

曼氏无针乌贼精子鞭毛蛋白1(Spefl)基因克隆 以及组织表达特异性

周林¹, 李颖¹, 吕振明¹, 迟长凤^{1*}, 吴常文¹, 史会来^{2*}

(1. 浙江海洋大学海洋科学与技术学院, 国家海洋设施养殖工程技术研究中心,

海洋生物种质发掘与利用国家地方联合实验室, 浙江舟山 316022;

2. 浙江省海洋水产研究所, 浙江舟山 316021)

摘要: 为研究精子鞭毛蛋白1(Spefl)在精子鞭毛结构的形成与组装中的生物学意义及功能, 本研究采用RACE技术和荧光定量PCR技术对曼氏无针乌贼*Spefl*(简称*SjSpefl*)基因cDNA全长进行克隆和组织表达特异性分析。结果显示, *SjSpefl* cDNA全长序列共1 135 bp, 5'和3'非编码区分别为178 bp和165 bp, 预测的开放阅读框(ORF)全长792 bp。编码的蛋白理论分子量为30.567 7 ku, 等电点7.03, 是一种亲水性蛋白。不存在跨膜区以及信号肽序列, 是在细胞内发挥作用的蛋白。二级结构分析发现该蛋白含有丰富的螺旋结构(49%)。氨基酸同源建模显示其蛋白的CH2结构域主要由4个螺旋结构组成, 并由多个loop结构串联而成。同源氨基酸序列比对发现, 它与加州双斑蛸的相似性最高且仅为59.49%, 表明*Spefl*在进化中并不保守。基于*Spefl*氨基酸序列构建的系统进化分析表明, 曼氏无针乌贼和加州双斑蛸进化关系最近。组织特异性分析表明*Spefl*在曼氏无针乌贼的精巢中有显著表达。*Spefl*基因的成功克隆以及组织表达特异性分析对于深入研究其细胞定位以及生物学功能具有重要意义。

关键词: 曼氏无针乌贼; *Spefl*; cDNA; 生物信息学; 荧光定量PCR

中图分类号: Q 786; S 917.4

文献标志码: A

精子尾部鞭毛对于正常精子的运动必不可少, 精子鞭毛为精子前进提供动力, 并决定精子的活力。研究发现多种蛋白质参与精子鞭毛的形成过程, 包括细胞骨架蛋白、信号蛋白以及其他一些未知功能的蛋白^[1]。人(*Homo sapiens*)和其他哺乳动物体内存在多种与精子鞭毛形成相关的蛋白, 包括MNS1^[2]、Spag16^[3]、Kif3a^[4]和Spef2^[5]。而精子鞭毛蛋白1(Spefl)是一种定位于精子鞭毛鞘中的蛋白, *Spefl*基因最初在小鼠(*Mus musculus*)2号染色体上某个含有*Cenpb*^[6]和*Cdc25C*^[7]基因的区域被发现。*Cdc25C*是一种细胞周期调节蛋白, 该蛋白主要在精巢中表达

并参与精巢的某些功能; *Cenpb*也被认为与精巢的某些生理功能有关。而*Spefl*作为*Cenpb*临近的下游基因, 很可能也在精巢的某些功能中扮演着重要角色。目前已有研究表明, 小鼠的*Spefl*参与精子的发生、形成和鞭毛结构的形成^[8]。综上, *Spefl*基因的研究有助于丰富物种的繁殖生物学, 进而在物种人工苗种繁育、资源增殖和保护等方面具有重要意义。曼氏无针乌贼(*Sepiella japonica*)作为一种具有较高药用、食用价值的海洋经济头足类物种, 在海洋生态系统的调节、稳定以及海洋生物多样性变动过程中扮演着重要角色^[9]。20世纪70年代以来, 过度捕捞和

收稿日期: 2017-04-10 修回日期: 2017-11-01

资助项目: 浙江省自然科学基金(LY15C190010); 舟山市科技专项(2016C41016); 浙江省科研院所专项(2015F50055); 浙江海洋大学“海洋科学”省重中之重学科开放课题(20160116)

通信作者: 迟长凤, E-mail: chicf@zjou.edu.cn; 史会来, E-mail: shihuilai1980@163.com

海洋环境的持续恶化造成其自然资源枯竭,针对此种情况,国内学者对其资源保护以及生殖发育开展了大量的研究,初见成效^[10-11]。本研究首次克隆了编码曼氏无针乌贼*Spefl*(定义为*SjSpefl*)的全长cDNA序列,结合生物信息学分析、探讨了*SjSpefl*基因以及蛋白结构与其可能存在的生物学功能的关系,并通过荧光定量PCR(qRT-PCR)技术分析了编码该蛋白的mRNA在不同组织中的表达特异性,为进一步研究*SjSpefl*的生物学功能以及作用机理奠定了基础,同时为曼氏无针乌贼的人工繁育、繁殖生物学和资源恢复奠定理论基础。

1 材料与方法

1.1 实验材料

实验对象 雄性性成熟的曼氏无针乌贼取自浙江温州苍南头足类种苗繁育基地,活体解剖取精巢、储精囊、输精管、前列腺、精荚、胃、肠、肝脏、鳃、视叶、脑、心脏及胰脏共13种组织样品保存于RNAstore中,以干冰装运至实验室,置于-80℃保存备用。

实验试剂 RNAstore购自CWbio公司;RNAiso Plus、Reverse Transcriptase M-MLV(RNase H⁻)、PCR Master Mix、SYBR Premix Ex Taq™ II Kit、SMART™ RACE Kit均购自TaKaRa公司。

1.2 实验方法

总RNA提取及cDNA第一链的合成 取保存于RNAstore中的新鲜精巢组织足量,根据RNAiso Plus试剂说明书提取总RNA。cDNA的合成按照M-MLV(RNase H⁻)反转录试剂盒推荐的方法进行。

