

草鱼 *Smad4* 基因的克隆、生物信息学分析及反义真核载体的构建

吕池波^{1,2}, 郁二蒙¹, 谢 骏¹, 王广军¹, 余德光¹, 李志斐¹, 王海英¹, 龚望宝¹

(1 中国水产科学研究院 珠江水产研究所, 农业部热带亚热带水产资源利用与养殖重点实验室, 广东 广州 510380; 2 上海海洋大学 水产与生命学院, 上海 201306)

摘要:为了研究 TGF- β 信号传导中的关键蛋白 Smad4 在草鱼 *Ctenopharyngodon idellus* 发育过程中的作用, 克隆出草鱼 *Smad4* 基因 cDNA 全序列, 其碱基长度为 2 974 bp, 其中开放阅读框 1 644 bp, 编码 547 个氨基酸. 生物信息学分析结果显示, 草鱼 Smad4 蛋白相对分子质量为 59 690.3, 分子式为 $C_{2\,632}H_{4\,073}N_{753}O_{791}S_{24}$, 理论等电点为 6.5, 含有 MH1 和 MH2 这 2 个高度保守的功能结构域. BLAST 相似性分析显示, 草鱼 Smad4 蛋白的氨基酸序列与鲤鱼 *Cyprinus carpio*、斑马鱼 *Danio rerio* 的相似性较高, 分别为 94.23% 和 92.97%. 还将草鱼 *Smad4* 基因的开放阅读框片段反向插入表达载体 pcDNA3.1(+), 成功构建了反义 *Smad4* 基因真核表达载体 pcDNA3.1(+)-Anti-*Smad4*.

关键词:草鱼; Smad4; 反义基因; 真核表达载体; 胚胎发育; 生物信息学分析
中图分类号:Q78; S917.4 **文献标志码:**A **文章编号:**1001-411X(2013)04-0517-07

Cloning of *Smad4*, Bioinformatics Analysis and Construction of Antisense *Smad4* Gene Expression Vector from *Ctenopharyngodon idellus*

LÜ Chibo^{1,2}, YU Ermeng¹, XIE Jun¹, WANG Guangjun¹, YU Deguang¹,
LI Zhifei¹, WANG Haiying¹, GONG Wangbao¹

(1 Key Laboratory of Tropical & Subtropical Fishery Resource Application & Cultivation, Ministry of Agriculture, Pearl River Fisheries Research Institute of Chinese Academy of Fishery Science, Guangzhou 510380, China;
2 College of Fisheries and Life, Shanghai Ocean University, Shanghai 201306, China)

Abstract: Smad4 is a key protein in TGF- β signal transduction, which can regulate the developmental processes like cell proliferation and differentiation. In order to study the role of Smad4 in developmental process of grass carp (*Ctenopharyngodon idellus*), a full length of 2 974 bp of *Smad4* cDNA sequence from grass carp was obtained, a 1 644 bp CDS (coding sequence) encoding a polypeptide of 547 amino acids was determined in the Smad4 of grass carp. Bioinformatics analysis showed that the relative molecular mass of grass carp Smad4 protein was 59 690.3, the molecular formula was $C_{2\,632}H_{4\,073}N_{753}O_{791}S_{24}$, isoelectric point was 6.5, and it contained both MH1 and MH2 highly conserved functional domains. BLAST homology analysis showed higher homology of grass carp Smad4 protein with common carp and zebrafish, which was 94.23% and 92.97% respectively. The open reading frame of grass carp *Smad4* gene was reversely inserted into the expression vector pcDNA3.1(+), which constructed the antisense *Smad4* gene eukaryotic expression vector pcDNA3.1(+)-Anti-*Smad4*.

Key words: *Ctenopharyngodon idella*; Smad4; antisense; eukaryotic expression vector; embryonic development; bioinformatics analysis

收稿日期:2012-10-17 网络出版时间:2013-09-09
网络出版地址:<http://www.cnki.net/kcms/detail/44.1110.S.20130909.1900.007.html>
作者简介:吕池波(1986—),男,硕士研究生;通信作者:谢 骏(1965—),男,研究员,博士,E-mail: xiejunhy@yahoo.com.cn
基金项目:现代农业产业体系建设专项(nycyx-49);公益性行业(农业)科研专项(201203083);广东省自然科学基金(10151038001000004);国家科技支撑计划项目(2012BAD25B04)

在 Smads 蛋白家族中已发现 9 个成员^[1], 根据结构和功能的差异可分为通用型、受体激活型和抑制型 Smads. 通用型 Smads 的主要功能是与受体激活型 Smads 结合形成异源寡聚体, 将信号转至细胞核内, 调节基因表达^[2], 目前仅发现 Smad4 这 1 种. Smad4 是 TGF- β /smads 信号转导通路中的中心转导分子, 它通过形成异源寡聚体, 将信号转导至细胞核内, 从而调控靶基因的转录并诱导下游基因的表达^[3], 因此它的表达情况与生长、发育等生命过程密切相关. 在鱼类研究中, 只有斑马鱼 *Danio rerio*、鲤 *Cyprinus carpio*、虹鳟 *Oncorhynchus mykiss*、金鱼 *Carassius auratus* 等少数物种的 *Smad4* 基因被克隆. 在鲤体内发现 *Smad4* 基因有不同的组织表达模型, 暗示着 *Smad4* 通过改变组织表达水平来调控 TGF- β 信号通路^[4]. 在斑马鱼胚胎发育研究中发现 *Smad4* 的时空分布广泛且表达量大, 而与之相关的 *Smad3* 基因呈现一种空间限制性表达模式^[5]. 将金鱼的 *Smad4* 结构与 *Smad2*、*Smad3* 及 *Smad7* 比较分析发现, *Smad4* 编码的氨基酸序列中间连接区域少了 42 个氨基酸, 这可能就是它在 LT-2 细胞中失活的原因^[6]. 目前, 有关草鱼 *Smad4* 基因的研究鲜见报道.

