

文章编号: 1674-5566(2011)05-0661-08

基因组步移技术克隆稀有鮡鲫角蛋白 18 基因序列及序列分析

叶莹, 蔡生力, 刘红

(上海海洋大学 农业部水产种质资源与养殖生态重点开放实验室, 上海 201306)

摘要:角蛋白 18 是角蛋白的一种,属于 S 型角蛋白,是在哺乳动物和两栖类动物胚胎发育中第一个表达的中间纤维蛋白。根据已发表的人、爪蟾及斑马鱼角蛋白 18 DNA 序列,通过 DNAMAN 同源性比较获得保守序列,并从中设计二对引物。从稀有鮡鲫基因组中 PCR 扩增得到 2 779 bp 的角蛋白 18 的 DNA 序列片段;通过基因组步移,分别获得该基因的上游和下游序列片段;最后将获得的全基因片段克隆至质粒载体 pMD19-T 进行序列测定检验。结果显示所克隆的稀有鮡鲫角蛋白 18 基因全长 3 765 bp。对该基因进行 NCBI 的 Blastn 同源性比较分析,表明稀有鮡鲫同斑马鱼比较角蛋白 18 DNA 序列存在 73.6% 同源相似性;同时对该基因进行生物学分析,开放阅读框 (Open Reading Frame, ORF) 预测表明稀有鮡鲫角蛋白 18 是含有 7 个外显子和 6 个内含子,共编码 367 个氨基酸的蛋白质。

研究亮点:首次成功克隆了稀有鮡鲫角蛋白 18 基因全长序列 3 765 bp,并且通过软件预测分析得出氨基酸序列,把氨基酸序列在 DNA 序列中进行标记,显示稀有鮡鲫角蛋白 18 基因全长序列含有 7 个外显子和 6 个内含子,共编码 367 个氨基酸。另外,没有采用常用的从 RNA 反转录为 cDNA 为模板获得基因全长,而是直接以 DNA 为模板出发,利用基因组步移的方法获得基因全长。

关键词:稀有鮡鲫;角蛋白 18;克隆;基因组步移

中图分类号: S 917

文献标志码: A

稀有鮡鲫 (*Gobiocypris rarus*) 又名金白娘、墨线鱼,属鲤形目,鲤科,鮡鲫属,是我国特有的一种小型鲤科鱼类。由于其具有适应性强,性成熟周期短(约 3~4 个月),能周年繁殖,产卵频率高,以及对于污染物敏感等特征,已作为一种新型的模式实验鱼应用于鱼类遗传学、鱼病学、环境科学以及转基因观赏鱼等研究领域^[1-7]。

角蛋白属于中间纤维蛋白。角蛋白 18 是角蛋白的一种,是在哺乳动物和两栖类动物胚胎发育中第一个表达的中间纤维蛋白,目前也是研究鱼类胚胎早期发育中基因表达的重要靶蛋白。至今为止,多种鱼类的角蛋白 18 的 cDNA 被成功克隆^[8-9],但有关稀有鮡鲫角蛋白 18 的研究未见报道。本研究旨在克隆鲤科鱼类稀有鮡鲫的角蛋白 18 基因序列并根据基因序列预测其氨基酸序列,为稀有鮡鲫的分子生物学研究及观赏鱼转

基因操作提供参考依据。

1 材料与方法

1.1 实验动物

实验用稀有鮡鲫由中国科学院水生生物研究所王剑伟教授惠赠,本实验室长期饲养保存。

1.2 主要试剂

胶回收试剂盒 (Agarose Gel DNA purification kit ver 2.0)、pMD19-T vector、500 bp Marker、LA Taq 酶购自 Takara 公司。UNIQ-10 柱式动物基因组 DNA 提取试剂盒、Agarose Blow EEO 购自上海生工生物工程有限公司 (上海生工)。Genome Walker Universal Kit 购自 Clontech 公司。其他试剂均采用国产分析纯。

收稿日期: 2011-04-02 修回日期: 2011-05-01

基金项目: 上海市教育委员会重点学科建设项目 (J50701)

作者简介: 叶莹 (1984—), 女, 硕士研究生, 研究方向为水产动物繁殖与发育生物学。E-mail: yeying3356@yahoo.com.cn

通讯作者: 蔡生力, E-mail: slcai@shou.edu.cn

1.3 DNA的提取

取稀有鮕鲫新鲜的肌肉组织约 50 mg, 按上海生工的 UNIQ-10 柱式动物基因组 DNA 提取试剂盒步骤进行 DNA 提取。

1.4 引物设计与合成

根据 GenBank 已发表的斑马鱼的角蛋白 18 的 DNA 序列与人、爪蟾的角蛋白 DNA 序列用 DNAMAN 进行同源性比较得到保守序列, 并以此序列作为模板利用 Premier 5.0 软件从中设计一对引物(上游引物 P: GAACGACCGTCTGGCCTCC TATCTG; 下游引物 Q: GATTTCTGCCTCCAGCTT ATC), 引物均由上海生工合成。

1.5 PCR 扩增

扩增总体系: 总体积为 25 μ L, DNA 模板 1 μ L, 2 \times Taq PCR mastermix 12.5 μ L, 正反向引物各 0.5 μ L, 加 ddH₂O 至 25 μ L。PCR 程序: 94 $^{\circ}$ C, 1 min; 94 $^{\circ}$ C, 40 s; 58.5 $^{\circ}$ C, 40 s; 72 $^{\circ}$ C, 3 min 30 s; 72 $^{\circ}$ C, 15 min。1% 琼脂糖凝胶电泳检测扩增效果, 凝胶成像系统拍照并记录。

