

准噶尔雅罗鱼 肌动蛋白基因启动子 克隆及序列分析

胡文革 陈创夫* 王远志 郝凤霞 曹旭东 盛金良

(石河子大学生命科学学院 石河子 832003; 石河子大学动物科技学院 石河子 832003;
石河子大学医学院 石河子 832003)

摘要:利用 PCR 方法克隆了准噶尔雅罗鱼 (*Leuciscus merzbacheri*) 的 α -actin 基因启动子片段 SZ21, 大小是 2 398 bp。对克隆的启动子序列进行了转录调控元件的生物信息学预测分析, 同时, 基于启动子中包含的开放阅读框和内含子序列, 探讨了准噶尔雅罗鱼与鲤鱼 (*Cyprinus carpio*)、草鱼 (*Ctenopharyngodon idella*)、青鱼 (*Mylopharyngodon piceus*)、团头鲂 (*Megalobrama amblycephala*)、泥鳅 (*Misgurnus mizolepis*) 间的系统进化关系。结果显示, 该启动子序列的 3 个核心启动子转录元件: CAA T-box, CA rG motif 和 TATA-box 分别在转录起始位点 (+1) 上游的 - 89、- 59、- 26 处, 序列中还含有 MEF2, SA TB, CHR F, NRE, MTEN, E-box, RU49, ZBP F, CREB, Enhance region, CEBP 位点等多种转录调控元件。在剪接体内含子中, 剪接位点遵循 GT...AG 法则。启动子 SZ21 序列含有 3 个内含子和 155 个氨基酸。内含子 1、内含子 2、内含子 3 的系统发育分析表明, 团头鲂与草鱼和青鱼的亲缘关系要比与准噶尔雅罗鱼的更近一些, 这与传统分类中的亲缘关系显示不一致, 其原因尚需探讨。

关键词: 准噶尔雅罗鱼; α -actin 基因启动子; 转录调控元件; 内含子; 系统发育

中图分类号: Q751 **文献标识码:** A **文章编号:** 0250-3263 (2010) 01-18-09

Cloning and Sequence Analysis of the α -actin Gene Promoter Region of *Leuciscus merzbacheri*

HU Wen-Ge CHEN Chuang-Fu* WANG Yuan-Zhi HAO Feng-Xia

CAO Xu-Dong SHENG Jin-Liang

(Life Science College, Shihezi University, Shihezi 832003;

Animal Science and Technology College, Shihezi University, Shihezi 832003;

Medical College, Shihezi University, Shihezi 832003, China)

Abstract Through PCR amplification, the α -actin gene promoter fragment SZ21 (2 398 bp) of *Leuciscus merzbacheri* was obtained. We forecasted its transcriptional regulation elements using bioinformatics method. In view of its ORF and introns sequence, the system development was analyzed between *L. merzbacheri* and other closely related fish species. Three core promoter transcriptional elements were situated in the transcriptional start site (+1) upper - 89, - 59, - 26, including the CAA T-box, CA rG motif, TATA-box. Results of putative transcription binding sites analysis of the promoters revealed the presence of several critical transcription binding

基金项目 国家自然科学基金项目 (No. 30860215), 石河子大学科技发展计划项目 (No. ZR KX200744);

*通讯作者, E-mail: ccf-xb@163.com;

第一作者介绍 胡文革, 女, 副教授; 研究方向: 动物分子遗传与育种; E-mail: hwg-t@163.com.

收稿日期: 2009-06-05, **修回日期:** 2009-11-04

sites such as the MEF2, SATB, CHRF, NRE, MTEN, E-box, RU49, ZBPF, CREB, Enhance region and CEBP in the further promoter. The splice sites abided by GT...AG rule in the spliceosome introns. The promoter SZ21 sequence included the intron 1, intron 2, intron 3, and 155 amino acids. The system development analysis of three introns showed that *Megalobrama amblycephala* had more close relationship to *Ctenopharyngodon idella* and *Mylopharyngodon piceus* than *L. merzbacheri*. This is not consistent with the traditional classification, and the reason needs further investigation.

