

## 中间球海胆 *Y-box* 基因分子特性和表达分析

田 焱, 张丙龙, 常亚青\*, 董萍萍

(大连海洋大学农业部北方海水增养殖重点开放实验室, 辽宁 大连 116023)

**摘要:** 为获得 *Y-box* 蛋白在中间球海胆的生物学功能, 实验克隆了中间球海胆 *Y-box* 基因的全长 cDNA, 并进行了定量表达模式分析。*Y-box* 基因全长 cDNA 为 2 425 bp, 该序列具有一个 81 bp 的 5' 非编码区, 1 348 bp 3' 非编码区, ORF 为 996 bp, cDNA 编码 331 个氨基酸的前体蛋白, 预测蛋白分子质量为 37.4 ku, 理论等电点是 8.55, 属于亲水性蛋白。中间球海胆 *Y-box* 前体蛋白的氨基酸序列包含 3 个结构域: 氨基酸 N 末端, 亲水结构域 C 末端和冷休克结构域 (cold shock domain CSD), 其中 CSD 区为 *Y-box* 蛋白家族中保守结构域。海胆 *Y-box* 蛋白质二级结构中无规则卷曲占 88.22%, 延伸链占 9.97%,  $\alpha$ -螺旋 1.98%, 无  $\beta$ -折叠。氨基酸序列与其他物种氨基酸序列的同源性为 22.37%~41.26%, 在保守结构域 CSD 区同源性达到 75% 以上。实时定量 PCR 检测, *Y-box* 基因在中间球海胆体腔液、管足、围口膜、雄性性腺、雌性性腺、肌肉和肠 7 个不同组织中均有表达, 但差异不显著。温度胁迫下, 中间球海胆 *Y-box* 基因的表达均表现为下调, 并且随着温度的升高, *Y-box* 基因的表达量呈现逐渐下降的趋势。

**关键词:** 海胆; *Y-box* 基因; 克隆; 表达

**中图分类号:** Q 785; S 917.4

**文献标志码:** A

*Y-box* 蛋白是从细菌到人类都高度保守的多功能蛋白家族, 在转录调节、翻译调控、mRNA 隐蔽、选择性剪接、DNA 修复、细胞增殖和再生等多方面具有重要作用<sup>[1-5]</sup>。*Y-box* 蛋白具有 3 个典型的结构域: 冷休克结构域 (cold shock domain, CSD), 此结构域高度保守, 约由 70 个氨基酸残基组成, 包含 RNP1 和 RNP2 两个 RNA 结合基序, 富含芳香族氨基酸和精氨酸<sup>[1,6]</sup>, 通过与 DNA 相互作用起到转录调控的作用, 作为 RNA 伴侣与 RNA 作用起到稳定 RNA 二级结构进而调控翻译过程; C-末端结构域, 位于 CSD 结构域后的亲水结构域, 富含精氨酸和甘氨酸<sup>[7]</sup>, 参与结合蛋白间的自身或异源聚合<sup>[8]</sup>; N-末端结构域, 此区域变异较大, 能与转录复合物中的蛋白质特异结合, 进而使转录活动调控具有专一性<sup>[9]</sup>。

关于 *Y-box* 蛋白的生物学功能研究很多, 其

中包括冷适应功能, 转录调节功能及蛋白质翻译调控功能等。冷休克蛋白 mRNA 能在低温条件下有效翻译, 而其他蛋白质的合成则被完全阻断, 主要是由于冷休克 mRNA 都含有一个下游盒, 在低温条件下能增强翻译起始<sup>[10]</sup>。在大肠杆菌中, 冷休克反应的前 4~6 h 内的所有冷休克蛋白表现为瞬间的冷休克诱导, 许多冷休克蛋白能够与核糖体或 RNA 相互结合。原核生物中冷休克蛋白通过选择性的与 *Y-box* 框序列 (CTGATTGGCCAA) 中的一条链结合来激活转录<sup>[11-12]</sup>。在真核生物中 *Y-box* 蛋白的转录调节作用分成转录激活与转录抑制两个方面。研究发现, *Y-box* 蛋白在金鱼 (*Carassius auratus*) 精细胞里表达非常高<sup>[13]</sup>, 在金鱼中是一种特别的精子 RNA 结合蛋白, 在非洲爪蟾 (*Xenopus tropicalis*) 中 *Y-box* 做为精细胞的特别转录因子抑制翻译。

收稿日期: 2012-08-24 修回日期: 2012-12-15

资助项目: 国家“八六三”高技术研究发展计划 (2012AA10A412); 水产种质资源与养殖技术重点开放实验室开放课题

通信作者: 常亚青, E-mail: yqchang@dlou.edu.cn

在小鼠卵母细胞中 *Y-box* 蛋白占总蛋白的 2%<sup>[14]</sup>,在小鼠精子细胞中占 0.7%<sup>[15]</sup>,在小鼠中研究发现缺少 *Y-box* 蛋白的小鼠形态生长发育正常但不育<sup>[16]</sup>。

随着分子生物学技术的快速发展,在无脊椎动物中成功克隆了多个 *Y-box* 基因家族的成员,在比较低等的锥虫 (*Trypanosoma brucei*) 中发现了 *Y-box* 基因家族的成员 RBP16,实验证明该蛋白与 RNA 作用并调节 RNA 的活性<sup>[17]</sup>;在涡虫 (*Dugesia etrusca*) 中克隆了 *DjY1* 基因<sup>[5,11]</sup>,在果蝇 (*Drosophila melanogaster*) 中克隆了 *Yps* 基因<sup>[18]</sup>,在摇蚊 (*Chironomus tentans*) 中发现了其 *YB-1* 基因,经过选择性剪接能够产生两个变体<sup>[19]</sup>,在血吸虫 (*Schistosoma mansoni*) 中成功提取了 *SMYB1*,并发现它与 DNA 和 RNA 发生作用<sup>[20]</sup>,在海鞘 (*Ciona intestinalis*) 中发现了 3 个选择性剪接变体 *CiYB1*、*CiYB2*、*CiYB3*,都属于 *Y-box* 基因家族成员<sup>[3,11]</sup>。同时已经获得 *Y-box* 蛋白的还有玻璃海鞘 (*Ciona intestinalis*)<sup>[21]</sup>、加州海兔 (*Aplysia californica*)<sup>[22]</sup>、秀丽线虫<sup>[10]</sup>、金鱼<sup>[13]</sup>、日本青鳉 (*Oryzias latipes*)<sup>[23]</sup>、红鳍东方鲀 (*Fugu rubripes*)<sup>[23]</sup>、日本七鳃鳗 (*Lampetra japonica*)<sup>[24]</sup> 等。

中间球海胆 (*Strongylocentrotus intermedius*) 属棘皮动物门,海胆纲,正形目,球海胆科。原产于日本北部,1989 年由大连海洋大学引进,因其生长快、品质佳、性腺指数高而受到青睐。本实验成功克隆了 *Y-box* 基因的全长 cDNA 序列,运用生物信息学方法,对该基因编码蛋白从氨基酸组成、理化性质、亲/疏水性、亚细胞定位及三级结构等方面进行预测和分析,并进行了组织及温度胁迫的定量表达分析,旨在为后续中间球海胆 *Y-box* 基因功能研究奠定基础。

## 1 材料与方 法

### 1.1 实验材料

实验材料为 2 龄中间球海胆,海胆体质量 (24.55 ± 0.27) g,取自大连海洋大学农业部北方海水增殖重点实验室。养殖用水为砂滤海水, pH 7.98,盐度 31。主要试剂有 Trizol 试剂, M-MLV 反转录酶、RNA 酶抑制剂、dNTP、EX Taq DNA 聚合酶、pMD 19-T Vector、*Hind* III、*Bam* HI、实时定量相关试剂购自宝生物(大连)公司,PCR

引物均由宝生物(大连)公司合成。BD SMART™ RACE cDNA Amplification Kit 试剂盒购自 Clontech 公司。AXYGEN 柱式凝胶回收试剂盒、AXYGEN 小量质粒提取试剂盒购自上海生工生物工程有限公司。

### 1.2 总 RNA 的提取与反转录

取中间球海胆的性腺,按 Trizol 法提取总 RNA。采用紫外分光光度计与 1.0% 琼脂糖凝胶电泳检测 RNA 的浓度及完整性, -80 °C 保存备用。使用 MMLV 反转录 cDNA。反转录体系:100 ng 总 RNA, 1 μL oligo(dT), 5 × M-MLV Buffer 2 μL, dNTP Mixture (各 10 mmol/L) 0.5 μL, RNase Inhibitor (40 U/μL) 0.25 μL, RTase M-MLV (200 U/μL) 0.25 μL, 用 RNase-free 的水定容至 10 μL。反应条件:42 °C 水浴 30 min, 95 °C 5 min。