研究基因功能常用的方法有基因转导技术、反义技术、基因敲除和敲入、基因诱捕技术等, 其中反义技术利用 DNA 或 RNA 分子通过 Watson-Crick 碱基配对原则与目的基因的 mRNA 互补结合, 通过各种机制使其降解或抑制其编码蛋白的翻译, 从而抑制特定目的基因的表达^[7]. 作为反义技术中的一种, 反义转基因技术已成为下调生物体内源目的基因表达效果的一种有效手段. 其中, 具有延熟保鲜特性的转基因番茄 *Solanum lycopersicum* 的培育是最经典的成功范例^[8]. 采用反义转基因技术导入反义重组基因, 以抑制机体内的基因表达, 可阻断其生理功能. 中国科学院水生生物研究所研究团队利用促性腺激素释放激素的反义基因抑制了转基因鲤鱼中促性腺激素的合成和性腺发育^[9]; 反义桃拉综合病毒外衣蛋白基因转基因虾显著增强了病毒抵抗力^[10]. 在有关 *Smad4* 基因的研究中, 通过构建反义 *Smad4* 真核表达载体可以抑制 *Smad4* 蛋白的表达, 从而减少 I 型胶原的表达量^[11]. 为此, 本研究拟通过 PCR 方法和 RACE 技术获得草鱼 *Smad4* cDNA 全序列, 并将开放阅读框 (Open reading frame, ORF) 序列反向插入真核表达载体 pcDNA3.1 (+) 构建 *Smad4* 基因反义真核表达载体 pcDNA3.1 (+)-Anti-*Smad4*, 期望为进一步研究 *Smad4* 基因在草鱼胚胎发育过程中的生理机制打下基础.

1 材料与方法

1.1 试验材料

2010 年 5 月在广东省中山食品水产集团有限公司南头养殖场采集 3 尾体质健康的草鱼, 体质量为 (2.0 ± 0.10) kg, 去鳞解剖后, 采集肌肉组织, 并立即置于液氮中保存, 运回实验室, 置于 -40°C 冰箱冷冻保存备用.

1.2 主要试剂

总 RNA 提取试剂盒 (百泰克公司); DNA Marker DL2000、PrimeScriptTM RT-PCR Kit、pMD19-T 载体、DH5 α 、LA Taq DNA 聚合酶、EcoR I、Xho I 和 T4 DNAligase (大连 TaKaRa 公司); SMARTTM RACE cDNA Amplification Kit (Clontech 生物工程公司); Plasmid Mini Kit I、Gel Extraction Kit (OMEGA 公司).

1.3 试验方法

1.3.1 总 RNA 提取与 cDNA 的合成 RNA 提取参照百泰克公司总 RNA 提取试剂盒说明书进行, 利用琼脂糖凝胶电泳检测 RNA 纯度, 然后使用微量分光光度计测总 RNA 浓度. 按照 TaKaRa PrimeScriptTM RT-PCR Kit 说明书进行反转录合成 cDNA: Oligo dT Primer 1 μL 、总 RNA 4 μL 、dNTP Mixture 1 μL 、RNase Free ddH₂O 4 μL , 反应程序为 65°C 5 min, 4°C 条件下保存; 然后加入 $5 \times$ PrimeScript II Buffer 4 μL 、RNase inhibitor 0.5 μL 、PrimeScript I RTase 1 μL 和 RNase Free ddH₂O 4.5 μL , 反应程序为 42°C 30 min, 95°C 5 min, 4°C 条件下保存.

1.3.2 草鱼 *Smad4* 基因 cDNA 全序列克隆 以反转录合成的 cDNA 为模板, PCR 扩增草鱼 *Smad4* 基因核心序列片段. PCR 反应体系 25 μL : ddH₂O 19.7 μL 、PCR Buffer 2.5 μL 、 $4 \times$ dNTP Mixture 0.5 μL 、cDNA 1 μL 、*Smad4* F (5'-TGTGTGACCATACAGAGAAC-3') 0.5 μL 、*Smad4* R (5'-TCCAGCAGGGCGTCTCTT-TA-3') 0.5 μL 、LA Taq 0.3 μL . PCR 反应条件: 预变性 94°C 5 min; 94°C 1 min, 65°C 30 s, 72°C 90 s, 30 个循环; 然后 72°C 延伸 8 min. PCR 产物经 $10 \text{ g} \cdot \text{L}^{-1}$ 琼脂糖凝胶电泳检测后回收并克隆、测序, 所得的 cDNA 片段与 GenBank 数据库中的已有物种的 *Smad4* cDNA 序列进行 BLAST 相似性分析.

3'RACE 试验按照 SMARTTM RACE cDNA Amplification 说明书进行. 用上述的逆转录反应液、 $5 \times$ PCR Buffer 10 μL 、ddH₂O 28.75 μL 、TaKaRa Ex Taq[®] HS 0.25 μL 、*Smad4*-3' RACE (5'-CCTGCTATCAGTTT-GTCT-3') 0.5 μL 、M13 Primer M4 0.5 μL 进行首次 PCR 反应, 反应条件: 94°C 预变性 3 min; 94°C 30 s, 55°C 30 s, 72°C 1 min, 共 30 个循环; 72°C 延伸 10

min. 巢式 PCR 使用的引物为 M13 Primer M4 和 *Smad4*-3' 2RACE (5'-ACCCACGGCAGAGCATTA-3'). PCR 产物在琼脂糖凝胶中电泳检测后进行胶回收,并克隆测序.