Key words: *Leuciscus merzbacheri*; α -actin gene promoter; Transcriptional regulation element; Intron; System development

启动子是真核生物转基因制作过程中表达载体构件的重要组成部分,启动子决定了基因表达的时空以及强度^[1]。启动子包含核心启动子区域和调控区域,核心启动子区域产生基础水平的转录,调控区域能够对不同的环境条件作出应答。对于启动子的深入研究,有助于优化控制外源基因在转基因生物中的表达。

肌动蛋白(actin)是所有真核生物细胞内的重要结构蛋白。 α -actin是最主要的非肌肉或胞质肌动蛋白异型体,它在真核生物的大多数非肌肉细胞以及未分化的成肌细胞中表达^[2-3]。 α -actin启动子具有广泛启动表达的功能,具有与SV40早期启动子相当甚至更强的启动转录活性^[4]。第一个被分离到的鱼类 α -actin基因启动子是鲤鱼(*Cyprinus carpio*)的,已证明其具有有效的启动作用^[5-6]。之后,草鱼(*Ctenopharyngodon idella*)、团头鲂(*Megalobrama amblycephala*)、青鳉(*Oryzias latipes*)、斑马鱼(*Danio rerio*)、罗非鱼(*Tilapia*)、泥鳅(*Misgurnus mizolepis*)、青鱼(*Mylopharyngodon piceus*)、唐鱼(*Tanichthys albonubes*)等的 α -actin启动子也先后被分离并验证了其驱动基因表达的活性^[7-9]。

准噶尔雅罗鱼(*Leuciscus merzbacheri*)又名新疆雅罗鱼,隶属鲤形目鲤科雅罗鱼属,是新疆特有土著鱼类,是新疆现有3种雅罗鱼属鱼即准噶尔雅罗鱼、贝加尔雅罗鱼(*L. leuciscus baicalensis*)和高体雅罗鱼(*L. idus*)中较古老的鱼类物种^[10-11]。该鱼已被列为新疆濒危鱼类,其基因资源优势目前尚没有被开发利用。本实验克隆了准噶尔雅罗鱼的 α -actin基因启动子,并对克隆的启动子序列进行了转录调控元件的

生物信息学预测分析,同时,基于启动子中包含的内含子序列,与近缘鱼类物种间进行了同源性比对和系统发育分析。为下一步进行转基因鱼研究提供基本材料。

1 材料与方法

1.1 实验材料 准噶尔雅罗鱼是采自新疆赛里木湖(东经 80°39'~81°30',北纬 44°27'~44°45')的土著野生鱼类。大肠杆菌 *E. coli* DH5⁺、pBS-T载体、long DNA Taq和普通琼脂糖凝胶 DNA回收试剂盒、DNA Marker购自天根生化科技有限公司。T₄DNA连接酶购自TaKaRa公司。其他常用试剂为本室自行配制。PCR仪为BD RAD公司生产的MyCycleTM热循环仪,凝胶成像仪为基因有限公司(Gene Company Limited)生产的G:BOX SYNGENE。引物合成和DNA序列测序由上海生工生物工程技术服务有限公司完成。

1.2 实验方法

1.2.1 准噶尔雅罗鱼基因组DNA提取 称取约0.05g准噶尔雅罗鱼肌肉组织块,放入1.5ml离心管,用手术剪刀将组织块尽量剪碎,加600 μ l DNA提取缓冲液[0.4 mol/L NaCl, 10 mmol/L Tris-Cl (pH 8.0), 2 mmol/L EDTA (pH 8.0), 40 μ g/ml RNase A]。用自制的离心管小锤把组织块研磨成糊状。分装400 μ l/管。每管各加10% SDS 40 μ l, 10 mg/ml蛋白酶K 17 μ l,小心混匀。60 $^{\circ}$ C水浴2.5 h后,取出加5 mol/L NaCl 400 μ l,混匀。离心后取上清液,加1/10体积的NaAc及等体积的异丙醇-20 $^{\circ}$ C沉淀1 h。70%乙醇洗涤沉淀,加入适量TE溶解。-20 $^{\circ}$ C保存备用。