1.6 克隆及测序

PCR 产物经 1.2% 琼脂糖凝胶电泳, 回收纯化后与 pMD19-T (Takara, Japan) 载体相连接, 转化大肠杆菌 DH5 α 感受态细胞, 涂布含氨苄青霉素 (100 μ g/mL) 的平板培养过夜, 挑取白色菌斑, 接种含氨苄青霉素的 LB 培养基, 37 $^{\circ}$ C, 220 r/min 振荡过夜, 提取质粒, 以 M13 正反向引物进行 PCR 鉴定后, 取菌液送往上海桑尼生物技术有限公司测序, 测序结果用 Blastn 软件 (<http://www.ncbi.nlm.nih.gov/BLAST/>) 进行同源性比对。

1.7 基因组步移文库的构建

分别选用 4 种限制性内切酶 Dra I、EcoR V、Pvu II 和 Stu I 消化稀有鮕鲫基因组 DNA。按照 Clontech 基因组步移文库说明书构建文库 D、E、P、S。

1.8 基因组步移

根据得到的 DNA 片段序列设计特异性引物 (F1, F2, R1, R2), 利用基因组步移技术对目的基因的 3' 和 5' 末端进行 PCR 扩增。

在 3' 端和 5' 扩增中, 以构建的文库为模板, 利用引物 F1 (5'-ACCAACAATTCAAACCTAACCTC

CTTGC-3') 与接头引物 AP1 以及 R1 (5'-TTCGCAGACAGATTCAGACATTGGAGA-3') 与接头引物 AP1, 进行 PCR 扩增, 反应条件: (94 $^{\circ}$ C, 25 s; 72 $^{\circ}$ C, 3 min) \times 7; (94 $^{\circ}$ C, 25 s; 67 $^{\circ}$ C, 3 min) \times 32; 67 $^{\circ}$ C, 7 min。所得 PCR 产物稀释 100 倍后取 1 μ L 用 F2 (5'-GCTCCAGGGTCCTCACTTTCTCCAGAT-3') 与接头引物 AP2 并且 R2 (5'-CTCCTGGGTTCCCTTTGCTTCATTCTC-3') 与接头引物 AP2, 进行巢式 PCR, 反应条件: (94 $^{\circ}$ C, 25 s; 72 $^{\circ}$ C, 3 min) \times 5; (94 $^{\circ}$ C, 25 s; 67 $^{\circ}$ C, 3 min) \times 20; 67 $^{\circ}$ C, 7 min。电泳、DNA 片段回收、连接、转化、测序同前述。

1.9 序列分析

利用 Seqman 软件进行序列拼接; 拼接结果用 Blastn 软件在 GenBank 数据库中 (<http://www.ncbi.nlm.nih.gov/BLAST/>) 进行同源性比对。利用 GEECEE (<http://bioweb.pasteur.fr/seqanal/interfaces/geecee.html>) 进行 GC 含量分析。用 GeneMark (<http://opal.biology.gatech.edu/GeneMark/>) 分析预测核酸序列的开放读码框。

2 结果与分析

2.1 角蛋白 18 基因的克隆

以引物 P 和 Q, 稀有鮕鲫基因组 DNA 为模板进行 PCR 扩增, 结果如图 1(a), 得到一条约为 2 800 bp 的特异性条带, 与预期相符。将 PCR 扩增得到的 2 800 bp 左右的产物进行测序, 得到一条 2 779 bp 的 DNA 序列, 此序列在 NCBI 上进行同源性搜索。结果显示所得到的 2 779 bp 的序列中, 6 到 188 bp 与金黄色鲫鱼的角蛋白 mRNA 有 83% 的序列相似性, 346 到 1 236 bp 与斑马鱼 23 号染色体的部分 DNA 序列有 84% 的序列相似性, 2 112 到 2 779 bp 与斑马鱼 23 号染色体的部分 DNA 序列有 83% 的序列相似性。此斑马鱼 23 号染色体的部分序列经 GenBank 查找, 属于斑马鱼角蛋白 18 基因的序列。在所扩增得到的 2 779 bp 的 DNA 序列中, 总共有 1 845 bp 与鱼类的角蛋白 18 基因存在 80% 以上的序列相似性, 因此可以初步判定: 实验中所获得的 2 779 bp 的 DNA 片段即为稀有鮕鲫角蛋白 18 基因的部分序列。