根据前一步克隆的 *Smad4* 基因的 cDNA 序列,设计 2 个下游引物 *Smad4*-5' RACE (5'-GTAGGGGTTCACACAGAC-3') 与 *Smad4*-5' 2RACE (5'-AGCGT-TCTCTGTATGGTC-3'),合成 5' RACE cDNA 方法参照试剂盒说明书进行. 首次 PCR 反应体系:10 × Advantage 2 PCR Buffer 5 μL、dNTP Mix (10 mmol/L) 1 μL、10 × UPM 5 μL、*Smad4*-5' RACE 1 μL、5'-RACE-Ready cDNA 2.5 μL、50 × Advantage 2 Polymerase Mix 1 μL、PCR-Grade water 34.5 μL. 反应程序与 3'RACE 一致,将退火温度改为 65 °C,循环数为 25. 取首次 PCR 产物 5 μL,用 Tricine-EDTA Buffer 稀释 50 倍. 然后取 5 μL 作为模板,引物为 *Smad4*-5' 2RACE 和 NUP 各 1 μL,进行巢式 PCR 反应. RACE 试验产物进行胶回收并克隆测序.

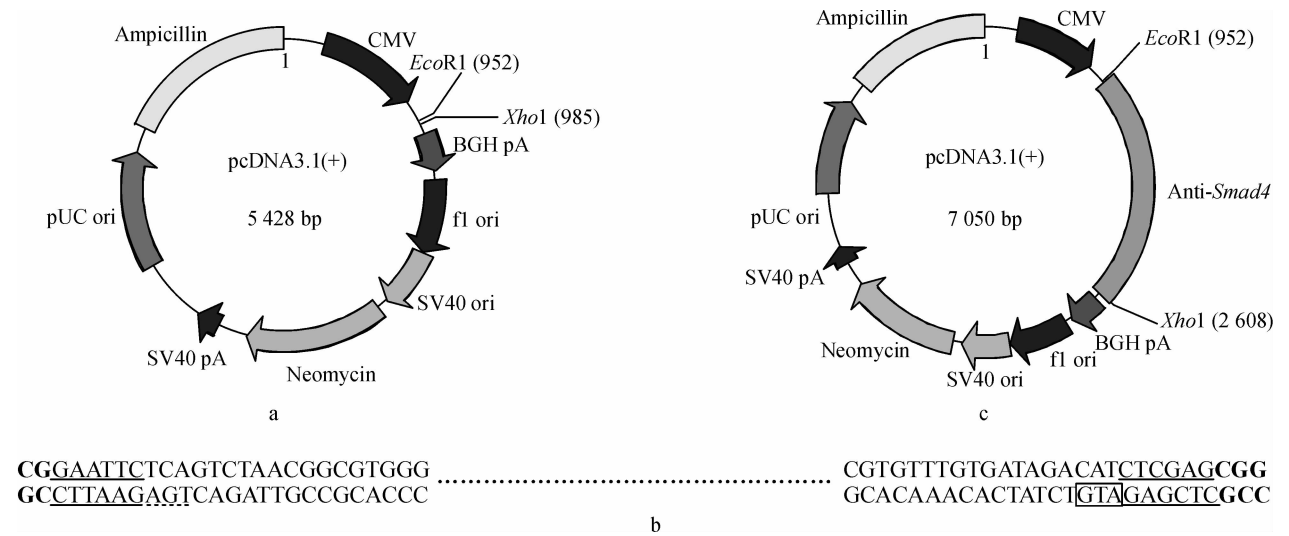
1.3.3 草鱼 *Smad4* 基因的生物信息学分析 *Smad4* 基因的分析、*Smad4* 蛋白的疏水性分析及功能结构域分析在 expasy 网站 (<http://www.expasy.org/>) 完成;蛋白功能位点及二级结构分析在 <https://www.predictprotein.org/> 完成;系统发育树使用 MEGA4.0 软件完成.

1.3.4 *Smad4* 基因 ORF 片段的获得与鉴定 以反

转录合成 cDNA 为模板,设计引物 P1 (5'-ATGTC-TATCACAACACGCCCACC-3')、P2 (5'-TCAGTCTA-ACGGCGTGGGGTC-3'),PCR 扩增 *Smad4* 基因 ORF 片段,产物经凝胶电泳检测后回收并克隆测序.

1.3.5 带酶切位点的 *Smad4* 基因 ORF 片段的获得与鉴定 将上述进行过验证的菌液提取质粒 pMD19-T-Anti-*Smad4*,并以此为模板,P3 (5'-CCGCTCGA GATGTCTATCACAACACG-3')、P4 (5'-CGGAATTCTCAGTCTAACGGCGTGGG-3') (加粗的为保护碱基,带下划线为酶切位点)为引物进行 PCR 扩增,反应体系及反应条件同 1.3.2. 获得两端含有 *Xho* I 和 *Eco*R I 2 个酶切位点的 *Smad4* ORF 片段. PCR 产物经凝胶电泳检测后回收并克隆测序.