曼氏无针乌贼*Spefl*基因核心片段的克隆 根据实验室已有的卵巢转录组数据^[12],设计用于克隆*SjSpefl*基因核心片段的引物(表1)。

PCR体系参考PCR Master Mix试剂说明书(25 μL),反应条件:94℃预变性5 min;94℃30 s,56℃30 s,72℃30 s,35个循环;72℃7 min,12℃10 min。PCR产物经1%琼脂糖凝胶电泳检测,含有目的片段的PCR产物送至深圳华大基因科技服务有限公司进行测序。

***SjSpefl*基因全长cDNA的克隆** 根据得到的*SjSpefl*基因核心序列,参考SMART™ RACE试

剂盒的要求设计用于5'和3'RACE的引物(表1)。3'RACE采用引物Spefl-3' outer和试剂盒提供的接头引物3' adapter进行第一轮PCR扩增,接着以引物Spefl-3' inner和3' adapter进行巢式PCR二轮扩增。5'RACE采用引物Spefl-5' outer和试剂盒接头引物5' adapter以加尾的cDNA为模板进行第一轮PCR扩增,然后以Spefl-5' inner引物和5' adapter进行巢式PCR第二轮扩增。PCR条件如前所述。PCR目的条带纯化回收后连接到pMD18-T载体,后转化至DH5α感受态细胞中培养,挑取阳性克隆送至深圳华大基因科技服务有限公司测序。

表1 *SjSpefl*基因cDNA克隆使用引物

引物 primer	序列(5'—3') sequence
Spefl-1F	AGCCAATGCTACTAAACC
Spefl-1R	TGATTAGGCAATGCTCG
Spefl-5' inner	GCAAATACTCGCCTGTTA
Spefl-5' outer	AATTCCTCTTTATGGCTCA
Spefl-3' inner	CTCAAGTGCAACGTCAGAC
Spefl-3' outer	GGTGCTGCAAGTCGTACAAA

曼氏无针乌贼*Spefl*基因全长及序列分析 将克隆得到的片段通过DNAMAN软件进行拼接,并用ORFfinder (<https://www.ncbi.nlm.nih.gov/o-rffinder/>)在线工具预测基因的开放阅读框(ORF)位置;然后利用Expasy-ProtParam (<http://web.expasy.org/protparam/>)^[13]将其翻译为氨基酸序列并预测蛋白的相对分子量和等电点;采用在线工具SignalP (<http://www.cbs.dtu.dk/services/SignalP/>)^[14]以及TMHMM Server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>)分别对其信号肽以及跨膜区进行预测;采用NetPhos (<http://www.cbs.dtu.dk/services/NetPhos/>)分析其磷酸化位点;采用NCBI在线搜索工具(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)分析保守结构域;采用Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>)^[15]对曼氏无针乌贼*Spefl*及其同源氨基酸序列进行比对;使用SWISS-MODEL在线工具(<https://www.swissmodel.expasy.org/>)^[16-17]构建蛋白三维模型;最后基于*Spefl*氨基酸序列分析不同物种之间的进化关系并利用贝叶斯法构建分子进化树。

曼氏无针乌贼精子表面蛋白*Spefl*的组织表

达特异性分析 在 *SjSpefl* 基因 cDNA 全长 ORF 区内设计用于 qRT-PCR 的引物, 以曼氏无针乌贼 β -actin 基因 (*S. japonica* JN564496.1) 作为内参基因 (表 2)。

表 2 *SjSpefl* 荧光定量引物

引物 primer	序列(5'-3') sequence
Spefl-YF	TTATTACTGCTTCAAAGTCGTCT
Spefl-YR	CTGATTAGGCAATGCTCGT
Actin-F	TGAGAGGGAGATTGTGCGTG
Actin-R	GAACATAGATTCTGGAGCACGG

各组织的 RNA 提取以及 cDNA 第一链合成方法, 如“总 RNA 提取及 cDNA 第一链的合成”所述。反转录后的 cDNA 利用核酸蛋白检测仪测定浓度, 并稀释至 100 ng/ μ L 备用, qRT-PCR 加样体系及反应条件按照 SYBR Premix Ex TaqTM II Kit^产品说明书进行。

2 结果

2.1 曼氏无针乌贼 *Spefl* 基因全长的获得及序列特征

以曼氏无针乌贼精巢 cDNA 为模板, 以 Spefl-1F/1R 引物扩增获得 303 bp 核心片段, 根据该核心序列, 采用 5'RACE 和 3'RACE 方法分别获得 464 bp 和 527 bp 的 2 个目的片段。利用 DNA-MAN 软件将目的片段进行拼接, 得到 *SjSpefl* 基因全长 cDNA 共 1 135 bp, 5'UTR 区 178 bp, 3'UTR 区 165 bp, 预测 ORF 792 bp, 共编码 263 个氨基酸 (图 1)。

2.2 曼氏无针乌贼 *Spefl* 氨基酸序列比对以及同源性分析

将曼氏无针乌贼 *Spefl* 与加州双斑蛸 (*Octopus bimaculoides* XP_014768591.1)、长牡蛎 (*Crassostrea gigas* EKC38895.1)、眼斑雀鳝 (*Lepisosteus oculatus* XP_006629185.1) 和矛尾鱼