1.2.2 引物设计 根据已发表的鲤鱼^[12]和草鱼的 *-actin* 基因序列 (NCBI序列号分别是 M24113和 M25013)相对保守区,参照 Hwang等^[13]使用的扩增鲤鱼 *-actin* 基因启动子的引物,设计合成 1对简并引物,上游引物为 A-P1: 5'-GTGT/AGTGACGCT/CGGACCAATC-3',下游引物为 A-P3: 5'-CCA GAGTCCA TCACGAT-ACCACT-3'。

1.2.3 准噶尔雅罗鱼 *-actin* 基因启动子的克隆 用 long DNA *Taq* 酶 PCR 扩增准噶尔雅罗鱼 *-actin* 基因启动子。扩增程序为 94 预变性 4 min; 94 变性 50 s, 63 退火 50 s, 72 延伸 1 min 50 s, 30 个循环; 72 延伸 10 min。1%的琼脂糖凝胶检测 PCR 产物,UV 凝胶仪拍照。切胶回收 PCR 产物,将 PCR 产物连接到克隆载体 pBS-T上。转化 *E. coli* DH5 感受态细胞,进行蓝白斑筛选,选取白斑阳性克隆,经菌落 PCR 和双酶切鉴定正确的,提取质粒测序。

1.2.4 启动子序列分析 利用软件 DNAMAN 5.2.2和 NCBI在线 blastn,通过与 Liu等^[12]发表的和已登陆的鲤鱼 M24113、草鱼 M25013、青鱼 AY289135、团头鲂 AY170122、泥鳅 AF2706-49的同源序列进行比对,再利用软件 GenScan (<http://genes.mit.edu/GENSCAN.html>),结合序列比对预测克隆的准噶尔雅罗鱼 *-actin* 基因启动子中的内含子和外显子的拼接位点。利用在线工具 MatInspector software (<http://www.genomatix.de/cgi-bin/matinspector.pl>)以及 Match program (<http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi>),预测启动子区域结构特征和转录调控元件,根据转录调控元件结合位点核心区的相关度,选取相关度系数(在线工具自动计算得出)大于 0.9 的结合位点,以确保其具有较高的特异性。分离准噶尔雅罗鱼 *-actin* 基因启动子以及鲤鱼、草鱼、青鱼、团头鲂、泥鳅启动子部分的内含子和外显子序列,利用软件 DNAMAN 5.2.2分析准噶尔雅罗鱼 *-actin* 基因启动子开放阅读框的蛋白质编码序列,然后,对于 *-actin* 基因在

不同鱼类物种中的蛋白质编码序列以及内含子序列分别做同源性比对和多序列分析,构建 observed divergency 系统树,通过自引导检验 (bootstrap) 获得系统分支的置信度 (1 000 次重抽样)。

2 结果与分析

2.1 准噶尔雅罗鱼 *-actin* 基因近端启动子及其部分开放阅读框的获得 对提取的准噶尔雅罗鱼基因组 DNA 利用引物 A-P1/A-P3 经 PCR 扩增,电泳检测,得到一条准噶尔雅罗鱼 *-actin* 基因启动子片段 (记为 SZ21),大小约为 2 400 bp。切胶回收纯化后,克隆到 pBS-T 载体上,记为 pBS-SZ21-P。蓝白斑筛选后,选取阳性克隆提取质粒,经 PCR 和双酶切鉴定正确的 (图 1),进行测序。测序后经拼接分析,克隆的准噶尔雅罗鱼 *-actin* 基因启动子 SZ21 序列大小为 2 398 bp。

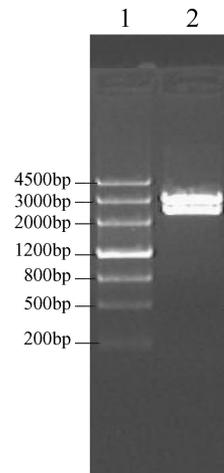


图 1 用 *Eco*R 和 *Hind* 双酶切质粒 pBS-SZ21-P 电泳结果

Fig 1 The electrophoresis picture of digested plasmid pBS-SZ21-P by *Eco*R + *Hind*