1.3.6 反义真核表达载体的构建 如图 1 所示,将 *Smad4* 片段反向插入 pcDNA3.1(+) 载体中. 同时用 *Xho* I 和 *Eco*R I 对质粒 pcDNA3.1(+) 进行双酶切,反应体系:*Xho* I 1 μL,*Eco*R I 1 μL,10 × H Buffer 2 μL,质粒 1 μg,加无菌水至 20 μL. 反应条件:37 °C,2 h. 利用相同的反应体系对 pMD19-T-Anti-*Smad4* 进行双酶切. 产物进行 10 g · L⁻¹ 琼脂糖凝胶电泳,按照 OMEGA 胶回收试剂盒说明书纯化回收目的片段. 连接载体和 *Smad4* 片段,反应体系:10 × T4 DNA Ligase buffer 2 μL,*Smad4* 酶切片段 15 μL,pcDNA3.1(+) 酶切片段 2 μL,T4 DNA Ligase 1 μL. 16 °C 连接 15 h,转化并将菌液送至上海英骏生物技术有限公司测序.



a:pcDNA3.1(+)载体;b:插入的带酶切位点的反义 *Smad4* ORF 片段(加粗的为保护碱基,带下划线的为酶切位点,带框的为起始密码子,带下划线虚线的为终止密码子);c:构建好的重组载体 pcDNA3.1(+)-Anti-*Smad4*.

图1 反义 *Smad4* 重组载体图谱

Fig. 1 Recombinant vector of antisense *Smad4*

1.3.7 重组载体的 PCR 检测和酶切鉴定 提取重组载体 pcDNA3.1(+) -Anti-*Smad4*,方法按照 OMEGA 质粒提取试剂盒说明书进行. 以 P1、P2 为引物,

抽提的质粒为模板进行 PCR 检测,反应体系和条件同 1.3.2. 用 *Xho* I 和 *Eco*R I 双酶切重组载体 pcDNA3.1(+) -Anti-*Smad4*,酶切产物进行 10 g · L⁻¹ 琼

脂糖凝胶电泳分析. 然后, 参照测序结果筛选阳性的反义表达载体.

2 结果与分析

2.1 草鱼 Smad4 基因 cDNA 全序列及推导的氨基酸序列分析

测序结果显示: 草鱼 *Smad4* cDNA 全长为 2 974 bp, 开放阅读框为 1 644 bp (第 367 ~ 2 010 bp), 编码 547 个氨基酸, 其相对分子质量为 59 690. 3, 分子式为 C_{2 632}H_{4 073}N₇₅₃O₇₉₁S₂₄, 理论等电点为 6. 5, 无信号

肽和跨膜序列. 草鱼 *Smad4* 基因含有 2 个 N-糖基化位点, 2 个蛋白激酶 C 磷酸化位点, 4 个酪蛋白激酶 II 磷酸化位点, 9 个 N-豆蔻酰化位点和 1 个酰胺化位点(图 2).

2.2 草鱼与其他生物 Smad4 蛋白氨基酸序列的相似性比较和进化树分析

草鱼 *Smad4* 氨基酸序列与其他生物的相似性比较高, 其全长序列与斑马鱼、鲤、非洲爪蟾 *Xenopus laevis*、小鼠和人的相似性分别达到了 95. 61%、94. 33%、84. 63%、85. 77% 和 85. 95% (表 1).

1 M S I T N T P T T S N D A C L S I V H S L M C H R Q G G E S E T F A K R A I E S L
41 V K K L K E K K D E L D S L I T A I T T N G A H P S K C V T I Q R T L D G R L Q
81 V A G R K G F P H V I Y A R L W R W P D L H K N E L K H V K Y C Q F A F D L K C
121 D S V C V N P Y H Y E R V V S P G I D L S G L T L S G S G P S G L M V K D E Y D
161 Y E S Q Q S L P S T E G H M Q T I Q H P P S R P V A P E P F N A P A M L P P A E
201 G S S S A S T S A F S S I A V G S T T Q P N S V L T G N H S S E S L L Q I A S G
241 T G Q G S Q Q N G F P P G Q P S T Y H H N A T S T W S R N S N F T P T V P H H Q
281 N G H L Q H H P P M A H P A H Y W P V H N E I A F Q P P I S N H P A P E Y W C S
321 I A Y F E M D V Q V G E **T F K** V P S S C P I V T V D G Y V D P S G G G D R F C L G
361 Q L S N V H R T E A I E R A R L H I G K G V Q L E C K G E G D V W V R C L S D H
401 A V F V Q S Y Y L D R E A G R A P G D A V H K I Y P S A Y I K V F D L R Q C H R
441 Q M Q Q Q A A T A Q A A A A A Q A A A V A G N I P G P G S V G G I A P A I S L S
481 A A A G I G V D D L R R L C I L R M S F V K G W G P D Y P R Q S I K E T P C W I
521 E I H L H R E L H L L D D V L H T M L I A D P T P L D *

阴影表示 N-糖基化位点; 加粗斜体表示蛋白激酶 C 磷酸化位点; 双下划线表示酪蛋白激酶 II 磷酸化位点; 方框表示 N-豆蔻酰化位点; 粗下划线表示酰胺化位点; “*”表示终止密码子.