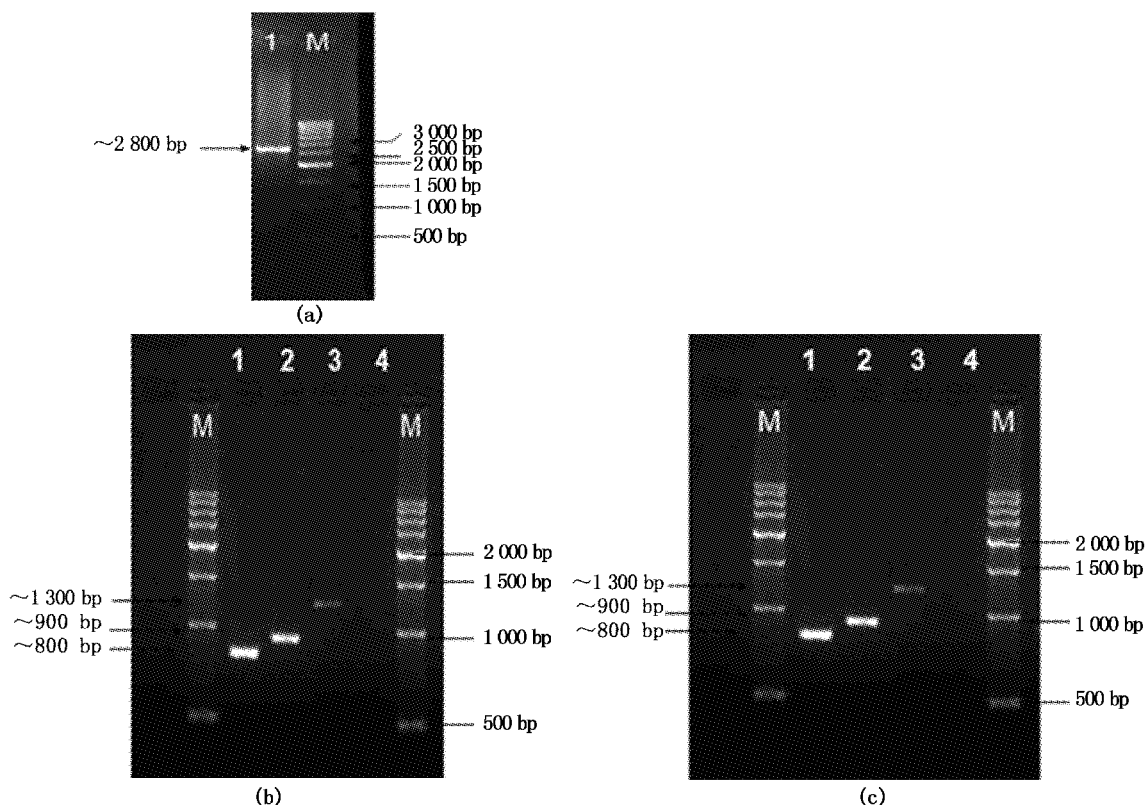


图1 稀有鮡鲫角蛋白 18 基因片段克隆结果和 5'上游侧翼序列及 3'下游侧翼序列基因组步移结果

Fig.1 The keratin 18 DNA segment of *Gobiocypris Rarus* and the PCR result of 5' and 3'-flanking sequence genome walking

M. 500 bp Marker ;图(a)中泳道 1 表示稀有鮡鲫角蛋白 18 基因片段;图(b)和图(c)中,泳道 1,2 ,3 ,4 分别表示 DnaI,EcoRV,Pvu II,和 Stu I 文库的 PCR 扩增产物。

再根据所获得的 2 779 bp 的稀有鮡鲫角蛋白 18 基因片段设计引物 F1、F2、R1 和 R2,并进行基因组双向步移实验。得到结果如图 1(b)和图 1(c)所示。上游侧翼序列基因组步移所得结果如图 1(b)所示,在 1,2 泳道都出现单一明亮条带,选取第 2 泳道 900 bp 左右 DNA 片段进行测序,得到一条 885 bp 的序列。下游侧翼序列基因组步移结果如图 1(c)所示,在 1,3 泳道都出现单一明亮条带,选取第 3 泳道 1 500 bp 左右 DNA 片段进行测序,得到一条 1 493 bp 的条带。再将双向步移所得的结果与前面所得的 2 779 bp 的 DNA 片段进行拼接,最后得到了一条 3 765 bp 长的稀有鮡鲫角蛋白 18 基因序列。

2.2 角蛋白 18 基因的序列分析

根据所获得的稀有鮡鲫角蛋白基因 3 765 bp 的序列进行 Blastn 的同源性搜索。所得到序列中 1 bp 到 1 038 bp 与斑马鱼角蛋白 18 DNA 序列有 84% 的序列相似性,1 914 到 2 804 bp 与斑马鱼

23 号染色体的部分 DNA 序列有 84% 的序列相似性,2 923 到 3 765 bp 与斑马鱼 23 号染色体的部分 DNA 序列有 81% 的序列相似性。此斑马鱼 23 号染色体的部分序列经 GenBank 查找,属于斑马鱼角蛋白 18 基因的序列。1 039 bp 到 1 914 bp 部分和 2 805 bp 到 2 922 bp 共 992 bp 相似性程度较低,可能为内含子区。由此可知,在 3 765 中,有 73.6% 的稀有鮡鲫角蛋白 18 DNA 序列(2 769 bp)与斑马鱼的角蛋白 18 DNA 序列有相似性,表明所得序列为稀有鮡鲫角蛋白 18 基因序列。