### 1.3 PCR 扩增、PCR 产物的克隆及测序

在中间球海胆 EST 数据库进行检索比对,将获得的与海胆 *Y-box* 基因有关的 EST (相似度在 85% 以上)用 Invitrogen 软件进行拼接,形成拼接体 (contig)。然后再以此拼接体再次在海胆 EST 数据库中进行同源检索拼接,反复进行比对连接直到拼接序列不再延伸为止,获得中间球海胆 *Y-box* 的 cDNA 序列片段。以组装后的 *Y-box* 的 cDNA 序列片段在 GenBank 中进行 Blast 同源搜索,进行序列比对,进而判断拼接得到的 *Y-box* 的 cDNA 片段的正确性。用拼接后的 cDNA 序列片段设计特异引物(表 1)。所有引物由上海生工生物工程技术服务有限公司合成。利用 BD SMART™ RACE cDNA Amplification Kit 试剂盒进行 RACE 克隆。反应体系:1 μL cDNA, 0.5 μL 正向引物, 0.5 μL 反向引物, 10 × Ex Taq Buffer (Mg<sup>2+</sup> Plus) 2.5 μL, dNTP Mixture (各 2.5 mmol/L) 2 μL, Ex DNA 聚合酶 (5 U/μL) 0.13 μL, 灭菌蒸馏水 18.37 μL。反应条件:94 °C 预变性 3 min; 94 °C, 30 s, 53 °C, 60 s, 72 °C 延伸 5 min, 35 个循环; 72 °C, 10 min, 4 °C 保温。回收 cDNA 片段后,与 pMD19-T vector 载体连接,将连接好的质粒转化感受态 DH5α,通过蓝白斑筛选和菌落 PCR 初步鉴定后,阳性菌落于 37 °C 震荡培养过夜。提取质粒后将鉴定为阳性的克隆送宝生物(大连)技术有限公司进行测序。

### 1.4 序列分析

序列同源性比对和相似性搜索用 Blast 软件

进行;序列拼接用 DNAMAN 软件进行,开放阅读框(ORF)的寻找用 NCBI 上的 ORF Finder 进行;信号肽查找用 Target P 1.1 Server 程序预测;利用 Tmhmm 对跨膜域进行预测;NetNGlyc 1.0 分析推测糖基结合位点;蛋白特征模体查找采用 Smart 软件;蛋白质性质分析采用 ExPASy server 软件;用 PSORT Prediction 蛋白的亚细胞定位,氨基酸组成用 Bioedit 分析,蛋白 3D 结构预测用 I-Tasser 在线软件,多序列比对采用 DNAMAN 软件,用 MEGA 4.0 的邻接法对中间球海胆 *Y-box* 氨基酸序列与 GenBank 上发表的氨基酸序列构建系统进化树,1 000 次重复计算自展分析值,利用 NCBI 在线软件预测结构特征。

表 1 本研究所用引物序列

Tab.1 Sequences of primers used in the study

引物 primer	引物序列(5'-3') primer sequence
O1	CAGACCCCGTAACTCGCAG
O2	CACAAGAGAGGCGATAAGCA
I1	CAGACCCCGTAACTCGCAG
I2	CACAAGAGAGGCGATAAGCA
RTF	GCGAGTTACGGGGTCTGC
RTR	TGAAGCCGCCGAGTGAG
18F	GTTCGAAGGCGATCAGATAC
18S	CTGTCAATCCTCACTGTGTC

### 1.5 *Y-box* 基因的定量表达分析

采用 RT-PCR 方法,检测 *Y-box* 基因在不同组织及温度胁迫下的转录表达情况。不同组织中表达分析每个组织取 5 枚海胆,分别取海胆的体腔液、管足、围口膜、雄性性腺、雌性性腺、肌肉和肠组织的 RNA,并反转录成 cDNA。温度胁迫的实验温度梯度分别设置为 19、22、25 和 28 °C,2 h 后每组取 5 枚海胆管足,提取 RNA,并反转录成 cDNA。各做 3 组平行对照,利用引物 RTF 和 RTR 对样品进行扩增,用 18S RNA 的 18F 和 18S 作内参引物。PCR 反应条件为 94 °C 5 min,94 °C 30 s,60 °C 30 s,72 °C 30 s,进行 28 个循环;循环结束后 72 °C 延伸 5 min;4 °C 保存。

## 2 结果

### 2.1 PCR 的扩增结果

以总 RNA 反转录得到的 cDNA 为模板,利

用设计好的特异性引物进行扩增,获得了与预期大小一致的目的片段(图 1),其中 Marker 为 DL2000。

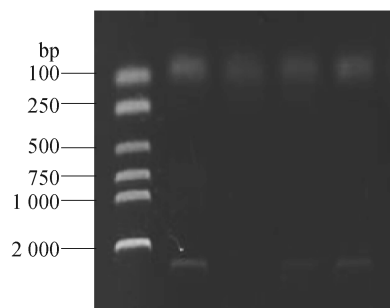


图 1 PCR 扩产物电泳

Fig.1 Gel electrophoresis of PCR product

### 2.2 中间球海胆的 *Y-box* 的 cDNA 全长序列

将 *Y-box* 基因的 EST 片段和 PCR 获得的 cDNA 序列拼接后获得中间球海胆 *Y-box* 基因的全长 cDNA。该 cDNA 全长为 2 425 bp,5'非编码区为 81 bp,3'非编码区为 1 348 bp,开放阅读框 ORF 为 996 bp。在 *Y-box* 5'端的第一个 ATG 为该基因的起始密码子;在起始密码子 ATG 的-3 位为 A,符合 Kozak 规律。同时 3'UTR 位有典型的加尾信号 AATAA(图 2),并有终止密码子 TAG 和 poly(A)符合有效翻译 *Y-box* 基因 cDNA 全长的特征。

### 2.3 中间球海胆的氨基酸组成特征

用 ProtParam 分析 *Y-box* cDNA 可编码 331 个氨基酸的前体蛋白,预测蛋白分子质量为 37.4 ku,理论等电点是 8.55。*Y-box* 的 331 个氨基酸中强碱性氨基酸(K,R)有 68 个,强酸性氨基酸(D,E)有 66 个,疏水氨基酸(A,I,L,F,W,V)有 42 个,不带电荷的极性氨基酸有(N,C,Q,S,T,Y)有 78 个,分子式为  $C_{1542}H_{2453}N_{547}O_{542}S_1$ ,氨基酸组成如图 3 所示。*Y-box* 蛋白质在中间球海胆体内半衰期大于 20 h,不稳定系数为 67.4,总平均亲水性为 -1.897。以 ProtScale 程序对 *Y-box* cDNA 编码的氨基酸分析,发现第 15~24、27~29、52~54、74~78、94~97、99~102 残基为疏水性的,第 5~14、30~51、55~70、79~93、103~327 残基为亲水性的(图 4)。亲水性残基所占比例远大于疏水性残基,而且疏水性最大值为 1.578,亲水性最大值为 3.844,因此推测 *Y-box* 编码蛋白质为亲水性。

```

1      ggatccccgggctgcaggaagaccattgtgcgaacgagtcggaatTTTTTccacctg
61     agaaaaacatcatccaggaagATGTCGAGGGCCAAGAAGAACAGCCCCAGACCATTGAG
1      M S E G Q E E Q P Q T I E
121    GCTCCGAGAAGATCGTTTTGGCCACAAAAGTTTCAGGAACAGTCAAGTGGTTCAACGTC
14     A P E K I V L A T K V S G T V K W F N V
181    AAGAATGGCTACGGGTTTCATCAACCGAGATGACACCAAGGAGGATGTATTCGTCATCAG
34     K N G Y G F I N R D D T K E D V F V H Q
241    TCGGCCATTGTAAGGAACAACCCACGCAAGTACCAGCGCAGTGTGGTGATGGAGAAGTG
54     S A I V R N N P R K Y Q R S V G D G E V
301    GTAGAGTTTGACGTTGTAGAAGGAACCAAGGGCAATGAAGCCGCCGAGTACTGGTCCT
74     V E F D V V E G T K G N E A A R V T G P
361    GAGGGTGCCCCAGTAGTAGGGAGCAAGTACGCGCCGACAAGCGCCGTTACCGCGCCGC
94     E G A P V V G S K Y A A D K R R Y R G R
421    CCCCCTACTACCGCCCGCAGACCCCGTAACTCGCAGGGAGACGAGGAGGAGATTCCC
114    P R Y Y R R R R P R N S Q G D E E E I P
481    TCAGACGAGAAGGAGGGAGGAGAAGAGAACGACGGTCCCAAGGAGGGGAGGAGACAGTAC
134    S D E K E G G E E N D G P K E G R R Q Y
541    CGTGATGGAGTAGGGGAGGCTATCGTGGTGGTCGTCGAAGGTATAGAGGTGGACGTAGA
154    R D G G R G G Y R G R R R Y R G G R R
601    CCAAGACCGGCCAGGATGGAGCACCGCCAGACGGGGAGCAAAATGAAGAGGAAGGGGAG
174    P R P G Q D G A P P D G E Q N E E E G E
661    CGACAGGATAGGGAAGGCGATGACAAGTCAAACAGCAACCGCCGTCGCCGATACCGTCCC
194    R Q D R E G D D K S N S N R R R R Y R P
721    AACCAGAGATACAACAACCGCCGCCCCAGTCCCGTGGCGACCAGAATGAGGAAGGTGGT
214    N Q R Y N N R R P Q S R G D Q N E E G G
781    GATAGGGAGAACGGGAAGGAGGTGAGGAAAGGAAGGTGGCGACGATCAACAACGTAGG

1      ggatccccgggctgcaggaagaccattgtgcgaacgagtcggaatTTTTTccacctg
61     agaaaaacatcatccaggaagATGTCGAGGGCCAAGAAGAACAGCCCCAGACCATTGAG
1      M S E G Q E E Q P Q T I E
121    GCTCCGAGAAGATCGTTTTGGCCACAAAAGTTTCAGGAACAGTCAAGTGGTTCAACGTC
14     A P E K I V L A T K V S G T V K W F N V
181    AAGAATGGCTACGGGTTTCATCAACCGAGATGACACCAAGGAGGATGTATTCGTCATCAG
34     K N G Y G F I N R D D T K E D V F V H Q
241    TCGGCCATTGTAAGGAACAACCCACGCAAGTACCAGCGCAGTGTGGTGATGGAGAAGTG
54     S A I V R N N P R K Y Q R S V G D G E V
301    GTAGAGTTTGACGTTGTAGAAGGAACCAAGGGCAATGAAGCCGCCGAGTACTGGTCCT
74     V E F D V V E G T K G N E A A R V T G P
361    GAGGGTGCCCCAGTAGTAGGGAGCAAGTACGCGCCGACAAGCGCCGTTACCGCGCCGC
94     E G A P V V G S K Y A A D K R R Y R G R
421    CCCCCTACTACCGCCCGCAGACCCCGTAACTCGCAGGGAGACGAGGAGGAGATTCCC
114    P R Y Y R R R R P R N S Q G D E E E I P
481    TCAGACGAGAAGGAGGGAGGAGAAGAGAACGACGGTCCCAAGGAGGGGAGGAGACAGTAC

