包含一段 19 aa 的信号肽(Dang *et al.*, 2017)。研究表明, MASP 的 SP 区(Serine protease domain)是催化功能区,含有所有丝氨酸蛋白酶活性中心高度保守的组氨酸(His)、丙氨酸(Asp)和丝氨酸(Ser)等 3 个氨基酸残基,与 MASP 的丝氨酸蛋白酶活性有关(Wallis *et al.*, 2000、2007)。在本研究中,SMART 结果显示, *CsMASP1* 含有 6 个结构域,分别是 2 个 CUB 结构域(47~168 aa, 215~327 aa)、1 个 EGF

结构域(169~212 aa)、2个CCP结构域(331~392 aa, 397~461 aa)和1个Tryp-SPc结构域(478~752 aa)。与文献报道的人MASP1(Matsushita *et al.*, 1998; Stover *et al.*, 2001; Dobo *et al.*, 2016)结构域相同,与草鱼MASP1(Dang *et al.*, 2017)结构域也相同。丝氨酸蛋白酶结构域(SP)活性中心有1个典型的组氨酸环。氨基酸序列比对结果显示,丝氨酸蛋白酶结构域完全相似,组氨酸His⁵²⁶、天冬氨酸Asp⁵⁸²和丝氨酸Ser⁷⁰⁵定位相当保守,与文献报道的文昌鱼和七鳃鳗(Endo *et al.*, 2003)的比对结果一致,说明该活性中心在进化上相对保守。这些特性与MASP/C1r/C1s家族的进化相关(Jia *et al.*, 2003)。系统进化分析结果显示,CsMASP1与伯氏朴丽鱼MASP1亲缘关系最近,与金目鲈和黄鳝等鱼类的MASP1亲缘关系次之。

3.2 CsMASP1在健康组织和鳗弧菌感染后的免疫组织中的表达分析

本研究利用qRT-PCR检测CsMASP1在半滑舌鳎11种健康组织和感染后6种免疫组织中的表达谱,结果显示,在所检测的11种组织中分布广泛且具有组织特异性,在肝中表达水平最高,脾中表达水平次之,在头肾中的表达相对其他组织也较高,推测其可能在免疫调控中发挥作用。与草鱼的分布结果不相同,草鱼MASP1在心脏中表达水平最高,在肝和脑中表达水平次之(Dang *et al.*, 2017)。

MASP1是构成先天性免疫防御体系的重要成员之一(Ammitzboll *et al.*, 2013)。本研究报道了CsMASP1参与鳗弧菌感染后机体反应的免疫应答。鳗弧菌感染半滑舌鳎72 h内,CsMASP1在血液、鳃、头肾、肠、肝和脾组织中均出现上调表达的情况,总体趋势表现为表达量先升高,达到一个峰值后,随感染时间的延长表达量逐渐降低恢复至正常水平,说明CsMASP1在病原入侵后参与了机体的免疫应答。鳃和肠中的CsMASP1表达量在鳗弧菌感染后6 h大幅上调,鳃在6 h达到最大值,肠比鳃稍微滞后,肠在12 h达到最高峰,分别是对照组的7.7倍和4.6倍,随后下降到与对照组无显著差异。肠是最大的免疫器官,肠上分布着大量的肠相关淋巴组织,肠道组织和鳃是腹腔注射鳗弧菌后直接接触的部位,所以肠和鳃CsMASP1的表达量均先于其他免疫组织上调。随后,病原菌进入其他免疫组织,激活其他免疫组织CsMASP1基因的表达。与Dang等(2017)研究发现草鱼感染嗜水气单胞菌后,草鱼MASP1在鳃和肠中表达量短时快速增加结果相一致;但草鱼MASP1在肠中感染后4、24、72 h显示3个表达高峰与本研究的

结果不太一致,推测草鱼MASP1在肠中感染后出现3个峰值可能是由于感染菌不同,也可能是由于鱼体差异等。脾CsMASP1的表达量随着时间增加而增加,在感染6 h后达到最大值,约是对照组的15倍,是所有免疫组织中增量最高的组织;这与肾、脾是鱼类免疫细胞生成的主要场所,也是产生免疫应答的主要部位这一观点是一致的(Wang *et al.*, 2018)。6种免疫组织表现出不同的表达模式可能与不同时间点的信号转导有关,在我们之前的研究中,感染后,补体激活途径、凝血级联激活途径和弧菌感染相关途径均被激活来抵御病原菌的入侵(Zhang *et al.*, 2015)。CsMASP1在免疫组织中的高表达暗示了CsMASP1在机体免疫反应中可能发挥着重要作用。本结果为深入研究MASP1基因在半滑舌鳎免疫防御系统中的作用奠定了基础。