角蛋白 18 基因序列的 GC 含量为 42%。A、C、G、T 的含量分别为 30.09%,20.61%,21.49%,27.81%。

2.3 稀有鮡鲫角蛋白 18 基因序列开放阅读框的分析预测

将稀有鮡鲫角蛋白 18 基因序列(3 765 bp)进行开放阅读框的预测,结果表明稀有鮡鲫角蛋

白 18 共有 7 个外显子,6 个内含子,共编码 367 个氨基酸(表 1),外显子与内含子的具体分布如图 2 所示。本预测所得的结果可以从侧面验证

2.2 中所推测的 1 039 到 1 914 bp 以及 2 805 到 2 922 bp 的两个 DNA 片段的共 992 bp 为内含子区域(表 1)。

表 1 稀有鮡鲫角蛋白 18 基因外显子与斑马鱼角蛋白 18 基因外显子的比较

Tab.1 The comparison between the exons of *Gobiocypris rarus* keratin 18 gene and zebrafish keratin 18 gene

| | 稀有鮡鲫 外显子范围(长度)/bp | 斑马鱼 外显子范围(长度)/bp | 稀有鮡鲫 内含子长度/bp | 斑马鱼 内含子长度/bp |
|---|----------------------|---------------------|------------------|-----------------|
| 1 | 89 ~ 307(219) | 1 ~ 509(508) | 92 | 1 710 |
| 2 | 399 ~ 527(129) | 2 219 ~ 2 305(86) | 1 424 | 84 |
| 3 | 1 951 ~ 2 033(83) | 2 389 ~ 2 546(157) | 103 | 164 |
| 4 | 2 136 ~ 2 292(157) | 2 644 ~ 2 808(164) | 115 | 80 |
| 5 | 2 407 ~ 2 571(165) | 2 888 ~ 3 020(132) | 92 | 88 |
| 6 | 2 663 ~ 2 788(126) | 3 108 ~ 3 329(221) | 175 | 88 |
| 7 | 2 963 ~ 3 187(225) | 3 417 ~ 3 898(481) | | |

3 讨论

3.1 稀有鮡鲫角蛋白 18 基因序列与斑马鱼角蛋白 18 基因序列的比较

将斑马鱼角蛋白 18 基因的 mRNA 序列与斑马鱼角蛋白 18 基因全序列进行比较,发现斑马鱼也有 7 个外显子和 6 个内含子,与稀有鮡鲫的角蛋白 18 基因结构相似(表 1)。将稀有鮡鲫和斑马鱼角蛋白 18 基因两条序列均去除其内含子部分,分别得到了 1 104 bp(稀有鮡鲫)和 1 293 bp(斑马鱼)的角蛋白 18 基因编码区,两者比较后发现:其序列相似性达到 91%。另外还发现,与斑马鱼相比,稀有鮡鲫角蛋白 18 基因编码区存在一段缺失序列,长度为 90 bp,可能导致 30 个氨基酸的缺失,相对应的缺失 DNA 序列位于第 1 个外显子和第 2 个外显子之间。

由于斑马鱼的角蛋白 18 基因全长 3 898 bp,包含 7 个外显子和 6 个内含子,编码 431 个氨基酸,位于斑马鱼第 23 号染色体上。而本研究获得的角蛋白 18 基因长为 3 765,也包含 7 个外显子和 6 个内含子,但是却编码 367 个氨基酸,除了中间缺失 30 个氨基酸,末尾也相差了 64 个氨基酸,经验证中间相差的 30 个氨基酸,确实是稀有鮡鲫缺少,而斑马鱼特有的。在对 DNA 序列分析发现,在第 7 个外显子后面出现终止密码子 TAA,以及之后在 3' 末端,有多聚腺苷酸信号 AATAAA^[10],因此可以基本确定本研究所获得的角蛋白 18 基因序列就是稀有鮡鲫角蛋白 18 基因的全长序列。

3.2 角蛋白的特性与功能

角蛋白属于中间纤维蛋白的 I 型和 II 型,最为复杂,也是上皮细胞分化的标记。按表达类型,角蛋白可分为 E 型角蛋白和 S 型角蛋白,其中 E 型角蛋白是上皮细胞的角化蛋白。S 型角蛋白是形成简单的上皮细胞组织的蛋白。含有 S 型角蛋白的上皮组织有上皮细胞、前肾管、消化管、脊动脉管和皮肤等。例如角蛋白 18 和角蛋白 8 都属于 S 型角蛋白^[8-9]。而本研究中的角蛋白 18 主要在鱼类的背动脉管和前肾管上皮细胞中表达^[8]。另外角蛋白的两种类型——I 型角蛋白和 II 型角蛋白,在角蛋白纤维的组装都需要 I 型和 II 型的 1:1 的组合。在哺乳动物中,至少有 30 个基因编码角蛋白。一般地,每种上皮细胞都表达一个特异的 I 型和 II 型结合的角蛋白,如角蛋白 18(k18)和角蛋白 8(k8)在许多上皮细胞中共同表达,并且形成必需的混合物,然后组装成角蛋白纤维^[11]。目前已有很多鱼类的角蛋白 18 的 cDNA 被成功克隆,包括虹鳟鱼、金鱼、鲤、斑马鱼和鲨鱼^[8]等,用于各项研究。

本研究也成功克隆了稀有鮡鲫角蛋白 18 DNA 序列,根据序列预测所得的稀有鮡鲫角蛋白 18 氨基酸序列与斑马鱼的角蛋白 18 氨基酸序列,具有 92.86% 的相似性,与哺乳类人的角蛋白 18 氨基酸序列比对具有 54.14% 相似性,与爪蟾的角蛋白 18 氨基酸序列比对发现具有 52.34% 的相似性。由此可见角蛋白在进化中相对比较保守。