```

```

134 S D E K E G G E E N D G P K E G R R Q Y
541 CGTGATGGAGGTAGGGGAGGCTATCGTGGTGGTCGTCGAAGGTATAGAGGTGGACGTAGA
154 R D G G R G G Y R G G R R R Y R G G R R
601 CCAAGACCGGGCCAGGATGGAGCACCGCCAGACGGGGAGCAAAATGAAGAGGAAGGCGAG
174 P R P G Q D G A P P D G E Q N E E E G E
661 CGACAGGATAGGAAGGCGATGACAAGTCAAACAGCAACCGCCGTCGCCGATACCGTCCC
194 R Q D R E G D D K S N S N R R R R Y R P
721 AACCAGAGATAACAACCGCCGCCAGTCCCCTGGCGACCAGAATGAGGAAGGTGGT
214 N Q R Y N N R R P Q S R G D Q N E E G G
781 GATAGGGAGAACGGGAAGGAGGTGAGGAAAGGGAAGGTGGCGACGATCAACAACGTAGG
234 D R E N G E G G E E R E G G D D Q Q R R
841 AGGCCGAGACGCCCTTACCGTCGCTACAACAATCGTAAAGGTGCCCCAGCAAGGAC
254 R P R R P Y R R R Y N N R K G A P S K D
901 GGAGAAGGGACCAAGGAGGAGAGAACCAAGGAGGAGAGAACCAAGGAGGAGAGAACCAA
274 G E G D Q G G E N Q G G E N Q G G E N Q
961 GGAGAGAACCAAGGAGACTGGAGGAGCCGGAATGAGCCCTCGGACAGACCGGTCATC
294 G E N Q G E T G G G R N E P S D R P V I
1021 AAGGAGGAGCCATCTGCTCCAGCAGATACCAAGGCCAGCATCAAGGTTGAAGAGTAGaag
314 K E E P S A P A D T K A S I K V E E ※
1081 gaggttaccaccaagaagcactgcaccctgcctttagcatttcacctagtaaatcatca
1141 tcggcacagtcatcgaagtttagacagacactgctcaccagcgatgcaagaacaacatca
1201 gcaacaagttacaacacccatggacaacatcatttctgcttttcttgattttacttgaatt
1261 caagtattttcattgctctctggattctaccaagcaccggactacaagaagtcgcgt
1321 tatcaaaattctttgcagaatgcatcattgacttctagtagtattgttatcaataagtgaa
1381 atatacaaggccttcatgacctaaccatattgctcaacaagcgtatggaccatttttgtc
1441 gcttgtatcaaagctctacactgatgacctatcgtcaatcgatttgcaaccttccaat
1501 aaaaatctggctctcttttttagtctaccattttaaatgagtacagggagtaaacattgg
1561 aaagataagtaagaggaaaagagcacagctaccaaaatctttctttttatgttgcgt

```

图2 中间球海胆 *Y-box* 基因 cDNA 序列及推导的氨基酸序列

框内为起始密码子 ATG、终止密码子 TAG, 灰色背景为冷休克区 (CSD 区), polyA 加尾信号 AATAA 用下划线标出。

Fig. 2 Full-length cDNA of *S. intermedius Y-box* and the deduced amino acid sequence

The letters in box indicate the start codon (ATG) and the stop codon (TAG). CSD is shaded by gray. The polyA special sequences AATAA is underlined.

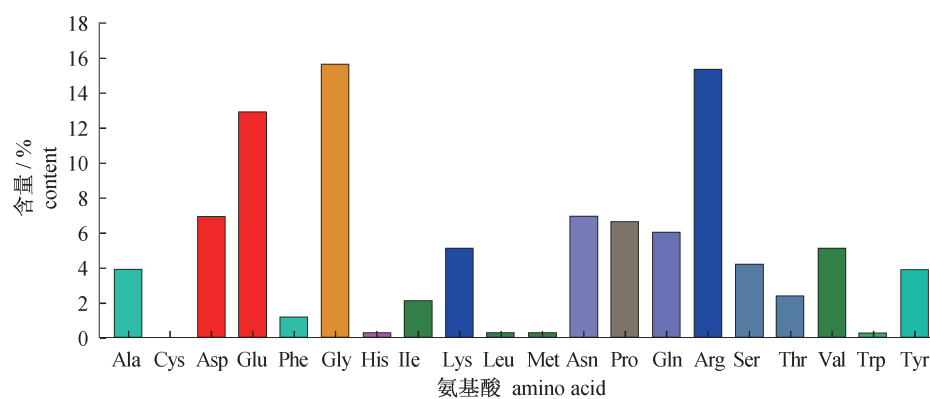


图3 中间球海胆 *Y-box* 蛋白的氨基酸组成

Fig. 3 The amino acid composition of *S. intermedius Y-box* protein

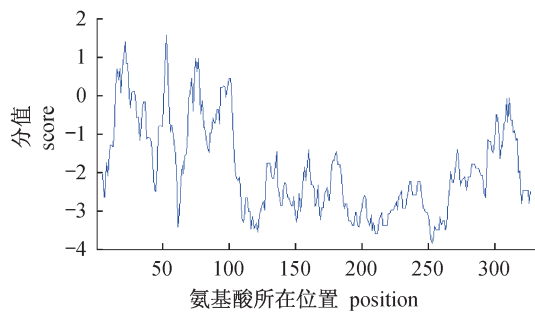


图 4 中间球海胆 Y-box 蛋白质亲水性/疏水性分析结果

Fig. 4 The Y-box protein hydrophilicity/hydrophobicity analysis results in *S. intermedius*

利用 Tmhmm 对 Y-box 前体蛋白预测跨膜域,结果表明 *Y-box* 基因无明显的跨膜区,推测该蛋白前体不属于跨膜蛋白类(图 5)。通过 NetNGlyc 1.0 分析推测糖基结合位点(NHS),推测该 Y-box 前体蛋白不是糖蛋白(图 6)。采用 TargetP 1.1 Server 预测 Y-box 氨基酸序列特征,结果表明(表 2),线粒体目标肽及分泌途径信号肽其分值分别为 0.066、0.072,预测可靠性为 1,说明该蛋白不存在信号肽酶切位点,无信号肽。故推测 Y-box 蛋白不属于分泌型蛋白,可能在游离核糖体上合成,经滑面型内质网加工而成。