参 考 文 献

- Ammitzboll CG, Steffensen R, Nielsen HJ, *et al.* Polymorphisms in the MASP1 gene are associated with serum levels of MASP-1, MASP-3, and MAp44. *PLoS One*, 2013, 8(9): e73317
- Cai XM, Zhao N, Zuo DM, *et al.* Construction and expression of prokaryotic expression vector of human MASP1 N-terminal fragment. *Cell and Molecular Immunology*, 2008, 24(6): 546~549 [蔡学敏, 赵娜, 左大明, 等. 人MASPIN端片段原核表达载体的构建及其表达. 细胞与分子免疫学杂志, 2008, 24(6): 546~549]
- Chen ZL, Lu X, Zhang LY, *et al.* Cloning and identification of human MBL-related serine protease 1 gene. *Cellular and Molecular Immunology*, 2000, 16(5): 417~419 [陈政良, 卢晓, 张丽芸, 等. 人MBL相关丝氨酸蛋白酶1 cDNA的克隆与鉴定. 细胞与分子免疫学杂志, 2000, 16(5): 417~419]
- Chen Y, Fan L, Liu TT, *et al.* Cloning and expression analysis of TRAF6 gene and TAK1 gene of half-smooth tongue sole. *Chinese Aquatic Science*, 2015, 22(5): 867~876 [陈燕, 樊琳, 刘田田, 等. 半滑舌鳎TRAF6基因和TAK1基因的克隆及表达分析. 中国水产科学, 2015, 22(5): 867~876]
- Dahl MR, Thiel S, Matsushita M, *et al.* MASP-3 and its association with distinct complexes of the mannan-binding lectin complement activation pathway. *Immunity*, 2001, 15(1): 127~135
- Dang YF, Shen YB, Xu XY, *et al.* Mannan-binding lectin-associated serine protease-1 (MASP-1) mediates immune responses against *Aeromonas hydrophila* *in vitro* and *in vivo* in grass carp. *Fish and Shellfish Immunology*, 2017(66): 93~102
- Dobo J, Pal G, Cervenak L, *et al.* The emerging roles of mannose-binding lectin-associated serine proteases (MASPs) in

- the lectin pathway of complement and beyond. *Immunological Reviews*, 2016, 274(1): 98–111
- Endo Y, Nonaka M, Saiga H, et al. Origin of mannose-binding lectin-associated serine protease (MASP)-1 and MASP-3 involved in the lectin complement pathway traced back to the invertebrate, amphioxus. *Journal of Immunology*, 2003, 170(9): 4701–4707
- Jia TJ, Li P, Zhang SM. Research status of structure and function of MASP. *Medical Review*, 2003, 9(4): 193–195 [贾天军, 李萍, 张庶民. MASP 的结构与功能研究现状. 医学综述, 2003, 9(4): 193–195]
- Lu Y, Wang QL, Li M, et al. Expression analysis of *TLR20* and *TLR21* genes in channel catfish *Ictalurus punctatus* challenged by different bacteria and virus. *Progress in Fisheries Science*, 2013, 34(6): 44–51 [路飚, 王启龙, 李敏, 等. 斑点叉尾鮰 *TLR20* 和 *TLR21* 基因在不同细菌和病毒感染后的表达特征. 渔业科学进展, 2013, 34(6): 44–51]
- Matsushita M, Fujita T. Activation of the classical complement pathway by mannose-binding protein in association with a novel Cls-like serine protease. *Journal of Experimental Medicine*, 1992, 176(6): 1497–1502
- Matsushita M, Endo Y, Nonaka M, et al. Complement-related serine proteases in tunicates and vertebrates. *Current Opinion in Immunology*, 1998, 10(1): 29–35
- Matsushita M, Thiel S, Jensenius JC, et al. Proteolytic activities of two types of mannose-binding lectin-associated serine protease. *Journal of Immunology*, 2000, 165(5): 2637–2642
- Matsushita M, Endo Y, Fujita T. MASPI (MBL-associated serine protease 1). *Immunobiology*, 1998, 199(2): 340–347
- Nicola F, Loriano B. Morula cells as key hemocytes of the lectin pathway of complement activation in the colonial tunicate *Botryllus schlosseri*. *Fish and Shellfish Immunology*, 2017, 63: 157–164
- Oroszlan G, Kortvely E, Szakacs D, et al. MASP-1 and MASP-2 do not activate pro-factor D in resting human blood, whereas MASP-3 is a potential activator: kinetic analysis involving specific MASP-1 and MASP-2 inhibitors. *Journal of Immunology*, 2016, 196(2): 857–865
- Schwaebi W, Dalai MR, Thiel S, et al. The mannose-binding lectin-associated serine proteases (MASPs) and MAp19: Four components of the lectin pathway activation complex encoded by two genes. *Immunobiology*, 2002, 205(4–5): 455–466
- Sha ZX, Wang LQ, Sun LM, et al. Isolation and characterization of monocyte/macrophage from peripheral blood of half-smooth tongue sole (*Cynoglossus semilaevis*). *Fish and Shellfish Immunology*, 2017, (65): 256–266
- Stover C, Endo Y, Takahashi M, et al. The human gene for mannan binding lectin associated serine protease-2, the effector component of the lectin route of complement activation, is part of a tightly linked gene cluster on chromosome 1p36.2-3. *Genes and Immunity*, 2001, 2(3): 119–127
- Sunyer JO, Lambris JD. Evolution and diversity of the complement system of poikilothermic vertebrates. *Immunological Reviews*, 1998, 166(1): 39–57
- Swierzko AS, Szala A, Sawicki S, et al. Mannose-binding lectin (MBL) and MBL-associated serine protease-2 (MASP-2) in women with malignant and benign ovarian tumours. *Cancer Immunology Immunotherapy*, 2014, 63(11): 1129–1140
- Takahashi M, Sekine H, Endo Y, et al. Comment on “Mannan-binding lectin-associated serine protease (MASP)-1 is crucial for lectin pathway activation in human serum, whereas neither MASP-1 nor MASP-3 is required for alternative pathway function”. *Journal of Immunology*, 2013, 190(6): 2477
- Wang LQ, Wang H, Chen YD, et al. Isolation and identification of T lymphocyte from peripheral blood of half-smooth tongue sole and analysis of *TCRβ* gene immune response. *Acta Hydrobiologica Sinica*, 2018, 42(3): 480–487 [汪林庆, 王航, 陈亚东, 等. 半滑舌鳎外周血 T 淋巴细胞的分离、鉴定及 *TCRβ* 基因免疫应答分析. 水生生物学报, 2018, 42(3): 480–487]
- Wallis R, Dodd RB. Interaction of mannose-binding protein with associated serine proteases: Effects of naturally occurring mutations. *Journal of Biological Chemistry*, 2000, 275(40): 30962–30969
- Wallis R, Dodds AW, Mitchell DA, et al. Molecular interactions between MASP-2, C4, and C2 and their activation fragments leading to complement activation via the lectin pathway. *Journal of Biological Chemistry*, 2007, 282(11): 7844–7851
- Yu MJ, Sun LM, Chen YD, et al. Cloning and expression of the *c-Jun* gene in immune response to pathogen challenge in half-smooth tongue sole (*Cynoglossus semilaevis*). *Progress in Fisheries Science*, 2017, 38(2): 31–39 [于孟君, 孙璐明, 陈亚东, 等. 半滑舌鳎(*Cynoglossus semilaevis*) *c-Jun* 基因的克隆及免疫应答分析. 渔业科学进展, 2017, 38(2): 31–39]
- Zhang X, Wang SL, Chen SL, et al. Transcriptome analysis revealed changes of multiple genes involved in immunity in *Cynoglossus semilaevis* during *Vibrio anguillarum* infection. *Fish and Shellfish Immunology*, 2015, 43(1): 209–218