1. DNA 分子量标准 ; 2. pBS-SZ21-P 双酶切产物。

1. DNA Marker ; 2. The digested plasmid pBS-SZ21-P.

2.2 准噶尔雅罗鱼 *-actin* 基因启动子序列特征分析 序列分析表明,克隆的 SZ21 启动子片段含 104 bp 的 5 侧翼序列近端启动子和 2 294 bp 的 *-actin* 基因部分开放阅读框序列。5 侧

翼序列含有与大多数鱼类相同的与转录密切相关的 3 个核心启动子转录元件: CAA T-box、CArG motif和 TA TA -box,其位置分别在转录起始位点 (+1)上游的 - 89、- 59、- 26 处。我们利用软件分析结合序列比对,预测了克隆的准噶尔雅罗鱼 *-actin* 基因启动子中的内含子和外显子的拼接位点。在剪接体内含子中,剪接位点都是 GT...AG(典型剪接位点)。SZ21 的开放阅读框包含不表达的第 1 外显子、表达的第 2 和第 3 外显子以及第 4 外显子的部分序列。在 SZ21 启动子序列中,表达的外显子、内含子 1、内含子 2、内含子 3 的 G + C 含量分别为 56%、39.3%、29.4%和 40.9%,这与一些文

献中认为的非编码区中的 A + T 含量高的结果一致^[14]。序列分析结果见图 2。

在线软件预测表明,全长 2 398 bp 的启动子 SZ21 序列含有 1 个 CAA T-box, 2 个 CC(A/T)₆ GG-box, 10 个 TA TA -box, 1 个 MEF2, 2 个 SA TB, 1 个 CHR F, 1 个 NRE, 1 个 MTEN, 2 个 E-box, 1 个 RU49, 2 个 ZBPF, 3 个 CREB, 1 个 Enhance region, 7 个 CEBP 位点等多种转录调控元件(除 1 个 MEF2 和 1 个 MTEN 外,相关度系数 >0.9)(表 1)。

2.3 克隆的启动子开放阅读框蛋白质编码序列与其他鱼类同源性分析 准噶尔雅罗鱼 *-actin* 基因 SZ21 启动子包含 2 294 bp 的 *-actin*

表 1 准噶尔雅罗鱼 *-actin* 基因启动子 SZ21 序列可能存在的重要转录调控元件预测

Table 1 Prediction of important transcriptional regulation elements for

Leuciscus merzbacheri -actin gene promoter SZ21

调控元件 The regulation elements	序列位点 (5 ~ 3, bp) (链) The sequence sites(5 - 3, bp) (strain)	Core 分值 Core score	Matrix 分值 Matrix score	特性功能 Character and function
CAAT box	- 89 ~ - 85	1.000	1.000	核心启动子的 CAA T 框
CC/(A/T) ₆ GG box	- 59 ~ 50 1 176 ~ 1 194 ^a	1.000 1.000	1.000 0.913	核心启动子的 CArG 基序
TA TA box	- 26 ~ - 22 105 ~ 121 ^a 116 ~ 132 ^a 186 ~ 202 450 ~ 466 ^a 453 ~ 467 ^a 834 ~ 850 ^a 893 ~ 909 1 176 ~ 1 192 ^a 1 372 ~ 1 388 ^a	1.000 1.000	1.000 >0.9	核心启动子的 TA TA 框
MEF2 ^b	273 ~ 295	1.000	0.773	肌细胞特异性增强子结合因子
SA TB	244 ~ 258 ^a 247 ~ 261	1.000	>0.9	富含 AT 的特殊序列结合蛋白
CHR F	300 ~ 312	1.000	>0.9	细胞周期调节子
NRE	429 ~ 439	1.000	>0.9	核心启动子起始因子元件
MTEN ^b	513 ~ 533	1.000	0.801	核心启动子基序十元件
E-box	1 306 ~ 1 318 1 467 ~ 1 481	>0.9	>0.9	E-box 结合因子
RU49	1 721 ~ 1 727	1.000	1.000	锌指转录因子 RU49
ZBPF	73 ~ 85 2 206 ~ 2 228	1.000	>0.9	锌结合蛋白因子
CREB	275 ~ 295 443 ~ 463 967 ~ 987	1.000	>0.9	cAMP 效应元件结合蛋白
Enhance region	1 175 ~ 1 225	1.000	0.941	增强子区域
CEBP	771 ~ 785 1 008 ~ 1 093 ~ 1 107 1 094 ~ 1 108 ^a 1 395 ~ 1 409 1 564 ~ 1 578 ^a 2 140 ~ 2 154	>0.9	>0.9	Ccaat/增强子结合蛋白