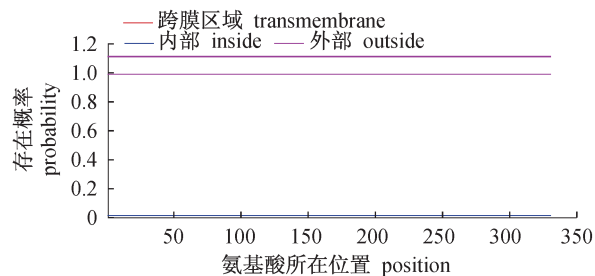


图 5 中间球海胆 Y-box 蛋白质跨膜区分析结果

Fig. 5 Prediction results of transmembrane region in Y-box in *S. intermedius*

表 2 中间球海胆 Y-box 导肽的预测

Tab. 2 Predicted leading peptide and cleavage site for Y-box in *S. intermedius*

名称 name	氨基酸数 len	线粒体 mTP	信号肽 SP	其他 other	位点 loci	可信度 RC
序列 sequence	331	0.066	0.072	0.934	-	1
截止值 cut off		0.000	0.000	0.000		

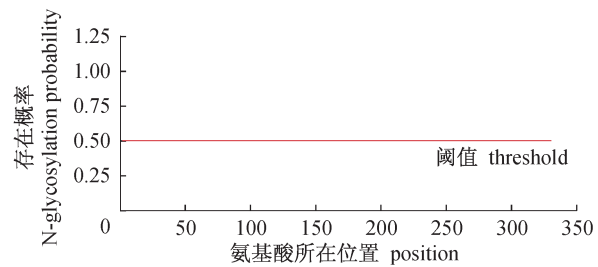


图 6 中间球海胆 Y-box 蛋白质 N-糖基化位点预测结果

Fig. 6 Prediction result of predicted N-glycosylation sites of Y-box in *S. intermedius*

## 2.4 中间球海胆 Y-box 蛋白的结构预测

中间球海胆 Y-box 蛋白质二级结构无  $\beta$ -折叠,无规则卷曲占 88.22%,延伸链占 9.97%, $\alpha$ -螺旋 1.98%,主要是以延伸链和无规则卷曲为主(图 7)。中间球海胆 Y-box 结合蛋白具有 3 个典型的结构域:氨基酸 N 末端,C 末端结构域,冷休克结构域,其中 CSD 结构域由 68 个氨基酸组成,是高度保守的核苷酸结合域,结合损伤 DNA 和非特异性 DNA,参与冷休克结合蛋白自身或异源的聚合,包含含有 RNP-1 (GYGFINR) 和 RNP-2 (EDV FVHQS) 两个 RNA 结合基序,这两个基序与其转录调控功能密切相关。蛋白质的拓扑结构对于保持蛋白质的功能是重要的,不同的拓扑结构将导致不同的功能。中间球海胆 Y-box 前体蛋白质的二级结构与人和斑马鱼 (*Danio rerio*) 的 Y-box 蛋白质在其结构及物理性质等多个方面都表现出相当高的保守性(表 3),说明蛋白质三维拓扑结构的保守性,则主要体现在保持某些共同的特有二级结构单元和折叠方式上。Y-box 前体蛋白的三级结构中包括 7 个  $\alpha$ -螺旋(红色部分)、8 处延伸链(绿色部分),其余为无规则卷曲(白色部分),无  $\beta$ -折叠(图 8)。

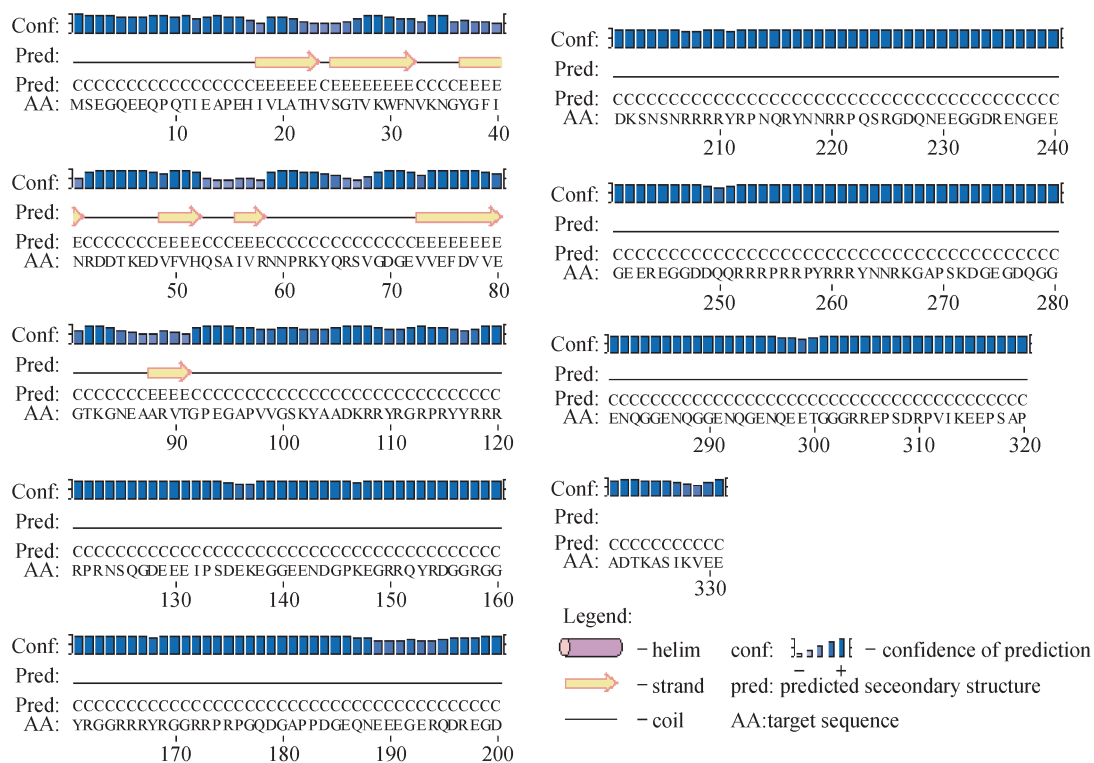


图7 中间球海胆 Y-box 蛋白质二级结构预测

Fig. 7 Prediction result of secondary structure of Y-box in *S. intermedius*

表3 3个物种 Y-box 蛋白的二级结构比较  
Tab. 3 Comparison of the secondary structure of Y-box in the three species

	中间球海胆 <i>S. intermedius</i>	人 <i>Homo sapiens</i>	斑马鱼 <i>Danio rerio</i>
亲/疏水性 hydrophilicity/ hydrophobicity	亲水性	亲水性	亲水性
信号肽 signal peptide	无	无	无
导肽 lending peptide	无	无	无
跨膜结构 transmembrane	无	无	无
亚细胞定位 location	细胞核	细胞核	细胞核
α-螺旋 α-helix	1.81%	2.97%	0.97%
无规则卷曲 random coil	88.22%	86.8%	90.00%
β-折叠 β-turn	无	无	无
延伸链 extended strand	9.97%	10.23%	9.03%

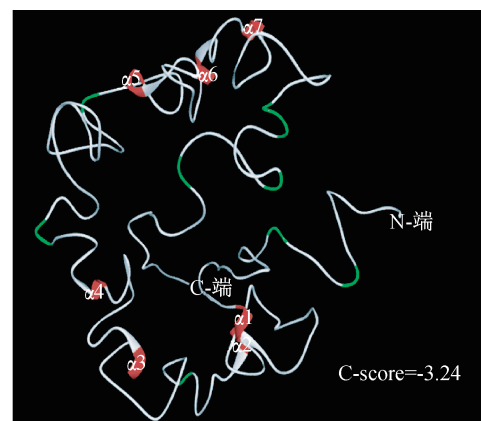


图8 中间球海胆 Y-box 蛋白的 3D 结构

红色部分为 α-螺旋、绿色为延伸链、白色为无规则卷曲。

Fig. 8 The sea urchin Y-box protein of 3D structure

The red part is the alpha helix, green for extensions chain, white curly for random coil.