图 2 草鱼 Smad4 基因全长 cDNA 序列推导的氨基酸序列

Fig. 2 Smad4 gene full-length cDNA sequence-deduced amino acid sequence from grass carp

表 1 草鱼 Smad4 蛋白氨基酸序列相似性比较
Tab. 1 Homology comparison of Smad4 amino acid sequence from grass carp

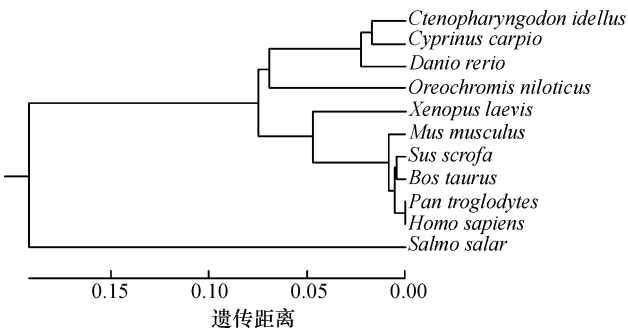
物种	与草鱼相似性/%			
	全长	MH1	连接区域	MH2
斑马鱼	95. 61	100. 0	89. 33	97. 83
鲤	94. 33	100	91. 57	93. 04
非洲爪蟾	84. 63	99. 2	58. 15	97. 83
小鼠	85. 77	99. 2	62. 30	97. 39
人	85. 95	99. 2	62. 30	97. 83

用 MEGA 软件构建了 Smad4 蛋白氨基酸序列的系统进化树(图 3). 分析结果显示, 草鱼和鲤鱼 Smad4 蛋白的亲缘关系比较近.

2.3 Smad4 蛋白疏水性分析

ProtScale 软件的疏水性预测结果如图 4 所示, Smad4 蛋白存在 6 个高分值 (Scare > 1. 5) 峰, 分布在第 16 ~ 18、473 ~ 474、477、480 ~ 481、483、497 氨基酸位点, 属于高疏水性域; 19 个低分值 (Scare < - 1. 5) 峰, 分布在第 26 ~ 28、45 ~ 50、100 ~ 101、103 ~ 104、106 ~ 107、158 ~ 164、179、243 ~ 245、247、250、255 ~

258、265、267、270、277 ~ 285、314、414 ~ 415、436 ~ 444、510 ~ 513 氨基酸位点, 属于高亲水性域.



Ctenopharyngodon idellus: 草鱼 (ADR80616); *Cyprinus carpio*: 鲤 (BAB71792); *Danio rerio*: 斑马鱼 (ACA58502); *Oreochromis niloticus*: 尼罗罗非鱼 (XP_003453910); *Xenopus laevis*: 非洲爪蟾 (NM_001097067. 1); *Mus musculus*: 小鼠 (AAM74472); *Sus scrofa*: 猪 (NM_214072. 1); *Bos taurus*: 牛 (NM_001076209. 1); *Pan troglodytes*: 黑猩猩 (XM_001155480. 2); *Homo sapiens*: 人 (BAB40977); *Salmo salar*: 大西洋鲑 (NP_001133775).

图 3 草鱼 Smad4 蛋白氨基酸序列系统进化树分析

Fig. 3 Phylogenetic analysis of amino acid sequence of Smad4 protein from grass carp