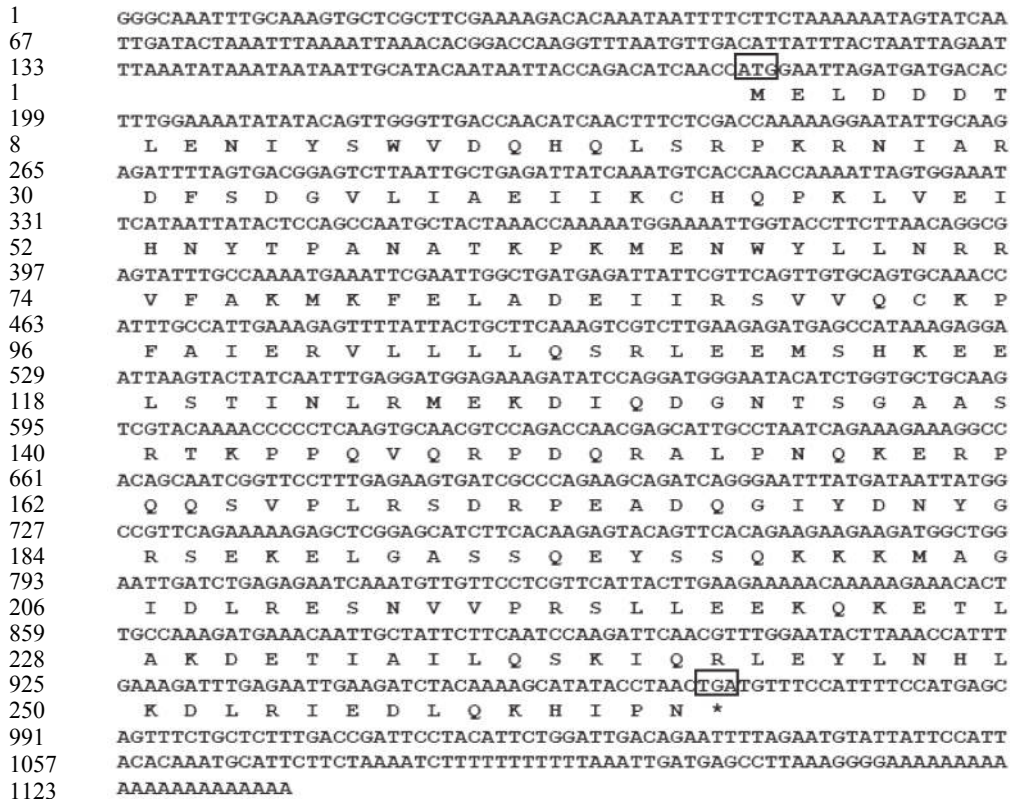


图 1 *SjSpefl* 基因 cDNA 序列及预测的氨基酸序列

预测的开放阅读框中起始密码子(ATG)和终止密码子(TGA)用黑色线框标示出, 翻译出的氨基酸序列在核苷酸序列下方给出

Fig. 1 The full length cDNA sequence and deduced amino acid sequence of *SjSpefl*

The putative start (ATG) and stop (TGA) codons are enclosed in solid lines, predicted amino acid sequence is under the cDNA sequence

(*Latimeria chalumnae* XP_014349819.1)的Spef1氨基酸同源序列进行比对,结果显示,它们的相似度分别为59.49%、36.45%、39.45和40.52%,发现组成该蛋白的氨基酸并不保守(图2)。保守结构域的分析发现,在不同物种的Spef1氨基酸序列中,靠近N端几乎同一位置均存在一个CH2结构域(2型CH结构域)^[18](图3),长度为90个氨基酸残

基,属于CH超家族成员,位于该结构域处的氨基酸在进化中比较保守。采用mrBayes软件^[19]对SjSpef1和来自于NCBI数据库的43条同源氨基酸序列构建系统进化树(图4)。结果表明,在进化上曼氏无针乌贼与加州双斑蛸的亲缘关系最近,其次是长牡蛎,与人以及大猩猩(*Gorilla gorilla gorilla*)的进化关系最远。

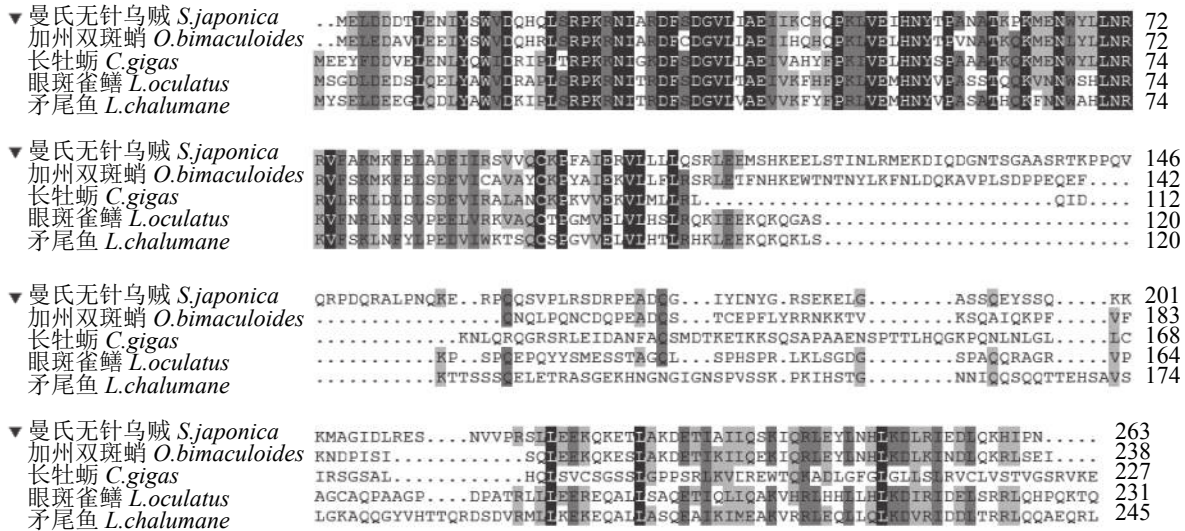


图2 Spef1氨基酸同源比对

相同氨基酸残基用黑色标出;保守氨基酸残基用灰色标出

Fig. 2 Multiple alignments of the deduced amino acid sequences of Spef1

Residues identical in all five sequences are boxed in black; conservative substitutions are boxed in gray