a 表示反义序列; b 由于 MEF2 和 MTEN 的作用重要,尽管两者的 matrix 分值分别只有 0.773 和 0.801,还是被列了出来。

a Indicates anti-sequence; b Because of the MEF2 and MTEN 's important function, they are listed, although their matrix scores are only 0.773 and 0.801, respectively.

基因部分开放阅读框序列,其蛋白质编码序列含有 155 个氨基酸。利用软件 DNAMAN 4.0,把鲤鱼、草鱼、青鱼、团头鲂、泥鳅的 *-actin* 基因

氨基酸序列与准噶尔雅罗鱼的 SZ21 进行同源性比对分析,结果见图 3。

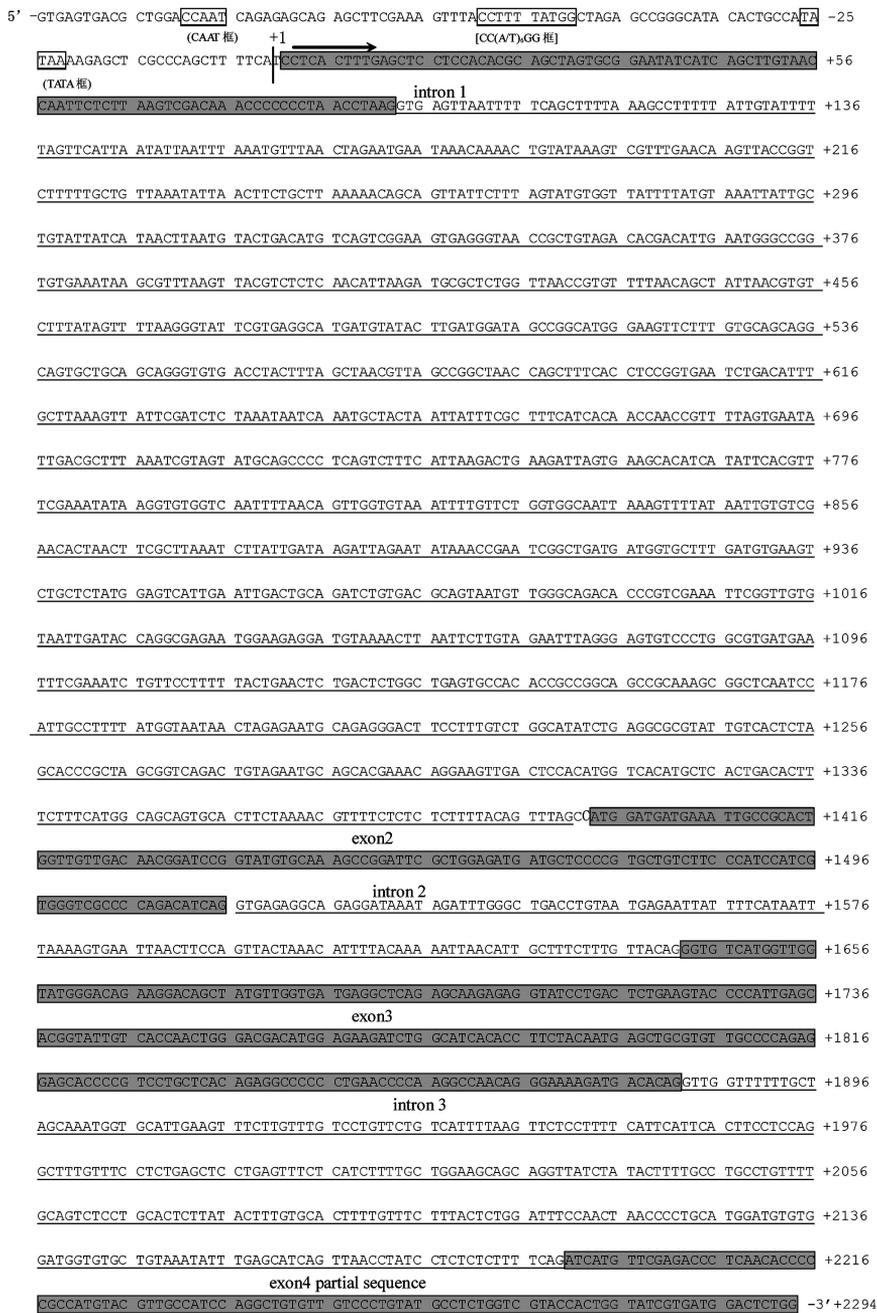


图 2 准噶尔雅罗鱼 β -actin 基因启动子 SZ21 的 3 个核心启动子转录元件及内含子、外显子序列结构示意图

Fig 2 The sequence composition sketch map of three core promoter transcriptional elements, introns, and exons of the β -actin gene promoter SZ21 of *Leuciscus merzbacheri*

+1 表示转录起始位点; “□”表示转录起点上游近端启动子的 CAAT框、CC(A/T)6GG框、TATA框;

“—”表示内含子部分; “//”表示外显子部分。

+1 indicates transcription start site; “□” indicates CAAT-box, and CC(A/T)6GG motif, and TATA-box of proximal promoter in transcription start site upper; “—” indicates intron region; “//” indicates exon region

准噶尔雅罗鱼	MDDI I AALVVDNNGSGMCKAGFAGDDAERAVF P SIVG R ERH	40
草鱼	MDDI I AALVVDNNGSGMCKAGFAGDDAERAVF P SIVG R ERH	40
鲤鱼	MDDI I AALVVDNNGSGMCKAGFAGDDAERAVF P SIVG R ERH	40
泥鳅	MDDI I AALVVDNNGSGMCKAGFAGDDAERAVF P SIVG R ERH	40
青鱼	MDDI I AALVVDNNGSGMCKAGFAGDDAERAVF P SIVG R ERH	40
团头鲂	MDDI I AALVVDNNGSGMCKAGFAGDDAERAVF P SIVG R ERH	40
Consensus	mdd i aalvvdngsgmckagfagddapravf p sivg r prh	
准噶尔雅罗鱼	QGV M V G M G C K D S Y V G D E A C S K R G I L T L K Y E I E H G I V T N W D	80
草鱼	QGV M V G M G C K D S Y V G D E A C S K R G I L T L K Y E I E H G I V T N W D	80
鲤鱼	QGV M V G M G C K D S Y V G D E A C S K R G I L T L K Y E I E H G I V T N W D	80