1~50 **T**TGCACTCCTCTGGCTGTGTGTAACGAGTTCTACCTCTCTCTCTCAG
 51~100 ⁺¹CCTTCATTCGCTCATCTGCTCACTCACTTCCAGCGACC**ATG**AGTATGAGA
 101~150 ACAAGCTACAGTGTGCGTTCCTCCACCTCCCAGGTGCCGGTATCCAGAT
 T S Y S V R S S T S Q V P V S Q M
 151~200 GTCCCAGATGTCTATCAAGCGTACCACCAATGTGCCAACTTACCGGGCTG
 S Q M S I K R T T N V P T Y R A
 201~250 CGAGCACCTATGGCGGCTGGAGGTCAAGGCACCCGTATCTCCTCTGCC
 A S T Y G G A G G Q G T R I S S A
 251~300 TCCTACTCAGGTGTCCGCAGTGGAAATGGGACTTCCCTCAATGTCCAGCTC
 S Y S G V R S G M G L P S M S S S
 301~350 CATCCAT**G**TGAGTGCCACCGGAGCCACAGGTGAGATCATGGCAATGAGA
 I H
 351~400 AGATGGCCCATGCAGAATTTGAACGACCGTCTGGCCTCCTATCTGGAGAA
 K
 401~450 AGTGAGGACCCTGGAGCAGGCCAACAGCAAGCTGGAGCTGAAGATCCGTG
 V R T L E Q A N S K L E L K I R
 451~500 AGGCCCTGGAGAAGAGAGGTCTGTATGTCCAGACTACAGCCGTTCCAG
 E A L E K R G P D V H D Y S R F Q
 501~550 CCCATCGTTGACGAGCTGCGCAAGGAG**G**TTAGTTGAATTGTTGGTACCG
 P I V D E L R K E
 551~600 ACAATTTATGGCTCTATAAAGAACTTAATGTGACAGGCTAATTGTGCTG
 601~650 CTTTGTGTTTATTAACAGTTTAGTGGATATTTATTGGATTTCAGGACAGA
 651~700 AACTAACAAACCAAAGGCATTTCTTTGTCAATTAAGTTGGATTTATTTA
 701~750 AAAGGGCTGAATGAGTTTACTGATGGCCAGCACTGATAGGCAGAGAACCA
 751~800 AAAAAATATGGGTGAAACTAAAAAAGGCATGGGGGAGGAGGTCTAAGTCA
 801~850 CTACAGGACGGCAGTCATTACTTCTGCTTACAGCCGTAATTCGGTGAAT
 851~900 AAGCATTTCCATCCTGTATTTATCCTGTCTGTTACACATCCAAATATT
 901~950 TGTTTATTCCTGTGGATGGAAACTCTGTGATCTGAGATGTTCAAGCATC
 951~1000 TGTGTGTGTGTGTCATAAGGGAAATGGAAAAGTCGAGGGAGTTCAGAC
 1001~1050 AGTGCAAAATCGGCCAGCTGTGAGACTTTTCTGAAAATGGCTGAACACT
 1051~1100 CTCAGAGATATTTGTGCTATTCTACATTTTCTTGCAAAGTTTGAGTTCAT
 1101~1150 AAAGCAGACACTGTTAAAGATAAAATACTGATTAGAGTCACCAAATGTTT
 1151~1200 CTTGAAGCTGCAATTTATAGCTAATATGTCACCTGATAACCAGACAGTGTT
 1201~1250 ATCAGAATGATTGGTTTTTCTTTCATCAGTCACAATGCAGGAAAGTGACA
 1251~1300 GGTAGATTAATAAACAATTAGATCCCTATTCCTCAGGGTTTTCTACA
 1301~1350 GTCACAAAATAAATGGGATTAGCCATGAGATTAACCAGATGAAAAGCTA
 1351~1400 GATGGCGTTTTGGCACTGAAAGGAATAGTCTAAAGGTTAAAAACTCAGTG
 1401~1450 AAGTACTGTTTATATAACAATAATAAATAAAAAAGTATAATATGTATAAAA
 1451~1500 CATACTACACATTTAAGTAGAACTATGTGTTAAAAGATACATAGCCAACAC
 1501~1550 CCCTTAGCATATGGAAAATTGTATTGAGCTCAGTAGTGTACATCTCCAC
 1551~1600 CCTTTCCATGATGACAGGATATGCAAACTGGCAAACATGGAAAGCATTT
 1601~1650 ACACAACAAAGTACATCAGGTTAGAAAAAACAAGGAAATCCAACACTCAA
 1651~1700 CCTCTGTCTTTGACACTACTCAACTAATTTACGACCGTCAAACAGACAC
 1701~1750 CAGAAGCTTTTCTAACGTGGCTGTTGGCTAATGATTATCCTACGAACTGT
 1751~1800 GAAAAGGACTTTGTCTCCAGTACCTCTTATCTAAGCAAAAACAAGGCGTC
 1801~1850 AGGGCAGAAGATGGTCTGTTTACACATCATGTACTACAATTTAAGGGT
 1851~1900 TAGGTGAGGTTTTATGTGTGATAATTACAGTTTTAATGGTTTCCTGACTC
 1901~1950 TTTCTGTGAACCTTTTGACTGATTCAAATGATATCTGTTTTGGTTGTAG
 1951~2000 ATCTTTGATGCCACCACAATAATGCCCGTCTAGTGCTCCAGATTGATAA
 I F D A T T N N A R L V L Q I D N
 2001~2050 CGCCCGTCTGGCAGCTGATGACTTCAGAGTCAAG**T**GAGTAACAAAATTC
 A R L A A A D D F R V K