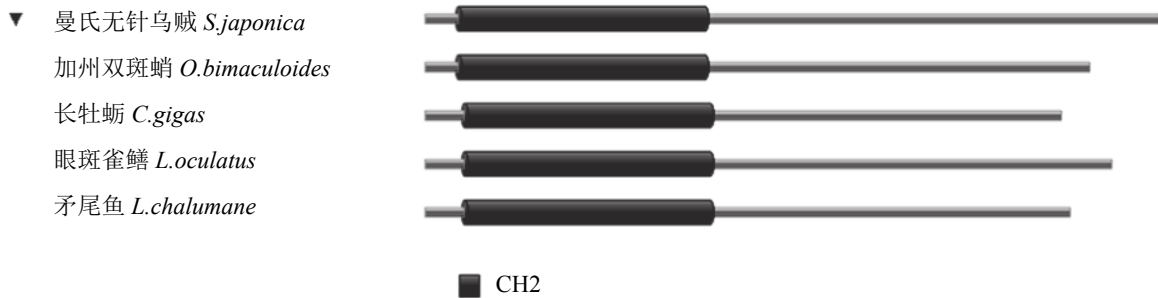


图3 不同物种Spef1 CH2结构域比对

Fig. 3 The comparison of Spef1 CH2 domain of different species

2.3 曼氏无针乌贼 *Spef1* 基因的生物信息学分析

分析表明, *SjSpef1* 基因编码的蛋白质包括214个亲水性氨基酸和49个疏水性氨基酸,为亲水性蛋白。理论分子量(MW)为30.567 7 ku,等电点(pI)为7.03。信号肽和跨膜区预测结果显示,

该蛋白无跨膜区以及信号肽结构。磷酸化位点分析表明,该蛋白存在24个可能发生磷酸化修饰的位点。采用Antheprot 5.0软件^[20]对SjSpef1蛋白的二级结构分析发现,该蛋白含有49%的螺旋(helix)、5%的折叠(sheet)、7%的转角(turn)以及38%的无规则卷曲结构(coil)(图5)。基于同源氨基

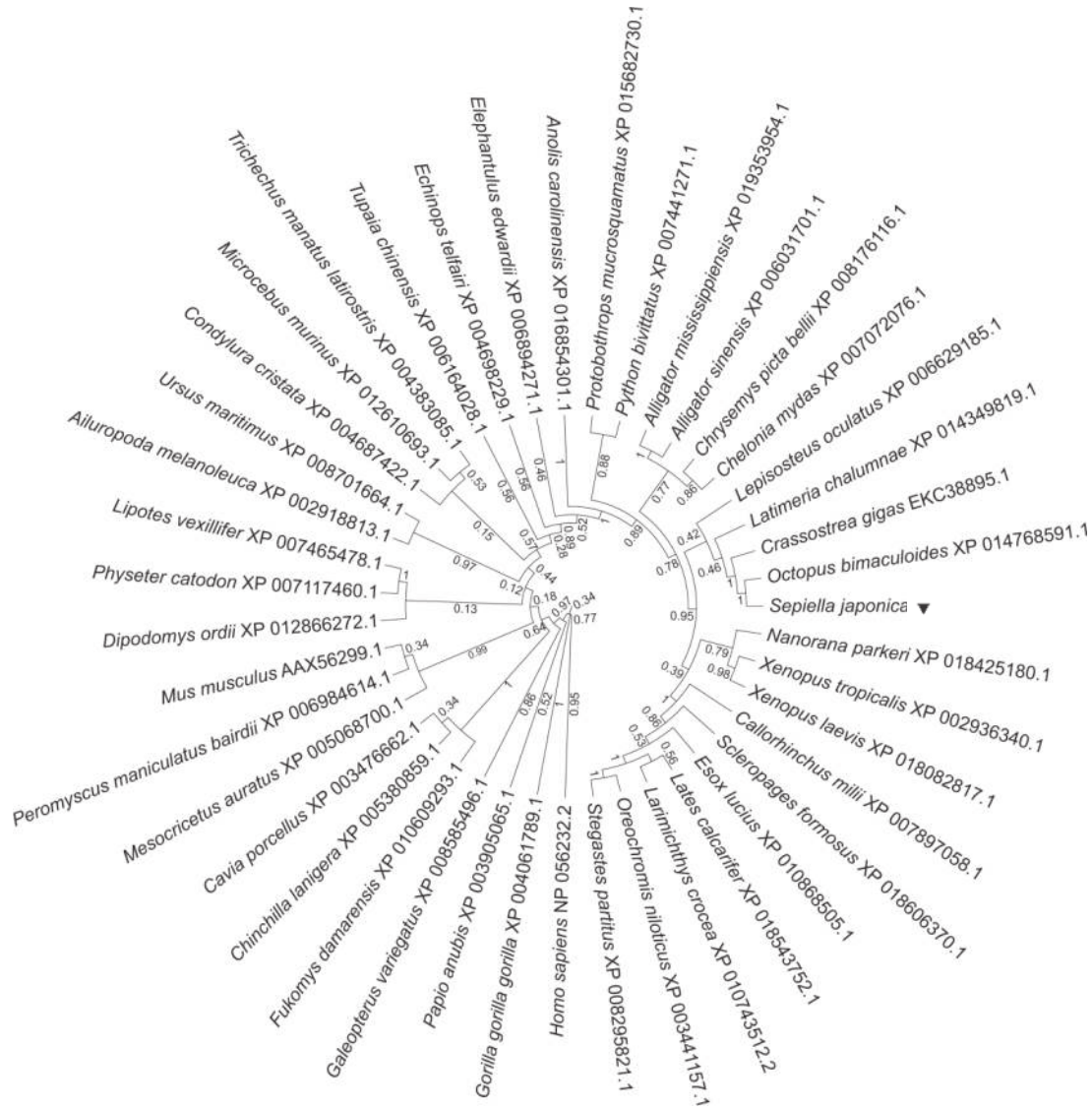


图 4 基于 Spefl 氨基酸序列构建的贝叶斯系统进化树

实心三角形标示 SjSpefl

Fig. 4 Bayesian phylogenetic tree depicting the relationship of Spefl amino acids sequence with other species