## 2.5 中间球海胆与其他物种 Y-box 氨基酸序列同源序列分析

使用DNAMAN程序检索分析了中间球海胆



氨基酸序列与其他物种的同源性(图9)。选取了加州海兔、非洲爪蟾、摇蚊、七鳃鳗、果蝇、大西洋鲑(*Salmo salar*)、玻璃海鞘(*Ciona intestinalis*)、栉孔扇贝(*Chlamys farreri*)、小家鼠(*Mus musculus*)、人(*Homo sapiens*)、虹鳟

(*Oncorhynchus mykiss*)、斑马鱼、金鱼、原鸡(*Gallus gallus*)及中间球海胆氨基酸比对结果相似度为22.37%~41.26%,但是在冷休克区相似度达到75%以上。

加州海兔	兔	.....	0
非洲爪蟾	蟾	.....	0
摇蚊	蚊	.....	0
日本七鳃鳗	鳗	.....	0
果蝇	蝇	.....MADSE	5
大西洋鲑	鲑	.....	0
玻璃海鞘	鞘	.....	0
栉孔扇贝	贝	.....	0
小家鼠	鼠	MSEAGEATTGG.TTLLPQAAADAPAAAAPPDPAPKSPAASGA	39
小家鼠	鼠	MSEAGEATTTTITLTPQAPTEAAAAAPQDPAPKSPVGS	40
虹鳟	鱼	.....	0
斑马鱼	鱼	.....	0
金鱼	鱼	.....	0
原鸡	鸡	.....MSSEAE	9
中间球海胆	胆	.....	0
Consensus			
加州海兔	兔	.....MADTEKQPEVEENQPDQEQ.....NEE	22
非洲爪蟾	蟾	.....MSSEVETQQQQTVALEKAGQ.....EPAAT	26
摇蚊	蚊	.....MTDTEAPAVEAQPQKT.....AAP	19
日本七鳃鳗	鳗	.....MSEAAAQPEKPR.....PA	14
果蝇	蝇	KQQQLPIEQQOPLAQEEQODEGQQAQG.....KPA	35
大西洋鲑	鲑	.....MSSEAE	29
玻璃海鞘	鞘	.....MSEQAGQLEQVT.....ENA	15
栉孔扇贝	贝	.....MVTPLLPSEAMSDEKQPE.....EKE	22
小家鼠	鼠	P..QAPAPAAIAGSPGGDAAPGPAPASSAP.....AGGE	72
小家鼠	鼠	PQAAAAPAAHVAGNPGGDAAPATGTAAAAAATAAGSE	80
虹鳟	鱼	.....MSSEAE	29
斑马鱼	鱼	.....MSSEAE	28
金鱼	鱼	.....MSDTEASEQFQ.....	11
原鸡	鸡	P..AAPVPAAPAAAPADSKPNGSGNGSSG.LASAAPPA	46
中间球海胆	胆	.....MSEGQEEQFQ.....TIE	13
Consensus			
加州海兔	兔	QKEKKIIASQVSGIVKWFNVKSGYGFINRNDTKEDVFVHQ	62
非洲爪蟾	蟾	VGEKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	66
摇蚊	蚊	PVQKEIIATKVIITGIVKWFNVKSGYGFINRNDTKEDVFVHQ	59
日本七鳃鳗	鳗	VPEKKIIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	54
果蝇	蝇	PPKKEIIATKVIITGIVKWFNVKSGYGFINRNDTKEDVFVHQ	75
大西洋鲑	鲑	AGDKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	69
玻璃海鞘	鞘	PVERKVLASHCSGVVWFNVKSGYGFVNRDDNKEDVFIHQ	55
栉孔扇贝	贝	EPKKNVIATKVIITGIVKWFNVKSGYGFINRNDTKEDVFVHQ	62
小家鼠	鼠	DAEKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	112
小家鼠	鼠	DAEKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	120
虹鳟	鱼	AGDKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	69
斑马鱼	鱼	AGDKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	68
金鱼	鱼	.PERKVIATGVEGIVKWFNVKSGYGFINRNDTKEDVFVHQ	50
原鸡	鸡	GGDKKVIATKVLGIVKWFNVKSGYGFINRNDTKEDVFVHQ	86
中间球海胆	胆	APEKIVLATKVSQIVKWFNVKSGYGFINRNDTKEDVFVHQ	53
Consensus		a g vkwfnv gygf nr d k d f hq	
加州海兔	兔	TAIVKNNPRKYLRSVGDGEKVEFDVVEGEKGN.EAANVTG	101
非洲爪蟾	蟾	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	105
摇蚊	蚊	SALIKNNPKKAVRSVGDGEVFEFDVVEGEKGS.EAANVTG	98
日本七鳃鳗	鳗	TAIKKNNPRKYLRSVGDGEVFEFDVVEGEKGA.EAANVTG	93
果蝇	蝇	SATARNNPKKAVRSVGDGELVEFDVVEGEKGN.EAANVTG	114
大西洋鲑	鲑	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	108
玻璃海鞘	鞘	TAIINKNNPKYLRSVGDGENVEFDVVEGEKGLPEAANVTG	95
栉孔扇贝	贝	TAIINKNNPRKYLRSVGDGEKVEFDVVEGEKGN.EASNVTG	101
小家鼠	鼠	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	151
小家鼠	鼠	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	159
虹鳟	鱼	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	108
斑马鱼	鱼	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	107
金鱼	鱼	TAIKKNNPRKFLRSVGDGEVFEFDVVEAAGKS.EAANVTG	89
原鸡	鸡	TAIKKNNPRKYLRSVGDGEIVFEFDVVEGEKGA.EAANVTG	125
中间球海胆	胆	SALVRNNPRKYQRSVGDGEVFEFDVVEGTTKGN.EAARVTG	92
Consensus		ai nnp k rsvgdge vefdvv kg ea vtg	



加非州海兔	PEGSNVQGSKYAADRRRFR...	133
州爪蟾	PGGVFVQGSKYAADRNHFRR...	139
加非州海兔	PEGEFVKGSFYAADKRRG...	132
州爪蟾	EGGTAVQGSRYAADRRRFRG...	126
加非州海兔	PSGEFVRSQFAADKRRNF...	149
州爪蟾	PAGVFPVQGSKYAADRNRYRR...	144
加非州海兔	ENGEFVKGSKYAADRRRYK...	127
州爪蟾	PECNPFVQGSKYAADRRRYR...	139
加非州海兔	PDGVFVEGSRYAADRRRYRRG...	183
州爪蟾	PDGVFVEGSRYAADRRRYRRG...	191
加非州海兔	PAGVFPVQGSKYAADRNRYRR...	144
州爪蟾	PGGVFVQGSKYAADRNRYRR...	141
加非州海兔	PGGLFVKGSRYAPNKRRFRRR...	124
州爪蟾	EGGVFVQGSKYAADRNHYRR...	159
加非州海兔	PEGAPVQGSKYAADKRRYRGR...	129
州爪蟾	Consensus	
加非州海兔	GGRRPQDM...	162
州爪蟾	YQNNESGE...	169
加非州海兔	TRVKKEGK...	162
州爪蟾	RNGGEGGD...	159
加非州海兔	NEGDELAP...	182
州爪蟾	YQSDGEGEGEPRPKSRREGA...	182
加非州海兔	GNEHQEGE...	161
州爪蟾	GGEGEEGV...	171
加非州海兔	..AGEIGE...	201
州爪蟾	..AGEIGE...	211
加非州海兔	YQSDGEGEGEDRPRPKSRREGA...	182
州爪蟾	YQSDPEAE...	175
加非州海兔	SGAAPAGP...	154
州爪蟾	YQNSSEGE...	187
加非州海兔	EETPSDEK...	162
州爪蟾	Consensus	
加非州海兔	QN...RRYF..GPPRRG...	185
州爪蟾	..RRVPPYYLRRPYGRRPQYSN...	206
加非州海兔	RR...FRRGF..GRRGGGG...	181
州爪蟾	..GRRYPPYYGG...	180
加非州海兔	PRQNYRRGP..GGGFGGP...	210
州爪蟾	GRRRYPPYF...	201
加非州海兔	RRGERKPCQ...	182
州爪蟾	GGGGYRGYFRGGNFRRGPL...	202
加非州海兔	..AHRNPTYR...	209
州爪蟾	..VHRNPTYR...	219
加非州海兔	GRRRYPPYF...	201
州爪蟾	GRRRYPPYFVRRRYGRRPPYIN...	215
加非州海兔	..RRRPVQPQEGE...	175
州爪蟾	..RRRYPPYYMRRPYGRRPQYSN...	223
加非州海兔	GG...RRRYRGGRRRFRPQ...	190
州爪蟾	Consensus	
加非州海兔	YQLQR...	220
州爪蟾	SQGDQGRPVQRQNYRGRFRFR...	240
加非州海兔	FGG..G...	208
州爪蟾	AGVDGTAPRRRPGGRGFRTRPC...	217
加非州海兔	PGGPPGPGQRRYNNYPRQTRRV...	250
州爪蟾	..PDQGNKPMRQNYRGRFRPSRG...	233
加非州海兔	DGDFEQKPEQRRRPRKPRPRFR...	219
州爪蟾	MVEEDGQYMDQFRRRRPFYRFR...	242
加非州海兔	.....PRFR...	225
州爪蟾	.....PRYRS...	236
加非州海兔	..PDQGNKPMRQNYRGRFRPSRG...	233
州爪蟾	..PDQGNKPMRQNYRGRFRPSRG...	245
加非州海兔	..AQQRPLPRFRFRFRRLRPRPP...	211
州爪蟾	..AGEQGRFVRQNYRGRFRFRSLT...	260
加非州海兔	EGERQDREGDDKSNRNRRRYRPN...	230
州爪蟾	Consensus	
加非州海兔	..LPRRTIQGRITSQARRRERPWGL...	253
州爪蟾	GNEEDKENQGDETQSQQAPQRR...	277
加非州海兔	GERNIRNGSGD...	240
州爪蟾	GGAEIDKENEAGEERPFRAQQQN...	256
加非州海兔	SDGVQRGGDQPRRGGVGGPQRRE...	289
州爪蟾	GEGENKENQGDGENQEPQROR...	271
加非州海兔	ADHVEGEDGEPQRORRPRKPR...	257
州爪蟾	GSFRGRPFGRGRGGRGRGRGNSR...	279
加非州海兔	GEAEDKENQQAANGFNQPSARR...	264
州爪蟾	GEAEDKENQQAATSGENQPSVRRG...	275
加非州海兔	GEGENKENQGDGENQEPQROR...	271
州爪蟾	GE..EDKENQSESGNQEPQROR...	281
加非州海兔	QESFPKGEINREQTEGQFQKP...	248
州爪蟾	GNEEDKENQGDETQSQQPPQRR...	297
加非州海兔	EGGDRENGEGGEEREGDDQRRR...	269
州爪蟾	Consensus	
加非州海兔	.....	253
州爪蟾	..KFDQKTKA...	306