| | |
|-----------|--|
| 2051~2100 | AAAGTATAACAAACATTTTAACTTCATGGAGAGTTTCTAGTTTTGAAGA |
| 2101~2150 | GAACAGCGACCCATTGACCTATGTTGTGCCTGCAGGTATGAATCCGAGCT Y E S E L |
| 2151~2200 | GTCTATCCGCCAAGGTGTGGAGGCTGACATTGTCGGCTTGAGAAAAGTCA S I R Q G V E A D I V G L R K V |
| 2201~2250 | TTGATGACACCAACATGAACCGAATGAATCTTGAAAGCGAGATTGAGGCC |
| 2251~2300 | I D D T N M N R M N L E S E I E A CTCAAGGAAGAGCTCATCTTCCTTAAGAAGAACCATGACAATGTGAGTAC L K E E L I F L K K N H D N |
| 2301~2350 | ACATCTGGCACATTTGGACATTAGCACATCTGGAACATTTGTCTACCAC |
| 2351~2400 | ATAAACCATATTGGCCGTTGCTTAAAGACATATTTTATTGTCTCTCCACC |
| 2401~2450 | CAACAGGAGGTAATGGAGCTTCGCAACCAGATCTCCATGTCCGGAGTGCA E V M E L R N Q I S M S G V Q |
| 2451~2500 | GGTGGATGTTGATGCTCCTAAGGGACAGGATATTGCCAGCTCATGGAGG V D V D A P K G Q D I A Q L M E |
| 2501~2550 | AGATGAGAGCCAAGTACGAGAAGATGGCCCTAAAGAACCAGGAGGAGTTG E M R A K Y E K M A L K N Q E E L |
| 2551~2600 | AAGAACTGGCATGAATCTCAGGTATATTCCAAAAGCCAATCATTTTCTTG K N W H E S Q |
| 2601~2650 | ATGTAGTCTTGCCATCAATGCCTTTGTTCTGGGGCTAATGTGAAATTTTG |
| 2651~2700 | CGCTCTCTCTAGATCACAGAGATTCAGGTACAAGTCACACAGAACACAGA I T E I Q V Q V T Q N T E |
| 2701~2750 | AGCCCTCCAGGGTCTCGTACAGAAGTCAACGAACTTCGCAGACAGATTC A L Q G A R T E V N E L R R Q I |
| 2751~2800 | AGACATTGGAGATTGAGCTGGAGTCACAGAGGAACCTGGTGAGTTGTCT Q T L E I E L E S Q R N L |
| 2801~2850 | TCTTCATTGAAAAGCTACCATTTTCCATCTGGTCTTTTCAATCCATAC |
| 2851~2900 | TCATGTTAAGATGTATTTTTTAAATATCGTGTCTGAATTTGTCTTGAG |
| 2901~2950 | AATTGTCTTAATGAAGTTGGCTCAAACCTAACTCCTGGGTTCCCTTTGC |
| 2951~3000 | TTCATTCTCCAGAAAGGATCCCTGGAAGCCACCCTGCGGGATACGAAAT K G S L E A T L R D T E M |
| 3001~3050 | CGGTTACAACATGGAGATCGAGAGTATCAACACCATTGTCTGCAACTGG R Y N M E I E S I N T I V L Q L |
| 3051~3100 | AGTCTGAGCTCACAAATCTGCGCAACAACATCCAACACCAGACACAGGAG E S E L T N L R N N I Q H Q T Q E |
| 3101~3150 | TACGAGGCTCTGCTGAACATCAAGATGAAGCTGGAGGCAGAAATCGCCAC Y E A L L N I K M K L E A E I A T |
| 3151~3200 | CTACAGGAGGCTTCTCGATGGTGGAGACTTCAAGTAAAGTGTCTTCTGAA Y R R L L D G G D F K |
| 3201~3250 | ACATACAAAGTTTGACAGCAGTTCACCATTGTGTAAACATGTAATGTG |
| 3251~3300 | GTTTTGTTTTGGTTGGTAACAAGCATGTTTTCTCCTGCAGGCTCCAGGAT |
| 3301~3350 | GCTCTTGATGAGCAAAAAGGTTAAAGTATGACTGTACGCAGACACT |
| 3351~3400 | GGTGGATGGGAAGGTGGTGTCTTCAAGCACAGAGACCAAAGAGAGGAAAC |
| 3401~3450 | TCTGAACATCAAATTCAGATCAACCTCTACAACACAACAGAAAAACCTT |
| 3451~3500 | ATGGCCAATTATAAAATTATCTAGACAGGTTGGACTTAACACATTAGAT |
| 3501~3550 | ATTTATTTAGTCCAAAGCGCCTTAATTCAGGGGTCTCAGAGCACATATGC |
| 3551~3600 | AATGAGATTTGACTAGCACATTTCAACACTAAAAAACACATTTAAAAATT |
| 3601~3650 | TAAATGACTATAGCCAAAAAGACTGAAAGTCAAAGAACACTATTGAAAC |
| 3651~3700 | TTTTTACATATTTTGTAAATGACAATGAATGATTGTTTACGCCAAATCA |
| 3701~3750 | CCAATTGTCCATTTGTTTCTATGTTGTGTTGTTTCATCATGTCTTGT |
| 3751~3765 | CTTTTCCATAAAAAG |