泥鳅	QGV M V G M G C K D S Y V G D E A C S K R G I L T L K Y E I E H G I V T N W D	80
青鱼	QGV M V G M G C K D S Y V G D E A C S K R G I L T L K Y E I E H G I V T N W D	80
团头鲂	QGV M V G M G C K D S Y V G D E A C S K R G I L T L K Y E I E H G I V T N W D	80
Consensus	qgv m v g m g c k d s y v g d e a c s k r g i l t l k y e i e h g i v t n w d	
准噶尔雅罗鱼	D M E K I W H H T F Y N E L R V A P E E H P V L L T E A P L N P K A N R E K M T	120
草鱼	D M E K I W H H T F Y N E L R V A P E E H P V L L T E A P L N P K A N R E K M T	120
鲤鱼	D M E K I W H H T F Y N E L R V A P E E H P V L L T E A P L N P K A N R E K M T	120
泥鳅	D M E K I W H H T F Y N E L R V A P E E H P V L L T E A P L N P K A N R E K M T	120
青鱼	D M E K I W H H T F Y N E L R V A P E E H P V L L T E A P L N P K A N R E K M T	120
团头鲂	D M E K I W H H T F Y N E L R V A P E E H P V L L T E A P L N P K A N R E K M T	120
Consensus	d m e k i w h h t f y n e l r v a p e e h p v l l t e a p l n p k a n r e k m t	
准噶尔雅罗鱼	Q I M F E T I N T E A M Y V A I Q A V L S L Y A S G R T T G I V M D	154
草鱼	Q I M F E T I N T E A M Y V A I Q A V L S L Y A S G R T T G I V M D	154
鲤鱼	Q I M F E T I N T E A M Y V A I Q A V L S L Y A S G R T T G I V M D	154
泥鳅	Q I M F E T I N T E A M Y V A I Q A V L S L Y A S G R T T G I V M D	154
青鱼	Q I M F E T I N T E A M Y V A I Q A V L S L Y A S G R T T G I V M D	154
团头鲂	Q I M F E T I N T E A M Y V A I Q A V L S L Y A S G R T T G I V M D	154
Consensus	q i m f e t n t p a m y v a i q a v l s l y a s g r t t g i v m d	