摇蚊	...	FRGG.....RRPGGPPGGPRGGNYSNNES...	264
日本七鳃鳗	GGAAA	GEGGKAAG...GNGGVASGAAATTTSGTAAASAT	293
果蝇	..YIQGG	QGGPPRQPFRRRQNRKPNPGGGGMEQQSQNG	327
大西洋鲑	..KPDG	GKEAKAD...GEPGLEKTSAPAAEQGGAE...	301
玻璃海鞘	EVPHEN	GEEGEQRPARRRRPRNRPKGPRRPKPTENENGN	297
栉孔扇贝	..SQEG	GDSNDAGA...ESSGAGDDRNDGGEQAE.....	308
小家鼠	..VSQD	GKRPRQVRHQLRTPQLPNRAVPSDFGSQAP...	299
人	..PSQD	GKEAKAG...EAPTENP..APPTQSSAE....	303
虹鳟	..KPDG	GKEAKAD...GEPGLEKTSAPAAEQGGAE....	301
斑马鱼	..KPDG	GKDSKA...ADASADKSAPEAEQGGAD....	310
金鱼	..ASKD	GDEPEKP...ESPAQASKPADVKSSEASPKQS	282
原鸡	..KPDG	GKETKT...AEPPEAENTSAPEAEQGGAE....	326
中间球海胆	..PSKD	GEGDQGGE...NOGGENQGGENQGENQGETGGGR	304
Consensus			
加州海兔	.....	.....	253
非洲爪蟾	.....	.....	306
摇蚊	.....	.....	264
日本七鳃鳗	TATST	GGEATAVAAAAAPPTPAAVAAAPSEPPAAAAA	330
果蝇	AQELQ	NITTESTA.....	340
大西洋鲑	.....	.....	301
玻璃海鞘	VPNQEN	ADAFVAQSQVFPKSE.....	320
栉孔扇贝	.....	.....	308
小家鼠	.....	.....	299
人	.....	.....	303
虹鳟	.....	.....	301
斑马鱼	.....	.....	310
金鱼	LKMDAL	PPTASKSSE.....	297
原鸡	.....	.....	326
中间球海胆	NEPSDR	PVIKEEPSAPADTKASIKVEE.....	331
Consensus			

图 9 不同物种 Y-box 氨基酸的多序列比对结果

框内为包含 70 个氨基酸残基的保守的 CSD。基序 RNP-1 (GYGFINR) 和 RNP-2 (EDVFVHQT) 用 \* 标出。

Fig. 9 Homologous comparison of amino acid sequences of the cold-shock domain in different species

The 70 amino acid residues of the cold-shock domain are boxed. The motifs of RNP-1 and RNP-2 are marked by \*.

2.6 系统进化树的构建

根据 GenBank 上已注册 Y-box 的氨基酸序列,使用 MEGA 软件构建系统进化树(图 10)。从系统进化树中可以看出大西洋鲑(NP 001133216.1)、虹鳟(NP 001158512.1)、斑马鱼(AAI68507.1)、非洲爪蟾(NP 001016677.1)、原鸡(NP 989745.1)、小家鼠(AAH27785.1)、人(NP 001138898)、七鳃鳗(ACF33226.1)、金鱼(BAA19850.1)聚为一支

属于脊椎动物,摇蚊(AAN 10049.1)、果蝇(XP 002048507.1)聚为一支属于陆地无脊椎动物,加州海兔(NP001191560)、栉孔扇贝(ACF90219.1)、玻璃海鞘(NP 001072010.1)、中间球海胆聚为一支为海洋无脊椎动物。海胆是一种相对古老的生物,从整个进化树可以看出中间球海胆 *Y-box* 基因出现的位置与其物种的进化地位是一致的,即处于无脊椎动物较原始的位置。

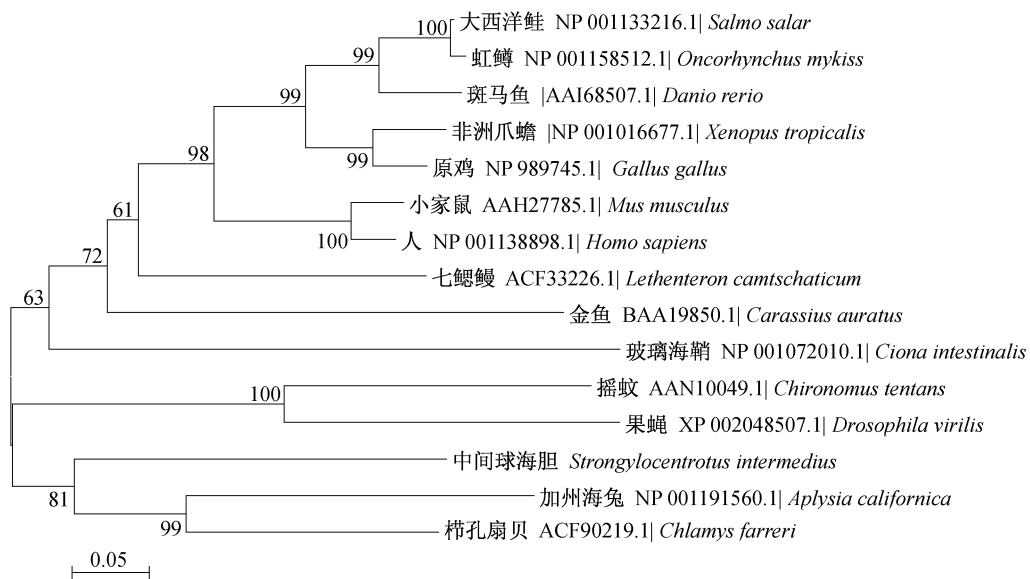


图 10 不同物种 Y-box 基因的系统进化树

Fig. 10 Phylogenetic tree of different species based on Y-box gene sequence

## 2.7 中间球海胆 *Y-box* 基因的组织表达分析

以 18S rRNA 为内参基因,利用实时定量 PCR 方法检测 *Y-box* 基因在中间球海胆体腔液、管足、围口膜、雄性性腺、雌性性腺、肌肉和肠 7 个不同组织中的表达,结果表明 *Y-box* 基因在围口膜中表达水平最高,其次是体腔液,其余组织中的表达量基本一致(图 11)。

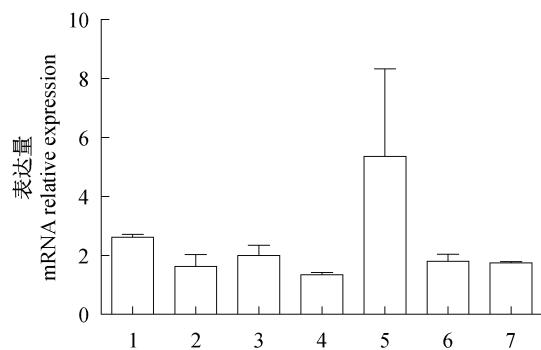


图 11 中间球海胆 *Y-box* 基因 mRNA 在不同组织的表达情况

1. 体腔液,2. 管足,3. 肌肉,4. 肠,5. 围口膜,6. 雄性性腺,7. 雌性性腺。

Fig. 11 Quantitative real-time PCR analysis of *Y-box* gene expression in different tissues from *S. intermedius*

1. coelome fluid,2. tube feet,3. muscle,4. intestine,5. peristomial membrane,6. male gonad,7. female gonad.