SjSpefl is denoted by solid triangle

酸构建的 SjSpefl 氨基酸三维模型(图6), 主要展示了 CH2 结构域的空间构型, 从图中可以看出 CH2 结构域主要由 4 个螺旋结构组成, 并由 3 个 loop 结构串接起来。

2.4 曼氏无针乌贼 Spefl 基因组织表达特异性分析

通过 qRT-PCR 的方法对 SjSpefl 基因在精巢、脑、储精囊等共 13 种组织中的表达特异性进行了分析, 以曼氏无针乌贼 β-actin 基因(S. japonica JN564496.1)作为内参基因, 选取心脏的表达量作为参照, 采用 2^{-ΔΔC_T} 方法计算相对表达量, 并用

SPSS 18 软件进行差异显著性分析。实验结果表明, SjSpefl 基因在性成熟时期的精巢中有显著表达(P<0.05), 而在其他组织中表达量极低(图7)。

3 讨论

Spefl 蛋白在自然界中分布比较广泛。截至目前, Spefl 的一些同源基因除了在哺乳动物如小鼠^[8, 21]、人^[22]的体内被发现之外, 在诸如脊索动物海鞘(Ciona intestinalis)^[23]以及单细胞原生动物衣藻(Chlamydomonas reinhardtii)^[24]的体内也相继被发现, 甚至在某些原生动物例如布氏锥虫

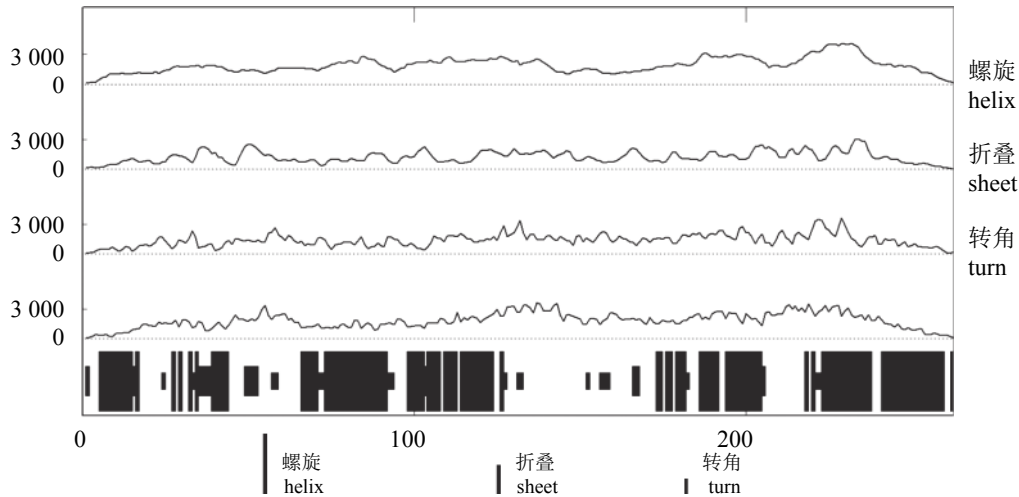


图 5 SjSpef1蛋白的二级结构预测

Fig. 5 Secondary structure prediction of SjSpef1

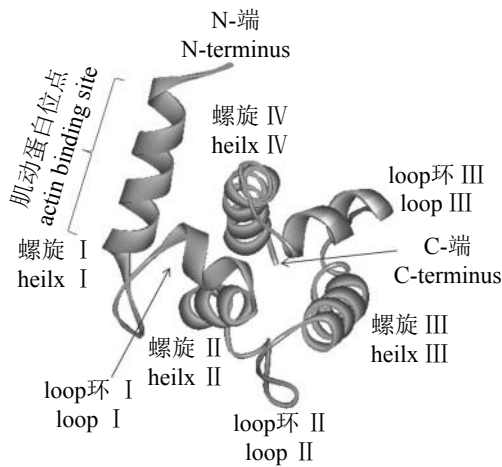


图 6 曼氏无针乌贼Spef1基于同源氨基酸建模结果

Fig. 6 The result of homology modeling of SjSpef1

(*Trypanosoma brucei*)和亚洲牛带绦虫(*Taenia asiatica*)^[25-27]中也发现了Spef1同源蛋白的踪迹, 这些证据表明, Spef1极有可能与鞭毛或者纤毛结构的形成及组装有关。SjSpef1 cDNA全长共1 135 bp, 预测ORF 792 bp, 编码263个氨基酸, 不含信号肽以及跨膜区结构。氨基酸比对发现SjSpef1与加州双斑蛸的同源性最高, 为59.49%, 其N端的部分序列在进化中比较保守, 基于Spef1氨基酸构建的系统进化关系显示曼氏无针乌贼与加州双斑蛸关系最近, 其次是长牡蛎, 与人以及大猩猩进化关系最远。针对不同物种的氨基酸序列分析发现, 在该保守序列处存在一个长度大约为90个氨基酸残基的2型CH结构域(CH2)^[18, 28]。CH结构域最早在钙调蛋白中被

发现, 它位于钙调蛋白的N端, 长度约为100个氨基酸残基, CH结构域依据其结构与功能又可分为多种类型。通常情况下, 肌动蛋白、细胞骨架蛋白以及信号蛋白这三个主要蛋白类别中也存在着不同数量的CH结构域^[29], 其中一部分CH结构域仍然发挥着它的基本功能, 即结合肌动蛋白并使肌动蛋白丝交联成束, 但另外一部分可能具有某种未知作用。通过对SjSpef1氨基酸同源建模, 发现其CH2结构域主要由4个螺旋结构组成, 其中的II、III、IV在空间上构成一个松散的螺旋结构束, 并由小的loop结构连接^[30], 螺旋结构I长度大约为14个氨基酸残基, 被认为是可能与F肌动蛋白结合的区域。虽然对于Spef1进行的大肠杆菌体外实验表明, Spef1的CH基序并不会直接与肌动蛋白结合, 但它能使肽链上串联的其他CH结构域与F肌动蛋白具有更强的亲和以及结合能力^[30-32]。有学者认为CH2结构虽不能直接与F肌动蛋白相互作用, 但它可能作为精子尾部鞭毛肌动蛋白丝结构中的定位器或者具有较低亲和力的对接点而发挥作用^[29]。因此, Spef1蛋白结构中的CH2结构功能还有待进一步研究。