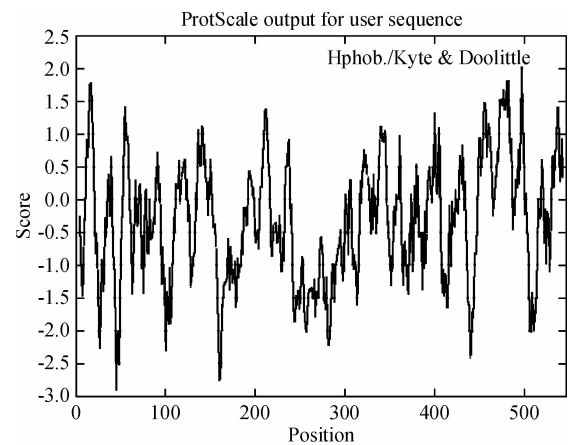


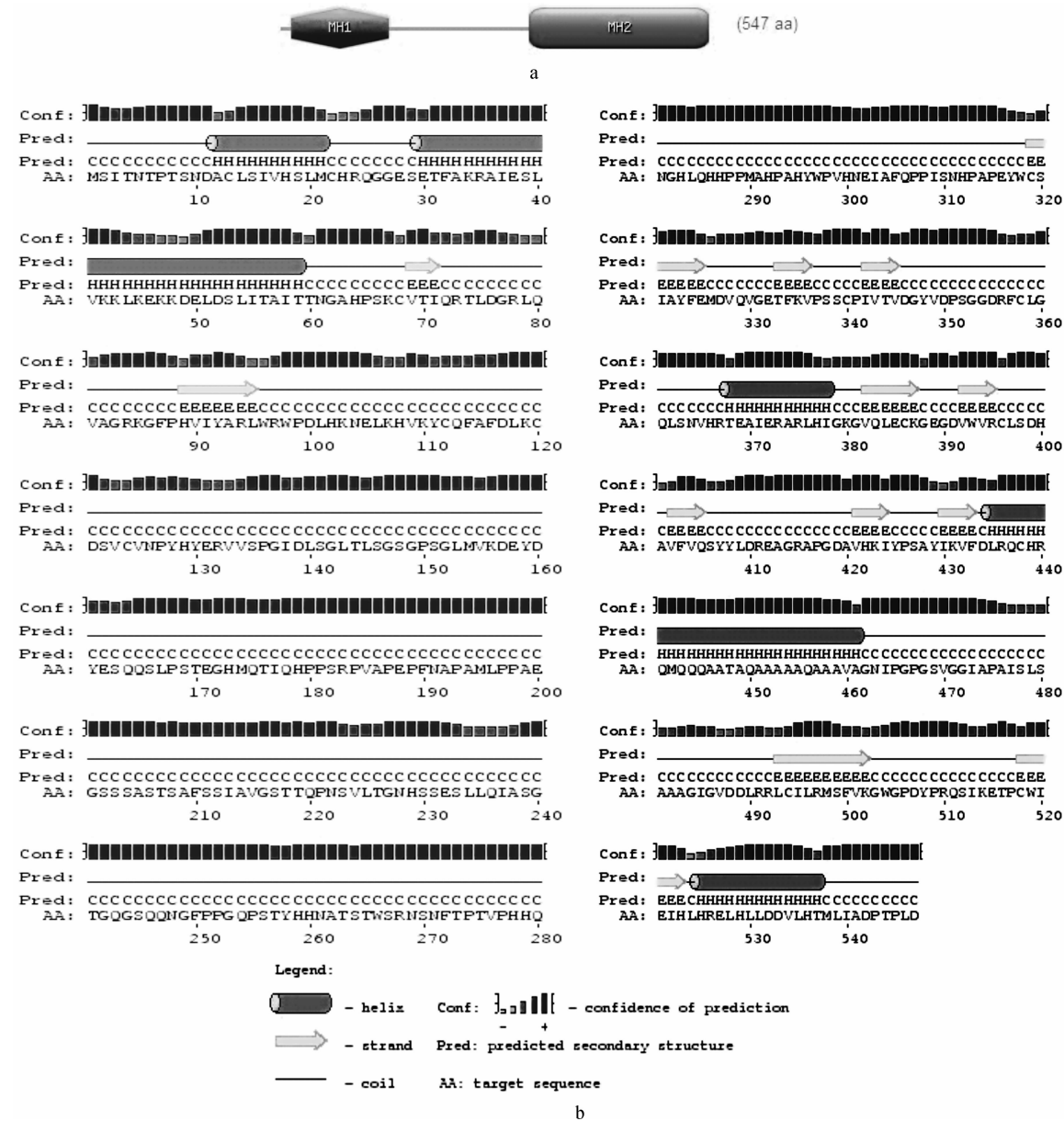
图 4 Smad4 蛋白氨基酸序列的疏水性分析

Fig. 4 Hydrophobicity analysis of amino acid sequence of Smad4 protein

2.4 Smad4 蛋白功能结构域和二级结构特征预测

用 ExPasy 服务器中的 Prosite 工具对 Smad4 蛋白进行预测,发现 2 个结构域(图 5a),第 1 个为 MH1 结构域,由第 15 ~ 139 共 125 个氨基酸残基组成;第 2 个是 MH2 结构域,由第 318 ~ 547 共 230 个氨基酸残基组成。

Predict protein 软件预测结果如图 5b 所示,Smad4 蛋白二级结构中有 5 个 α 螺旋结构域,分布在第 12 ~ 21、30 ~ 60、368 ~ 378、435 ~ 461、525 ~ 537 氨基酸区域;有 12 个 β - 折叠结构域,分布在第 69 ~ 71、89 ~ 95、319 ~ 325、333 ~ 336、342 ~ 345、382 ~ 387、392 ~ 395、402 ~ 405、421 ~ 424、430 ~ 433、493 ~ 502、518 ~ 523 氨基酸区域,其余区域均为无规则卷曲域。



a:Smad4 蛋白功能结构域;b:Smad4 蛋白序列二级结构

图 5 草鱼 Smad4 蛋白序列二级结构和功能结构域预测分析

Fig. 5 Prediction of domains and secondary structure of Smad4 protein from grass carp

2.5 重组质粒构建与 PCR 检测和酶切鉴定

以 pcDNA3.1 (+)-Anti-Smad4 为模板的 PCR 扩增,可见 1 条长度约 1.6 kb 的目的条带.将测序结果与已知的草鱼 Smad4 基因 ORF 序列比对,序列一致,Smad4 片段已经成功地反向插入了 pcDNA3.1 (+) 载体.如图 6 所示,pcDNA3.1 (+) 共由 5 428 bp 组成,EcoR I 和 Xho I 分别位于 952 和 985 bp 处,经双酶切后除去 33 bp,再加上新插入的片段 1 655 bp,重组载体的长度为 7 050 bp.重组表达载体 pcDNA3.1 (+)-Anti-Smad4 经 Xho I 和 EcoR I 双酶切,可见在 1.6 kb 处有 1 条目的条带.

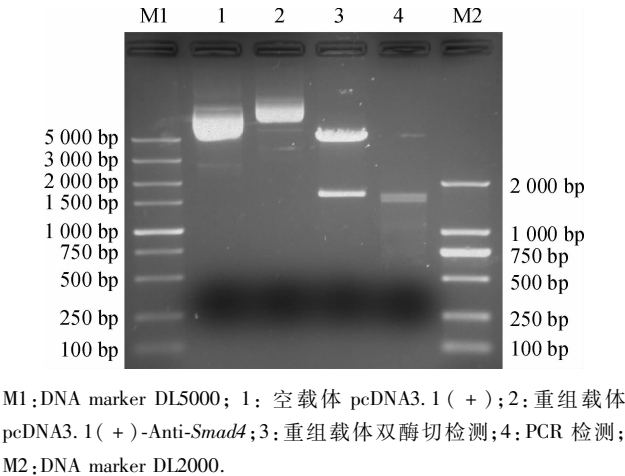


图 6 草鱼 Smad4 反义真核表达载体的 PCR 检测和双酶切鉴定
Fig. 6 PCR and double enzyme identification of recombinant vector pcDNA3.1 (+)-Anti-Smad4 from grass carp

3 讨论与结论

从系统进化树分析结果来看,Smad4 基因推导的氨基酸序列系统进化树主要分成 2 大支:鱼类和哺乳类.系统进化树显示,草鱼与鲤、斑马鱼等鲤科鱼类具有较近的亲缘关系,其中与鲤的亲缘关系最近.进化树中各个物种 Smad4 蛋白的亲缘远近与各自的分类地位相符合,说明 Smad4 蛋白的氨基酸序列具有很高的保守性.相似性比较结果显示,草鱼 Smad4 基因符合 Smad4 基因高度保守的特征,与其他鱼类相比,氨基酸序列的相似性都在 90% 以上;与人和小鼠等其他物种相比,氨基酸序列的相似性则达到了 85% 左右.草鱼 Smad4 基因推导的氨基酸序列系统进化树分析结果和相似性分析都与传统的鱼类分类地位相吻合.