图 3 准噶尔雅罗鱼 α -actin 基因部分氨基酸序列的同源性分析结果图

Fig 3 The homology analysis chart of partial amino acid sequence for *Leuciscus merzbacheri* α -actin gene

准噶尔雅罗鱼: *Leuciscus merzbacheri*; 草鱼: *Ctenopharyngodon idella*, M25013; 鲤鱼: *Cyprinus carpio*, M24113;

泥鳅: *Misgurnus mizolepis*, AF270649; 青鱼: *Mylopharyngodon piceus*, AY289135;

团头鲂: *Megalobrama amblycephala*, AY170122.

图中 Consensus 行的小写字母为 6 种鱼一致的氨基酸。

The small letters of consensus rows are the identical amino acids of six fish

2.4 准噶尔雅罗鱼 α -actin 基因启动子部分的内含子系统发育分析 准噶尔雅罗鱼 α -actin 基因启动子 SZ21 含有第 1、第 2 和第 3 个完全的内含子。利用软件 DNAMAN 5.2.2 做多序

列比对,将克隆的准噶尔雅罗鱼 α -actin 基因启动子 SZ21 部分的第 1、第 2、第 3 三个内含子与鲤鱼、草鱼、青鱼、团头鲂、泥鳅的进行同源性比对(表 2),同时构建了系统发育树(图 4)。

表 2 准噶尔雅罗鱼 α -actin 基因 SZ21 启动子中的氨基酸、内含子与其他鱼类的同源性比较分析

Table 2 The homology comparison analysis of amino acid and intron of α -actin gene between *Leuciscus merzbacheri* SZ21 and others fishes

相比较鱼的种类 The compared species of fish		同源性 Similarity (%)			
		氨基酸 Amino acid	内含子 1 Intron 1	内含子 2 Intron 2	内含子 3 Intron 3
草鱼 <i>Ctenopharyngodon idella</i>	M25013	99.4	81.0	79.2	91.0
青鱼 <i>Mylopharyngodon piceus</i>	AY289135	99.4	80.8	77.7	93.2
团头鲂 <i>Megalobrama amblycephala</i>	AY170122	99.4	78.5	75.0	92.2
鲤鱼 <i>Cyprinus carpio</i>	M24113	99.4	71.7	67.0	84.8
泥鳅 <i>Misgurnus mizolepis</i>	AF270649	98.1	34.4	45.2	44.3

表 2 显示,准噶尔雅罗鱼与相比较鱼类物种的同源性由近至远依次为:氨基酸序列分析结果是草鱼、青鱼、团头鲂、鲤鱼等同;内含子 1 序列分析结果是草鱼、青鱼、团头鲂、鲤鱼、泥鳅;内含子 2 序列分析结果是草鱼、青鱼、团头鲂、鲤鱼、泥鳅;内含子 3 分析结果是青鱼、团头