本研究对Spef1基因的组织表达特异性分析发现, 它在精巢中有显著表达。之前研究人员曾通过免疫荧光的方法对Spef1基因在小鼠体内的表达进行了定位研究, 结果显示Spef1在小鼠精巢的生精上皮组织中有大量表达, 进一步的实验发现Spef1在不同时期的曲精小管腔内也均

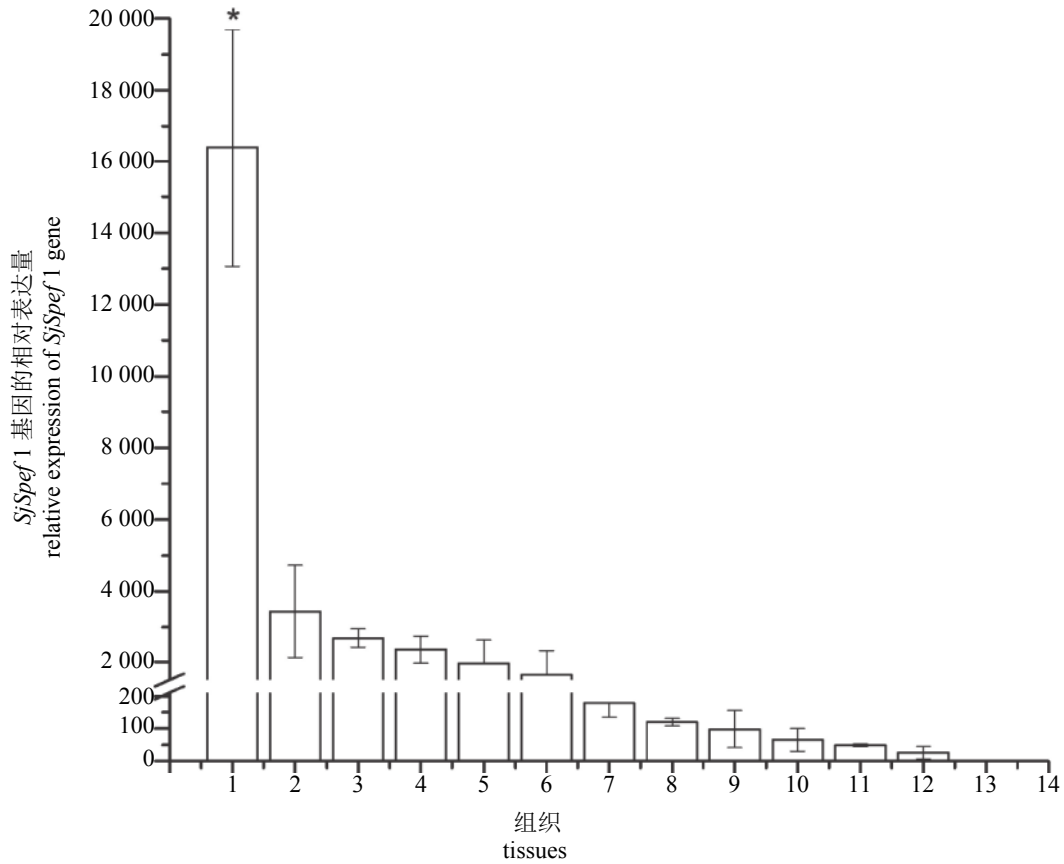


图 7 *SjSpef1*基因在不同组织中的表达特异性

1. 精巢; 2. 精荚; 3. 储精囊; 4. 输精管; 5. 肝脏; 6. 视叶; 7. 胰脏; 8. 脑; 9. 肠; 10. 鳃; 11. 前列腺; 12. 胃; 13. 心脏; “*”表示有显著差异($P<0.05$)

Fig. 7 Expression of *SjSpef1* mRNA in various tissues

1. testis; 2. spermatophore; 3. seminal vesicle; 4. spermaduct; 5. liver; 6. optic lobe; 7. pancreas; 8. brain; 9. gut; 10. gill; 11. prostate glands; 12. stomach; 13. heart; (mean \pm SE, $n=3$); bars with ‘*’ are significantly different($P<0.05$)

会产生离散信号, 即 *Spef1* 基因在精子产生的不同阶段都有表达, 但最早是在精子产生的第9个阶段检测到 *Spef1* 信号^[33], 并且随着精子逐渐发育成熟, 蛋白信号也不断增强。除了精巢, *Spef1* 基因在其他含有纤毛结构的组织中也有微量表达^[8], 由于鞭毛和纤毛结构在生物体内可能有着相同的发育起点, 因此, *Spef1* 在这些具有鞭毛或者纤毛结构的组织中可能具有类似功能。通过大量氨基酸序列的分析发现, 根据其序列的保守性差异可以将这类蛋白区分为两种类型, 即 *Spef1* 和 *Spef2*, 它们之间唯一的区别是 *Spef2* 氨基酸序列中位于羧基末端的40个氨基酸残基在进化中极不保守, 而且大量的研究发现 *Spef2* 主要在脊索动物中表达^[8]。研究人员通过对猪 (*Sus scrofa*) 精子细胞研究发现, *Spef2* 基因突变会导致雄性个体出现短尾精子从而造成精子运动活力

的大幅度减弱^[34], 进而影响受精过程。这类突变型精子通常会呈现出精子鞭毛中央微管结构、外围纤维鞘结构以及外部致密纤维结构与线粒体鞘结构异常, 最严重的会导致精子鞭毛主体结构的部分缺失^[35]。因此, 推测 *Spef1* 也可能与精子鞭毛结构的形成与组装有关。