(编辑 冯小花)

Cloning and Expression Analysis of the *MASP1* Gene from the Half-Smooth Tongue Sole (*Cynoglossus semilaevis*)

WANG Qian¹, ZHANG Xue¹, CHEN Yadong^{2,3}, SHA Zhenxia^{1,2①}

(1. College of Life Science, Qingdao University, Qingdao 266071; 2. Laboratory for Marine Fisheries Science and Food Production Processes, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266071;
3. Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Key Laboratory for Sustainable Development of Marine Fisheries, Ministry of Agriculture and Rural Affairs, Qingdao 266071)

Abstract Mannan-binding lectin-associated serine protease 1 (MBL associated serine protease 1, MASP1) is an important activator in the complement lectin pathway. In this study, the cDNA of the *MASP1* gene from *Cynoglossus semilaevis* (*CsMASP1*) was cloned using a RACE method based on the partial sequence of *CsMASP1*, and gene expression was performed by qRT-PCR. The results showed that the full length of cDNA was 2507 bp in size, including a 5'-untranslated region (UTR) of 82 bp, 3'-UTR of 142 bp, and a complete open reading frame (ORF) of 2283 bp, encoding 760 amino acids. The theoretical isoelectric point (pI) of the predicted protein was 5.33 and the molecular weight was 84.95 kDa. Homologous alignment showed that the amino acids sequence of *CsMASP1* had a high identity with those of other species, approximately 42%~76%. The *CsMASP1* gene was expressed in all tested tissues (liver, intestine, spleen, head-kidney, gill, blood, brain, skin, heart, trunk-kidney, and gonad) in the healthy *C. semilaevis* and the highest expression was in the liver (185.4). To study the expression patterns of the *CsMASP1* gene in an immune response, the specific expression of *CsMASP1* was performed after *Vibrio anguillarum* infection. The results showed that the expression of the *CsMASP1* gene was up-regulated in the liver, gill, blood, intestine, head-kidney, and spleen after *V. anguillarum* infection. The most significantly up-regulated expression and the peak level at 6 h reached 15.4 times baseline in the spleen. The results indicate that the *CsMASP1* genes are involved in the immune response.

Key words *Cynoglossus semilaevis*; *MASP1*; Gene cloning; Gene expression; Immune response

① Corresponding author: SHA Zhenxia, E-mail: shazhenxia@163.com