Smads 蛋白非常保守,它们结构相似,均由保守的 N 端 MH1 结构域^[12]、C 端 MH2 结构域^[13]和富含脯氨酸的连接区域构成. MH1 结构域是 Smads 蛋白的 1 个 DNA 绑定区域,通过 β -发夹结构与 DNA 结合, MH2 结构域位于 Smads 蛋白的 C 端,对与 I 型受体结合、磷酸化触发、异源低聚物形成及转录激活等都有作用,对

于 Smad4 与其他 Smad 分子形成低聚物及调节转录有十分重要的作用.结构域 MH1 和 MH2 可相互作用, MH2 结构域抑制 MH1 结构域的 DNA 结合功能, MH1 结构域抑制 MH2 结构域的转录激活功能^[14]. 本试验克隆的草鱼 Smad4 基因经软件分析发现,由它推导的氨基酸序列 MH1 结构域位于第 15 ~ 139 个氨基酸残基处, MH2 结构域位于 318 ~ 547 个氨基酸残基处,中间为连接区域,是典型的 Smads 蛋白结构.

已有研究发现 Smad4 蛋白的 2 个功能结构域 (MH1、MH2) 比较保守,而富含脯氨酸的连接区域是可以进行不同剪切的,不同的选择性剪切可以造成 Smads 蛋白的连接区域长短不同.在非洲爪蟾体内发现了 2 种 Smad4 的剪切体,尽管二者在连接区域存在差异,但是两端的功能结构域都非常保守^[15],而且 2 个剪切体都具有相似的转录活性^[16],在鲤中也发现了 4 种不同的 Smad4 剪切体^[12],它们的 MH1 结构域均含有 141 个氨基酸残基, MH2 结构域氨基酸数量也相差不多.草鱼与其他几个物种的 Smad4 氨基酸序列相似性比较显示, MH1 结构域的相似性极高,与斑马鱼和鲤达到了 100%, MH2 结构域的相似性也很高,除了鲤以外,相似性均在 97% 以上,而中间连接区域的相似性则相对较低,除了与斑马鱼和鲤相比在 90% 以外,与其他几个物种相比均在 60% 左右.据此推测,Smad4 蛋白结构和功能的保守性可能与 MH1 和 MH2 这 2 个功能结构域的高度保守有关.

Smad4 是 TGF- β 信号传导的中心环节,它在分子伴侣 TRAP1 (TNFR associated protein1, 肿瘤坏死因子受体相关蛋白 1)^[17] 的作用下与受体激活型 Smads 结合形成异源复合物,将信号传导至细胞核内,研究发现,受体激活型 Smad 与 Smad4 形成的异源复合物是由 2 个受体激活型 Smads 和 Smad4 组成的异源三聚体^[18],异源三聚体进入细胞核后通过 3 种方式作用于靶基因.第 1 种是异源三聚体直接与 DNA 结合, Smad4 通过识别靶基因上的 5'-AGAC-3' 序列,并通过 N-端 MH1 结构域利用 β -发夹结构与之结合,从而调节靶基因的转录^[19];第 2 种是 Smads 蛋白通过与其他 DNA 结合因子相互作用而与 DNA 结合^[20];第 3 种是 Smads 蛋白复合物与非 DNA 结合因子相互作用^[16].根据软件预测结果,草鱼 Smad4 蛋白的 MH1 结构域也存在 2 个 β -折叠,具备形成 β -发夹结构的条件,暗示草鱼 Smad4 蛋白可能与其他物种的 Smad4 蛋白一样具有 DNA 结合功能^[19].

1984 年, Lzant 等^[21] 就发现反义 RNA 具有下调目的基因表达的功能.目前,反义基因技术已在鱼类研究中得到了应用, Uzbekova 等^[22] 研究发现带重组反义 sGnRH (促性腺激素释放激素) 载体的转基因大西洋鲑鱼 Salmo salar 的 sGnRH 表达量远远低于非转基因鱼.有关鱼类的 Smad4 反义基因技术的研究鲜见报

道,本试验构建了草鱼反义 *Smad4* 基因真核表达载体,以探究草鱼 *Smad4* 基因在其胚胎发育等过程中的作用.在构建反义真核表达载体的过程中,本文选择真核表达载体 pcDNA3.1(+)^[23],载体上人巨噬细胞介导的启动子(CMV)保证重组蛋白高效地表达^[24],丰富的多克隆位点有助于目的片段的有效插入和扩增,牛生长激素 poly A 尾巴(BGH pA)能高效地终止转录^[25],载体上的 pUC ori 则保证了载体在大肠埃希菌 *Escherichia coli* 体内的增殖和合成.因此,本试验构建的草鱼反义 *Smad4* 基因真核表达载体 pcDNA3.1(+)-Anti-*Smad4* 可为进一步研究草鱼 *Smad4* 基因的功能及在胚胎发育过程的作用机理奠定基础.