本研究通过qRT-PCR以及RACE技术克隆得到了 *SjSpef1* 基因cDNA的全长序列, 并对其编码的长度为263个氨基酸的蛋白质进行了一系列生物信息学分析, 探讨了蛋白结构与其可能具有的生物学功能之间的关系, 并分析了 *SjSpef1* 基因的组织表达特异性, 最终推测它可能与精子鞭毛的形成与组装有关。后续有望通过原位杂交以及基因敲除的方法分别对 *SjSpef1* 的细胞定位以及生物学功能进行更加深入的探索, 为头足类精子相关蛋白的研究奠定理论基础。

参考文献:

- [1] Inaba K. Molecular architecture of the sperm flagella: molecules for motility and signaling[J]. *Zoological Science*, 2003, 20(9): 1043-1056.
- [2] Zhou J, Yang F, Leu N A, *et al.* MNS1 is essential for spermiogenesis and motile ciliary functions in mice[J]. *PLoS Genetics*, 2012, 8(3): e1002516.
- [3] Lesich K A, Zhang Z B, Kelsch C B, *et al.* Functional deficiencies and a reduced response to calcium in the flagellum of mouse sperm lacking SPAG16L[J]. *Biology of Reproduction*, 2010, 82(4): 736-744.
- [4] Corbit K C, Shyer A E, Dowdle W E, *et al.* Kif3a constrains β -catenin-dependent Wnt signalling through dual ciliary and non-ciliary mechanisms[J]. *Nature Cell Biology*, 2008, 10(1): 70-76.
- [5] Sironen A, Kotaja N, Mulhern H, *et al.* Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia[J]. *Biology of Reproduction*, 2011, 85(4): 690-701.
- [6] Sullivan K F, Glass C A. CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins[J]. *Chromosoma*, 1991, 100(6): 360-370.
- [7] Wu S, Wolgemuth D J. The distinct and developmentally regulated patterns of expression of members of the mouse *Cdc25* gene family suggest differential functions during gametogenesis[J]. *Developmental Biology*, 1995, 170(1): 195-206.
- [8] Chan S W, Fowler K J, Choo K H A, *et al.* Spef1, a conserved novel testis protein found in mouse sperm flagella[J]. *Gene*, 2005, 353(2): 189-199.
- [9] 董正之. 我国海洋生物多样性保护与可持续利用[J]. *世界科技研究与发展*, 1998, 20(2): 99-101.
Dong Z Z. Conservation and sustainable utilization of marine biological diversity in China[J]. *World Sci-tech R & D*, 1998, 20(2): 99-101(in Chinese).
- [10] 乌贼资源增殖研究课题组. 浙江近海曼氏无针乌贼资源增殖及繁殖保护的研究[J]. *浙江水产学院学报*, 1986, 5(2): 99-104.
Squid Resources Propagation Research Group by Tang Yimin. Protection of stocks and propagation of *Sepiella maindroni* de rochebrune in Zhejiang inshore waters[J]. *Journal of Zhejiang College of Fisheries*, 1986, 5(2): 99-104(in Chinese).
- [11] 张建设, 夏灵敏, 迟长凤, 等. 人工养殖曼氏无针乌贼 (*Sepiella maindroni*) 繁殖生物学特性研究[J]. *海洋与湖沼*, 2011, 42(1): 55-59.
Zhang J S, Xia L M, Chi C F, *et al.* Study on reproductive biology characteristics of cultured *Sepiella maindroni* (Rochebrune)[J]. *Oceanologia et Limnologia Sinica*, 2011, 42(1): 55-59(in Chinese).
- [12] Lü Z M, Liu W, Liu L Q, *et al.* De novo assembly and comparison of the ovarian transcriptomes of the common Chinese cuttlefish (*Sepiella japonica*) with different gonadal development[J]. *Genomics Data*, 2016, 7: 155-158.
- [13] Gasteiger E, Gattiker A, Hoogland C, *et al.* ExpASY: the proteomics server for in-depth protein knowledge and analysis[J]. *Nucleic Acids Research*, 2003, 31(13): 3784-3788.
- [14] Bendtsen J D, Nielsen H, Von Heijne G, *et al.* Improved prediction of signal peptides: SignalP 3.0[J]. *Journal of Molecular Biology*, 2004, 340(4): 783-795.
- [15] Thompson J D, Gibson T J, Plewniak F, *et al.* The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. *Nucleic Acids Research*, 1997, 25(24): 4876-4882.
- [16] Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models[J]. *Bioinformatics*, 2011, 27(3): 343-350.
- [17] Biasini M, Bienert S, Waterhouse A, *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information[J]. *Nucleic Acids Research*, 2014, 42(W1): W252-W258.
- [18] Full S J, Deinzer M L, Ho P S, *et al.* Phosphoinositide binding regulates α -actinin CH2 domain structure: analysis by hydrogen/deuterium exchange mass spectrometry[J]. *Protein Science*, 2007, 16(12): 2597-2604.
- [19] Ronquist F, Huelsenbeck J P. MrBayes 3: bayesian phylogenetic inference under mixed models[J]. *Bioinformatics*, 2003, 19(12): 1572-1574.
- [20] Geourjon C, Deléage G, Roux B. ANTHEPROT: an interactive graphics software for analyzing protein structures from sequences[J]. *Journal of Molecular Graphics*, 1991, 9(3): 188-190.
- [21] Sironen A, Hansen J, Thomsen B, *et al.* Expression of SPEF2 during mouse spermatogenesis and identification