图2 稀有鳊角蛋白18基因组序列及编码氨基酸序列

Fig. 2 Keratin 18 genome sequences and coded amino acid sequences of *Gobiocypris Rarus*

图中阴影部分和 +1 表示转录起始位点; 方框表示翻译起始位点; 下划线表示内含子的 GT-AG 法则位点; 方框加阴影表示终止密码子; 下划线加阴影表示多聚 A 尾。

3.3 基因组步移技术研究方法

本研究获得稀有鮕鲫角蛋白没有采用常用的从 RNA 反转录为 cDNA 为模板获得基因全长,而是直接从 DNA 为模板出发,采用基因组步移的方法获得。基因组步移的方法在克隆基因序列中已经很成熟,如生秀杰等利用基因组步移克隆小鼠 Doc-1R 基因组序列^[10],孙玉英等利用基因组步移技术克隆芽孢杆菌 *Bacillus* sp. S-1 壳聚糖酶基因^[12]。巢伟等利用基因组步移技术扩增嗜水气单胞菌 J-1 株脂酶基因全长^[13]。实践证明,使用基因组步移的方法获得基因全长^[11-28]的技术越来越成熟,大大简化了克隆全长基因的步骤,提高了实验效率。

感谢上海海洋大学邹曙明教授对本实验研究的指导与帮助。

参考文献:

- [1] 王剑伟. 稀有鮕鲫的繁殖生物学[J]. 水生生物学报, 1992, 16(2): 165-174.
- [2] 曹文宣, 王剑伟. 稀有鮕鲫——一种新的鱼类实验动物[J]. 实验动物科学与管理, 2003, 20(s): 96-99.
- [3] 王铁辉, 刘沛霖, 陈宏溪, 等. 稀有鮕鲫对草鱼出血病病毒敏感性的初步研究[J]. 水生生物学报, 1994, 18(2): 144-149.
- [4] 贾方钧, 魏芸. 稀有鮕鲫的染色体核型初报[J]. 水生生物学报, 2001, 25(4): 425-426.
- [5] 周永欣, 成水平, 胡炜, 等. 稀有鮕鲫——一种新的鱼类毒性试验材料[J]. 动物学研究, 1995, 16(1): 59-63.
- [6] 彭扣, 王玉凤, 刘绪生, 等. 稀有鮕鲫 β -肌动蛋白基因片段的克隆及其同源性分析[J]. 华中师范大学学报: 自然科学版, 2006, 40(4): 580-584.
- [7] 叶斐菲, 刘红, 蔡生力, 等. 稀有鮕鲫 β -肌动蛋白启动子的克隆及其驱动活性的初步检测[J]. 渔业科学进展, 2010, 31(2): 80-87.
- [8] WANG Y H, CHEN Y H, LIN Y J, et al. Spatiotemporal expression of zebrafish keratin 18 during early embryogenesis and the establishment of a keratin 18: RFP transgenic line [J]. *Gene Expression Patterns*, 2006, 6(4): 335-339.
- [9] WOO S J, EUN J K, HYUN J R, et al. Expression of a novel type I keratin, DAPK-1 in the dorsal aorta and pronephric duct of the zebrafish embryos[J]. *Gene*, 2003, 312: 145-150.
- [10] 生秀杰, 姜莉, 周伟强, 等. 应用基因组步移克隆小鼠 Doc-1R 基因组序列[J]. 中华医学遗传学杂志, 2001, 18(4): 314-316.
- [11] GONG Z Y, JU B S, WANG X K, et al. Green fluorescent protein expression in germ-line transmitted transgenic zebrafish under a stratified epithelial promoter from keratin 8 [J]. *Developmental Dynamics*, 2002, 223(2): 204-215.
- [12] 孙玉英, 张继泉, 王淑军, 等. 基因组步移技术克隆芽孢杆菌 *Bacillus* sp. S-1 壳聚糖酶基因与序列分析[J]. 中国生物工程杂志, 2009, 29(5): 72-77.
- [13] 巢伟, 刘永杰, 陆承平, 等. 基因组步移技术扩增嗜水气单胞菌 J-1 株脂酶基因全长及原核表达[J]. 动物医学进展, 2010, 31(9): 1-6.
- [14] 朱志明, 陈晖, 李盛霖, 等. 半番鸭 POU1F1 基因序列的克隆与生物信息学分析[J]. 福建畜牧兽医, 2010, 32(1): 1-4.
- [15] 李竹红, 刘德培, 梁植权. 改进的反向 PCR 技术克隆转移基因的旁侧序列[J]. 生物化学与生物物理进展, 1999, 26(6): 600-602.
- [16] 蹇文婴, 东秀珠. 利用反向 PCR 方法扩增细菌热激蛋白 HSP60 基因[J]. 微生物学报, 2002, 42(1): 56-61.
- [17] 刘博, 苏乔, 汤敏谦, 等. 应用于染色体步移的 BCD 扩增技术的研究进展[J]. 遗传, 2006, 28(5): 587-595.
- [18] 简清. 斑马鱼 Mylz2 启动子的克隆和转荧光蛋白基因斑马鱼、唐鱼的构建[D]. 上海: 上海水产大学, 2005.
- [19] 卫正国, 陈玉华, 李兵, 等. 家蚕 P450 基因 CYP305B1 的基因组序列克隆及结构分析[J]. 蚕业科学, 2009, 35(1): 144-147.
- [20] 郭旭东, 毛舒燕, 宝明涛, 等. 绵羊角蛋白关联蛋白 KA P621 基因 5'端调控区的分子克隆及测序结果比较[J]. 内蒙古大学学报: 自然科学版, 2007, 38(1): 58-63.
- [21] 郭旭东, 侯冬霞, 毛舒燕, 等. 小鼠超高硫角蛋白基因启动子的分子克隆及序列分析[J]. 科技通报, 2010, 23(4): 471-474.
- [22] 李薇, 寻萌, 楚雍烈, 等. 人细胞角蛋白 8 (CK8) 基因的克隆及其在大肠杆菌中的表达[J]. 细胞与分子免疫学杂志, 2010, 26(1): 41-43.
- [23] 王卫芳, 邱丽华, 黄建华, 等. 斑节对虾亲环素 A (cyclophilin A) 基因的克隆及启动子序列分析[J]. 中国水产科学, 2009, 16(3): 324-330.
- [24] 彭书明, 雷泞菲, 陈放, 等. 苦瓜中与 AGAMOUS 相似基因启动子的克隆和表达载体的构建[J]. 植物生理学通讯, 2008, 44(4): 693-698.
- [25] 袁磊, 孙红炜, 杨崇良, 等. 转基因玉米 MON88017 旁侧序列分析及定性 PCR 检测[J]. 作物学报, 2010, 36(2): 361-364.
- [26] 罗鹏, 胡超群. 溶藻弧菌溶血素基因反向 PCR 克隆及其原核表达[J]. 水产学报, 2009, 33(3): 410-415.
- [27] 何晓兰, 刘桂华, 倪万潮, 等. 三角叶滨藜甜菜碱醛脱氢酶基因侧翼序列的克隆与分析[J]. 江苏农业学报, 2004, 20(2): 65-69.
- [28] 尹俊, 李金泉, 张燕军, 等. 一个山羊 I 型毛角蛋白基因的序列及其在皮肤中的表达[J]. 畜牧兽医学报, 2006, 37(1): 18-22.