参考文献:

- [1] GRONEBERG D A, WITT H, ADCOCK I M, et al. Smads as intracellular mediators of airway inflammation [J]. *Exp Lung Res*, 2004, 30(3):223-250.
- [2] CHEN Xin, WEISBERG E, FRIDMACHER V, et al. Smad4 and FAST-1 in the assembly of activin-nesponsive factor[J]. *Nature*, 1997,389(6646):85-89.
- [3] CHEN X, RUBOCK M J, WHITMAN M. A transcriptional partner for MAD protein in TGF-beta signaling[J]. *Nature*, 1996, 383(6602):691-696.
- [4] WANG Zhiyong, FUTAMI K, NISHIHARA A, et al. Four types of Smad4 found in the common carp, *Cyprinus carpio*[J]. *J Exp Zool: Part B: Mol Dev Evol*,2005, 304(3):250-258.
- [5] DICK A, MAYR T, BAUER H, et al. Cloning and characterization of zebrafish *smad2*, *smad3* and *smad4*[J]. *Gene*, 2000, 246(1/2):69-80.
- [6] LAU M T, GE Wer. Cloning of *Smad2*, *Smad3*, *Smad4*, and *Smad7* from the goldfish pituitary and evidence for their involvement in activin regulation of goldfish FSH promoter activity [J]. *Gen Comp Endocr*, 2005, 141(1):22-38.
- [7] CROOKE S T. Progress in antisense technology[J]. *Annu Rrv Med*, 2004(55):61-95.
- [8] OELLER P W, LU M W, TANLOR P L, et al. Reversible inhibition of tomato fruit senescence by antisense RNA [J]. *Science*, 1991,254(5030):437-439.
- [9] HU Wei, LI Shuangfei, TANG Bin, et al. Antisense for gonadotropin-releasing hormone reduces gonadotropin synthesis and gonadal development in transgenic common carp (*Cyprinus carpio*) [J]. *Aquaculture*, 2007, 271(1/2/3/4):498-506.
- [10] LU Yuanan, SUN P S. Viral resistance in shrimp that express an antisense Taura syndrome virus coat protein gene [J]. *Antiviral Res*, 2005, 67(3):141-146.
- [11] XU X B, HE Z P, LIANG Z Q, et al. Obstruction of TGF-beta1 signal transduction by anti-Smad4 gene can therapy experimental liver fibrosis in the rat[J]. *Chin J hepatol*, 2004, 12(5):263-266.
- [12] GRISHIN N V. MH1 domain of Smad is a degraded homing endonuclease[J]. *J Mol Biol*, 2001, 307(1):31-37.
- [13] SHI Yigong. Structural insights on Smad function in TGF-beta function[J]. *Bioessays*, 2001, 23(3):223-232.
- [14] SHI Yigong, MASSAGUE J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus[J]. *Cell*, 2003, 113(6):685-700.
- [15] HOWELL M, ITOH F, PIERREUX C E, et al. Xenopus Smad4beta is the co-smad component of developmentally regulated transcription factor complexes responsible for induction of early mesodermal genes[J]. *Dev Biol*, 1999, 214(2):354-369.
- [16] MASUYAMA N, HANAFUSA H, KUSAKABE M, et al. Identification of two Smad4 proteins in *Xenopus*: Their common and distinct properties[J]. *J Biol Chem*, 1999, 274(17):12163-12170.
- [17] WURTHNER J U, FRANK D B, FELICI A, et al. Transforming growth factor-beta receptor-associated protein 1 is a Smad4 chaperone [J]. *J Biol Chem*, 2001,276(22):19495-19502.
- [18] CHACKO B M, QIN B Y, TIWARI A, et al. Structural basis of heteromeric smad protein assembly in TGF-beta signaling[J]. *Mol Cell*, 2004, 15(5):813-823.
- [19] DAI Jiale, BANSAL R K, KERN S E. G₁ cell cycle arrest and apoptosis induction by nuclear Smad4/Dpc4:Phenotypes reversed by a tumorigenic mutation [J]. *Cell Biol*, 1999,96(4):1427-1432.
- [20] DERYNCK R,ZHANG Y E. Smad-depended and Smad-independed pathways in TGF-beta family signaling[J]. *Nature*, 2003, 425(6958):577-584.
- [21] IZANT J G, WEINTRAUB H. Inhibition of thymidine kinase gene expression by anti-sense RNA: A molecular approach to genetic analysis[J]. *Cell*, 1984, 36(4):1007-1015.
- [22] UZBEKOVA S, CHYB J, FERRIERE F, et al. Transgenic rainbow trout expressed sGnRH-antisense RNA under the control of sGnRH promoter of Atlantic salmon[J]. *J Mol Endocrinol*, 2000,25(3):337-350.
- [23] CHEN Rongyi, CHEN Hongxiang, LIN Jiayi, et al. In-vivo transfection of pcDNA3.1-IGFBP7 inhibits melanoma growth in mice through apoptosis induction and VEGF down expression [J]. *J Exp Clin Canc Res*, 2010, 29(1):13-20.
- [24] MICHAEL B, FRANK W, GERHARD J, et al. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus[J]. *Cell*, 1985, 41(2):521-530.
- [25] GOODWIN E C, ROTTMAN F M, The 3'-flanking sequence of the bovine growth hormone gene contains novel elements required for efficient and accurate polyadenylation[J]. *J Biol Chem*, 1992, 267(23):16330-16334.