## 2.8 中间球海胆 *Y-box* 基因在不同温度处理下的表达

以 18S rRNA 为内参基因,利用实时定量 PCR 方法检测 *Y-box* 基因在温度胁迫 2 h 后在管足组织中的表达差异情况(图 12),结果发现 *Y-box* 基因的表达均表现为下调,并且随着温度的升高,*Y-box* 基因的表达量呈现逐渐下降的趋势。

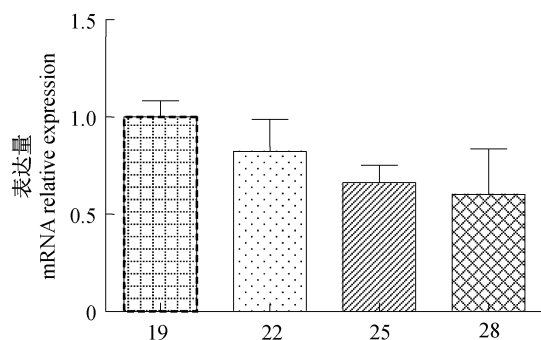


图 12 中间球海胆 *Y-box* 基因 mRNA 在温度胁迫下的表达情况

Fig. 12 Quantitative real-time PCR analysis of *Y-box* gene expression under temperature stress in *S. intermedius*

## 3 讨论

### 3.1 中间球海胆 *Y-box* 基因全长序列及组成特性

研究获得的中间球海胆 *Y-box* 基因 cDNA 全长 2 425 bp,该序列包括 81 bp 的 5'非编码区,1 348 bp 的 3'非编码区,996 bp 开放阅读框,可编码 331 氨基酸的前体蛋白。到目前为止,已经获得 *Y-box* 蛋白物种有玻璃海鞘<sup>[21]</sup>、加州海兔<sup>[22]</sup>、秀丽线虫<sup>[10]</sup>、金鱼<sup>[13]</sup>、日本青鳉<sup>[23]</sup>、红鳍东方鲀<sup>[23]</sup>、日本七鳃鳗<sup>[24]</sup>。其中加州海兔<sup>[22]</sup> *Y-box* 基因全长 cDNA 为 1 723 bp,5'非编码区 75 bp,3'非编码区 886 bp,开放阅读框 ORF 为 762 bp,编码 253 个氨基酸,与中间球海胆氨基酸相似度为 29.75%,CSD 区相似度为 85.5%。玻璃海鞘<sup>[21]</sup> *Y-box* 基因全长 cDNA 为 1 852 bp,5'非编码区 60 bp,3'非编码区 829 bp,开放阅读框 ORF 为 963 bp,编码 320 个氨基酸,与中间球海胆氨基酸相似度为 38.89%,CSD 区相似度为 77.8%。金鱼<sup>[13]</sup> *Y-box* 基因家族中 *Y-box1* 全长 1 430 bp,5'非编码区 119 bp,3'非编码区 375 bp,开放阅读框 ORF 为 936 bp,编码 311 个氨基酸,总氨基酸相似度为 25.5%,CSD 区相似度为 89.8%。金鱼<sup>[13]</sup> *Y-box2* 全长 1 503 bp,5'非编码区 14 bp,3'非编码区 595 bp,开放阅读框 ORF 为 894 bp,编码 297 个氨基酸,总氨基酸相似度为 40.66%,CSD 区相似度为 83.8%。日本青鳉<sup>[23]</sup> *Y-box1* 全长 1 464 bp,5'非编码区 136 bp,3'非编码区 407 bp,开放阅读框 ORF 为 921 bp,编码 306 个氨基酸,总氨基酸相似度为 41.36%,CSD 区相似度为 85.3%。红鳍东方鲀<sup>[23]</sup> *Y-box1* 开放阅读框 ORF 为 792 bp,编码 263 个氨基酸,总氨基酸相似度为 39.88%,CSD 区相似度为 85.3%。日本七鳃鳗<sup>[24]</sup> *Y-box1* 蛋白总氨基酸相似度为 26.62%,CSD 相似度为 92.7%。

### 3.2 中间球海胆 *Y-box* 基因全长的结构域预测

中间球海胆 *Y-box* 结合蛋白具有 3 个典型的结构域:氨基酸 N 末端,C 末端结构域,冷休克结构域,这些结构都符合 *Y-box* 前体蛋白的氨基酸序列特征。其中,CSD 结构域由 68 个氨基酸组成,是高度保守的核苷酸结合域,保守的冷休克结构域决定了 *Y-box* 结合蛋白的重要功能<sup>[11]</sup>。目前,几乎所有关于序列与结构比较的研究都支持这样的观点:对于功能重要的残基是保守的,并且位于拓扑结构

的等价位置。通过基因突变分析,证实 CSD 是与 DNA 和 mRNA 结合的结构域,负责结合损伤 DNA 和非特异性 DNA,参与冷休克结合蛋白自身或异源的聚合。CSD 包含 RNP1 和 RNP2 两个 RNA 结合基序,相当于 RRM(RNA recognition motif)蛋白家族中保守的 RNA 结合位点,其芳香环与 RNA 的碱基相互作用,甘氨酸和靠近芳香环残基的短链使其结合表面更易接近核酸,另外,*Y-box* 结合蛋白碱性区的正电荷对核酸也有潜在的吸引力<sup>[25]</sup>。高度保守的结构域也进一步证明,在进化过程中保持蛋白质功能的残基比维持蛋白质结构的残基更加保守。

氨基酸 N 末端氨基酸组成的不同是 *Y-box* 结合蛋白分子间的最大区别。由于 N-末端的不同,*Y-box* 结合蛋白才能与转录复合物中其他蛋白质特异结合,进而实现转录调控的专一性,但是 N 末端具体行使其功能的机理仍然不清楚<sup>[25]</sup>。亲水 C 末端氨基酸结构域,由酸碱区域交替组成,调节蛋白与蛋白以及蛋白与 RNA 之间相互作用的结构域。目前已知 C 末端的亲水结构域主要参与结合蛋白间的自身或异源聚合,聚合是通过精氨酸束与邻近分子中的酸性螺旋相互作用而实现的,为电荷拉链聚合<sup>[8]</sup>。

### 3.3 中间球海胆 *Y-box* 基因的位点分析与同源性分析

实验预测中间球海胆的 *Y-box* 蛋白无信号肽,没有糖基位点,说明中间球海胆 *Y-box* 蛋白不属于分泌性蛋白,这与人的 *Y-box* 蛋白功能很相似,人的 *Y-box* 蛋白通常在细胞基质与细胞核之间穿梭,以应激颗粒的形式存在于细胞基质中,主要功能是处理加工输出 mRNA,在细胞核中 *Y-box* 蛋白受到应激反应以后会发生移动包括腺病毒感染,DNA 损伤以及 PI3K-Akt 信号激活<sup>[26]</sup>。中间球海胆 *Y-box* 氨基酸序列与其他物种的氨基酸序列的相似性为 22.37%~41.26%,但在 CSD 区相似性在 75% 以上。进化树显示,大西洋鲑、虹鳟、斑马鱼、非洲爪蟾、原鸡、小家鼠、人、七鳃鳗、金鱼聚为一支属于脊椎动物,摇蚊、果蝇聚为一支陆地无脊椎动物,加州海兔、栉孔扇贝、玻璃海鞘、中间球海胆聚为一支海洋无脊椎动物。金萍<sup>[24]</sup>发现 *Y-box* 基因在有颌类出现以前只有一种类型,有颌类出现以后,*Y-box* 基因分化成 *YB1*、*YB2* 和 *CSDA* 3 种类型。在金鱼<sup>[13]</sup>中已经克隆 *YB1*、

*YB2*。从系统发育分析表明,*Y-box* 基因的进化和物种进化的拓扑结构一致。

### 3.4 中间球海胆 *Y-box* 基因的定量表达分析

实时定量 PCR 检测 *Y-box* 基因在中间球海胆体腔液、管足、围口膜、雄性性腺、雌性性腺、肌肉和肠 7 个不同组织中均有表达,但是差异不显著。*Y-box* 结合蛋白的 mRNA 在其他物种的大多数组织中也能检测到,并且 *Y-box* 结合蛋白是一类多功能结合蛋白,推测中间球海胆 *Y-box* 基因的这种组成型表达方式也从另一个角度说明其功能的多样性。*Y-box* 蛋白的生物学功能研究表明其主要具有转录调节功能及蛋白质翻译调控功能<sup>[22]</sup>。人类 *YB1* 蛋白直接结合在 *Y-box* 框或结合在和 *Y-box* 作用的转录因子上(如 TFs)<sup>[1]</sup>,*YB1* 蛋白与其它转录因子相互作用或者是促进或者是抑制,它结合在单链 DNA 启动子上,对其它的转录因子起到正调控或负调控的作用。中间球海胆 *Y-box* 基因的这种组成型表达方式也可能说明 *Y-box* 基因这种对转录和翻译的调控功能在所有组织中都具有重要作用。

温度胁迫下中间球海胆 *Y-box* 基因的表达均表现为下调,并且随着温度的升高,*Y-box* 基因的表达量呈现逐渐下降的趋势。*Y-box* 的生物学功能除了作为转录调节因子结合 DNA 调节转录活动外,还具有对 RNA 二级结构起到稳定的作用,并在低温条件下调控翻译活动。在 RNA 识别和翻译水平调控方面,有着类似于热休克蛋白(热激蛋白)的分子伴侣作用,能帮助生物体渡过胁迫环境,作为分子伴侣阻止 mRNA 错误折叠,保证细胞在低温时能正常翻译<sup>[18]</sup>。本实验结果表明,温度胁迫可诱导中间球海胆 *Y-box* 基因的表达,并做出相应的应激反应。海胆作为研究受精发育的模式生物,一直以来都是研究的热点。对于中间球海胆的 *Y-box* 基因在基因表达和调控机制及其生物学功能进行深入系统的研究,对揭示基因的表达调控具有深远的理论和实践意义。

### 参考文献:

- [1] Matsumoto K, Wolffe A P. Gene regulation by *Y-box* proteins: coupling control of transcription and translation[J]. Trends in Cell Biology, 1998, 8(8): 318-323.
- [2] Bouvet P, Wolffe A P. A role for transcription and FRGY2 in masking maternal mRNA within *Xenopus* oocytes[J]. Cell, 1994, 77(6): 931-941.