Cloning and analysis of keratin 18 gene from fish (carp) *Gobiocypris rarus*

YE Ying, CAI Sheng-li, LIU Hong

(Key Laboratory of Aquatic Genetic Resources and Aqua-cultural Ecosystem Certified by the Ministry of Agriculture, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Cytokeratin 18 (K18) is an 'S' keratin and is the first of the intermediate filament proteins to be expressed during mammalian and amphibian development. According to the published conserved DNA sequences of keratin 18 from human, xenopus and zebrafish, two pairs of primers were designed to obtain 2 779 bp of keratin 18 gene derived from the genome DNA of *Gobiocypris rarus* by PCR. And then through the genome walking upstream and downstream sequences were also amplified. And then the full amplified fragment of keratin 18 was inserted into *Escherichia coli* cloning vector pMD19-T and sequenced. The result indicated that the keratin 18 of *Gobiocypris rarus* was successfully obtained about 3 765 bp. The analysis of NCBI's blastn showed that comparing *Gobiocypris rarus* with zebrafish, the keratin 18 DNA sequence was highly conserved with 73.6% similarity. Finally, biologically analysis and ORF prediction showed that the keratin 18 DNA sequence contained 7 exons and 6 introns, and coded for a protein with 367 amino acids.

Key words: *Gobiocypris rarus*; keratin 18; cloning; genome walking

2012 年《中国水产科学》征订启事

《中国水产科学》为中国水产科学研究院主办,科学出版社出版的学术性期刊,目前已成为中国水产科学研究领域的重要学术期刊。本刊在促进中国的水产科学研究、加强国际间学术交流、展示中国水产科学研究领域最新科研成果与研究进展等方面发挥了重要作用。期刊影响因子逐年递增,2008 年中国科技期刊引证报告统计的影响因子值为 0.994, CNKI 期刊统计源影响因子为 1.143; 期刊多次获得“中国百种杰出学术期刊”奖。

本刊主要报道水产生物学基础研究、水生生物病害及其防治、水产生物营养及饲料、渔业生态保护及渔业水域环境保护、水产品保鲜与加工综合利用、水产资源、海淡水捕捞、水产养殖与增殖以及设施渔业等方面的最新进展、最新成果、最新技术和方法。

本刊为双月刊, A4 开本, 每期 200 页, 单月出版, 国内外公开发行。

国内定价 60 元/期, 全年 360 元(含邮费)。邮发代号: 18-250,

国内统一刊号: CN 11-3446/S, 国际标准刊号: ISSN 1005-8737, 国外代号 4639Q。

直接向编辑部订阅可享受 8 折优惠, 也可在当地邮政局(所)办理订阅手续(可破季订阅)。

漏订或补订当年和过期期刊, 请直接向编辑部订阅。

编辑部地址: 北京市丰台区青塔村 150 号(中国水产科学研究院内)

邮政编码: 100141

联系电话: 010-68673921, 010-68673931

传 真: 010-68673931

E-mail: zgscckx@cafs.ac.cn; jfishok@publica.bj.cninfo.net

投稿与查询网址: www.fishscichina.com