- of IFT20 as an interacting protein[J]. *Biology of Reproduction*, 2010, 82(3): 580-590.
- [22] Werner M E, Mitchell J W, Putzbach W, *et al.* Radial intercalation is regulated by the Par complex and the microtubule-stabilizing protein CLAMP/Spefl[J]. *The Journal of Cell Biology*, 2014, 206(3): 367-376.
- [23] Satou Y, Yamada L, Mochizuki Y, *et al.* A cDNA resource from the basal chordate *Ciona intestinalis*[J]. *Genesis*, 2002, 33(4): 153-154.
- [24] Asamizu E, Nakamura Y, Sato S, *et al.* A large scale structural analysis of cDNAs in a unicellular green alga, *Chlamydomonas reinhardtii*. I. Generation of 3433 non-redundant expressed sequence tags[J]. *DNA Research*, 1999, 6(6): 369-373.
- [25] 黄江, 黄灿, 胡旭初, 等. 亚洲牛带绦虫Spefl-Like基因及其蛋白结构特征的生物信息学分析[J]. *中国人兽共患病学报*, 2007, 23(12): 1212-1215.
- Huang J, Huang C, Hu X C, *et al.* Informatics analysis for the structure and function of spefl-like gene from *Taenia saginata asiatica*[J]. *Chinese Journal of Zoonoses*, 2007, 23(12): 1212-1215(in Chinese).
- [26] Gheiratmand L. Biochemical characterization of the bilobe complex in *Trypanosoma brucei*[D]. Singapore: National University of Singapore, 2013.
- [27] 王杰, 戴佳琳, 黄江, 等. 亚洲牛带绦虫Spefl-Like基因克隆、表达及纯化[J]. *中国公共卫生*, 2009, 25(7): 827-828.
- Wang J, Dai J L, Huang J, *et al.* Cloning and expression of Spefl Like gene from *Taenia asginata asiatica*[J]. *Chinese Journal of Public Health*, 2009, 25(7): 827-828(in Chinese).
- [28] Korenbaum E, Rivero F. Calponin homology domains at a glance[J]. *Journal of Cell Science*, 2002, 115(18): 3543-3545.
- [29] Gimona M, Djinojic-Carugo K, Kranewitter W J, *et al.* Functional plasticity of CH domains[J]. *FEBS Letters*, 2002, 513(1): 98-106.
- [30] Bañuelos S, Saraste M, Carugo K D. Structural comparisons of calponin homology domains: implications for actin binding[J]. *Structure*, 1998, 6(11): 1419-1431.
- [31] Keep N H, Winder S J, Moores C A, *et al.* Crystal structure of the actin-binding region of utrophin reveals a head-to-tail dimer[J]. *Structure*, 1999, 7(12): 1539-1546.
- [32] Norwood F L M, Sutherland-Smith A J, Keep N H, *et al.* The structure of the N-terminal actin-binding domain of human dystrophin and how mutations in this domain may cause Duchenne or Becker muscular dystrophy[J]. *Structure*, 2000, 8(5): 481-491.
- [33] Russell L D, Ettlin R A, Hikim A P S, *et al.* Histological and histopathological evaluation of the testis[J]. *International Journal of Andrology*, 1993, 16(1): 83.
- [34] Sironen A, Thomsen B, Andersson M, *et al.* An intronic insertion in KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in the pig[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(13): 5006-5011.
- [35] Sukura A, Mäkipää R, Vierula M, *et al.* Hereditary sterilizing short-tail sperm defect in Finnish Yorkshire boars[J]. *Journal of Veterinary Diagnostic Investigation*, 2002, 14(5): 382-388.

Sperm flagellar protein 1 (Spef1) gene clone and expression analysis in *Sepiella japonica*

ZHOU Lin¹, LI Ying¹, LÜ Zhenming¹, CHI Changfeng^{1*}, WU Changwen¹, SHI Huilai^{2*}

(1. National and Provincial Joint Laboratory of Exploration and Utilization of Marine Aquatic Genetic Resources, National Engineering Research Center of Marine Facilities Aquaculture, School of Marine Science and Technology, Zhejiang Ocean University, Zhoushan 316022, China;

2. Marine Fisheries Research Institute of Zhejiang, Zhoushan 316021, China)

Abstract: To study the biological significance and function of sperm flagellar protein 1 (Spef1) in the formation and assembly of spermatozoa, rapid amplification of cDNA ends (RACE) and qRT-PCR techniques were used to clone the gene cDNA of the Spef1 in *Sepiella japonica* and analyse the tissue expression specificity. The results showed that a 1 135 bp full-length cDNA of *Spef1* gene from *S. japonica* (described as *SjSpef1*) was obtained, which consisted of a 178 bp 5' untranslated region (UTR) and a 165 bp 3'UTR. The predicted open reading frame (ORF) was 792 bp, molecular weight of the deduced protein was 30.567 7 ku, and its pI was 7.03. The *SjSpef1* was hydrophilic, without signal peptide sequence and transmembrane regions, and it might play a intracellular role. Secondary structure analysis revealed that it contained rich spiral structures (49%). Homology modeling indicated that CH2 domain of Spef1 mainly consisted of four spiral structures, which were connected by some loop structures. The multiple sequence alignment showed that *SjSpef1* had the highest similarity with *Octopus bimaculoides* which was 59.49%, which showed that Spef1 was not very conserved in evolution. Phylogenetics analysis based on Spef1 demonstrated that *S. japonica* had the closest relationship with *O. bimaculoides*. *SjSpef1* gene was expressed significantly in testis by qRT-PCR. Its cloning and analysis of tissue specific expression will be of great significance to exploring its cellular localization and biological function in the future.

Key words: *Sepiella japonica*; Spef1; cDNA; bioinformatics; qRT-PCR

Corresponding author: CHI Changfeng. E-mail: chief@zjou.edu.cn;

SHI Huilai. E-mail: shihuilai1980@163.com

Funding projects: Natural Science Foundation of Zhejiang Province, China (LY15C190010); Science and Technology Program of Zhoushan (2016C41016); Special Program for Zhejiang Provincial Research Institute (2015F50055); Open Foundation from Marine Sciences in the Most Important Subjects of Zhejiang (20160116)