- [ 3 ] Chang B E, Lin C Y, Kuo C M. Molecular cloning of a cold-shock domain protein, zFY1, in zebrafish embryo [ J ]. *Biochimica et Biophysica Acta*, 1999, 1433(1-2):343-349.
- [ 4 ] Karlson D, Nakaminami K, Toyomasu T. A cold-regulated nucleic acid-binding protein of winter wheat shares a domain with bacterial cold shock proteins [ J ]. *The Journal of Biological Chemistry*, 2002, 277(38):35248-35256.
- [ 5 ] Valadao A F, Fantappie M R, LoVerde P T, *et al.* *Y-box* binding protein from *Schistosoma mansoni*: interaction with DNA and RNA [ J ]. *Molecular and Biochemical Parasitology*, 2002, 125 ( 1 - 2 ): 47-57.
- [ 6 ] Tafuri S R, Wolffe A P. DNA binding, multimerization, and transcription stimulation by the *Xenopus Y-box* proteins *in vitro* [ J ]. *The New Biologist*, 1992, 4(4):349-359.
- [ 7 ] Sommerville J. RNA-binding proteins: masking proteins revealed [ J ]. *Bioessays*, 1992, 14 ( 5 ): 337-339.
- [ 8 ] Wolffe A P. Chromatin: structure & function [ J ]. Academic Press, 1998, 3:343.
- [ 9 ] Yamanaka K, Fang L, Inouye M. The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation [ J ]. *Molecular Microbiology*, 1998, 27 ( 2 ):247-255.
- [ 10 ] Moss E G, Lee R C, Hmbros V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA [ J ]. *Cell*, 1997, 88(5):637-646.
- [ 11 ] Wolffe A P, Tafuri S, Ranjan M, *et al.* The *Y-box* factors: a family of nucleic acid binding proteins conserved from *Escherichia coli* to man [ J ]. *The New Biologist*, 1992, 4(4):290-298.
- [ 12 ] Ranjan M, Tafuri S R, Wolffe A P. Masking mRNA from translation in somatic cells [ J ]. *Genes & Development*, 1993, 7(9):1725-1736.
- [ 13 ] Katsu Y, Yamashita M, Nagahama Y. Isolation and characterization of goldfish *Y box* protein, a germ-cell-specific RNA-binding protein [ J ]. *European Journal of Biochemistry*, 1997, 249(3):854-861.
- [ 14 ] Yu J, Hecht N B, Schultz R M. Expression of MSY2 in mouse oocytes and preimplantation embryos [ J ]. *Biology of Reproduction*, 2001, 65(4):1260-1270.
- [ 15 ] Lasham A, Moloney S, Hale T, *et al.* The *Y-box*-binding protein, YB1, is a potential negative regulator of the p53 tumor suppressor [ J ]. *The Journal of Biological Chemistry*, 2003, 278 ( 37 ): 35516-35523.
- [ 16 ] Yu J, Deng M, Medvedev S, *et al.* Transgenic RNAi-mediated reduction of MSY2 in mouse oocytes results in reduced fertility [ J ]. *Developmental Biology*, 2004, 268(1):195-206.
- [ 17 ] Matsumoto K, Meric F, Wolffe A P. Translational repression dependent on the interaction of the *Xenopus Y-box* protein FRGY2 with mRNA. Role of the cold shock domain, tail domain, and selective RNA sequence recognition [ J ]. *The Journal of Biological Chemistry*, 1996, 271 ( 37 ): 22706-22712.
- [ 18 ] Tanaka K J, Matsumoto K, Tsujimoto M, *et al.* CiYB1 is a major component of storage mRNPs in ascidian oocytes: implications in translational regulation of localized mRNAs [ J ]. *Developmental Biology*, 2004, 272(1):217-230.
- [ 19 ] Nashchekin D, Masich S, Soop T, *et al.* Two splicing isoforms of the *Y-box* protein ctYB-1 appear on the same mRNA molecule [ J ]. *The FEBS Journal*, 2007, 274(1):202-211.
- [ 20 ] Ulusu N N, Tezcan E F. Cold shock proteins [ J ]. *Turkish Journal of Medical Sciences*, 2001, 31 ( 4 ): 283-290.
- [ 21 ] Imai K S, Hino K, Yagi K, *et al.* Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks [ J ]. *Development*, 2004, 131:4047-4058.
- [ 22 ] Skehel P A, Bartsch D. Characterization of a *Y-box* factor from *Aplysia californica* [ J ]. *Gene*, 1994, 145 ( 2 ):231-235.
- [ 23 ] Ajusch E Z, Hornung U, Burgtorf C. *et al.* Isolation and characterization of cold-shock domain protein genes, *Oryzias latipes Y-box* protein 2 (*OlaYP2*) and *Fugu rubripes Y-box* protein 1 (*FruYP1*), in medakafish and pufferfish [ J ]. *Gene*, 2002, 296 ( 1 - 2 ):111-119.
- [ 24 ] 金萍, 张玮玮, 黄惠芳, 等. 日本七鳃鳗 (*Lampetra japonica*) Lyb 基因的克隆与分析 [ J ]. *遗传*, 2008, 30(12):1597-1602.
- [ 25 ] Kong W H, Gao C L, Zuo J K. Masking mRNA by *Y-box* binding protein [ J ]. *Chinese Journal of Cell Biology*, 2000, 22:1-8.
- [ 26 ] Kedersha N, Anderson P. Mammalian stress granules and processing bodies [ J ]. *Methods in Enzymology*, 2007, 431:61-81.

## Molecular characteristics and expression analysis of *Y-box* gene from sea urchin *Strongylocentrotus intermedius*

TIAN Yi, ZHANG Binglong, CHANG Yaqing\*, DONG Pingping

(Key Laboratory of Mariculture, Ministry of Agriculture, Dalian Ocean University, Dalian 116023, China)

**Abstract:** The *Y-box* protein family is highly conserved nucleic acid binding proteins which are conserved from bacteria to human. The complete cDNA of *Y-box* was cloned through degenerate primer PCR amplification and Smart<sup>TM</sup> Race technology. The length of the cDNA of sea urchin *Y-box* gene is 2 425 bp, which contains a 996bp open reading frame(ORF), a 1 348 bp 3' UTR and an 81 bp 5' UTR. The AATAA special sequences in the 3' UTR and polyA regions were detected. The full-length *Y-box* cDNA encoded 331 amino acids constituting a 37.4 ku protein molecule with an isoelectric point of 8.55. The precursor protein was hydrophilic. The N-glycosylation site, hydrophilic C-tail end and cold shock domain CSD were evaluated. The *Y-box* protein family was very conservative in CSD domain. The secondary structure of the precursor protein consisted primarily of a 1.98% alpha helix, with a 9.97% strand chain and random coil 88.22%, without  $\beta$  folding. The comparison revealed that the similarity of the *Y-box* precursor protein ranged from 22.37% to 41.26% with other animals. The evolutionary tree indicated that the *Y-box* precursor protein of sea urchin was clustered with *Aplysia californica* and *Chlamys farreri*. The *Y-box* gene expression pattern was detected in coelomic fluid, tube feet, peristomial membrane, male gonad, female gonad, muscle and intestinal tissue. Tissue distribution pattern showed that the difference was not significant in sea urchin. The expression of *Y-box* gene under temperature stress was reduced, and the expression quantity of *Y-box* gene decreased gradually with the rise of temperature. The results provide reference for investigating the function of this gene.

**Key words:** *Strongylocentrotus intermedius*; *Y-box*; gene cloning; gene expression

**Corresponding author:** CHANG Yaqing. E-mail: yqchang@dlou.edu.cn