鲂、草鱼、鲤鱼、泥鳅。

由图 4 系统发育分析图的拓扑结构和分枝长度可以看出,就 α -actin 基因内含子 1、2、3 序列而言,团头鲂与草鱼和青鱼的亲缘关系要比准噶尔雅罗鱼与草鱼和青鱼的关系更近一些。

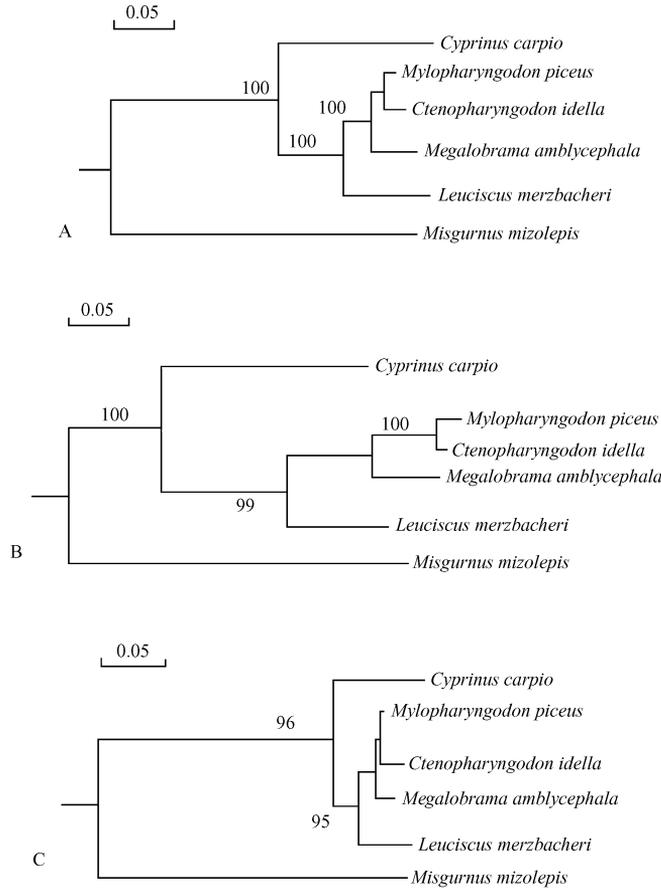


图 4 基于 α -actin 基因内含子序列构建的 observed divergency 系统发育树

Fig 4 The phylogenetic tree of the observed divergency based on α -actin gene intron

A. 内含子 1; B. 内含子 2; C. 内含子 3. A. Intron 1; B. Intron 2; C. Intron 3.

Mylopharyngodon piceus: 青鱼, AY289135; *Ctenopharyngodon idella*: 草鱼, M25013;

Cyprinus carpio: 鲤鱼, M24113; *Leuciscus merzbacheri*: 准噶尔雅罗鱼, 本实验克隆的序列;

Megalobrama amblycephala: 团头鲂, AY170122; *Misgurnus mizolepis*: 泥鳅, AF270649.

3 讨论

肌动蛋白基因是一个持家基因,在所有的真核生物细胞中高度保守。鱼类的 α -actin 基因调控序列被广泛应用于鱼类的转基因研究。与该基因转录下调相关的元件位于 3 非翻译

区,不带此 3 非翻译区下调相关元件的 α -actin 基因启动子,在真核生物的大多数非肌肉细胞以及未分化的成肌细胞中组成性表达^[9, 15]。 α -actin 基因启动子包含典型的 5 侧翼区调控元件 CAA T 框、TATA 框及 CA rG 基序(即 CC(A/T)₆GG 序列)。 α -actin 启动子的活性除了与上

述 5 侧翼区调控元件密切相关外,也与 5 非翻译区有关。已证实 *-actin* 基因的第 1 个内含子中含有增强子^[15]。研究显示,在人的 *-actin* 基因中,含有 5 侧翼序列以及第 1 个内含子的 *-actin* 基因启动子具有与 SV40 早期启动子相当的或更强的启动转录活性^[4]。参照 Liu 等^[12]对鲤鱼 *-actin* 基因启动子分析的结果并比对鲤鱼 *-actin* 基因序列 (NCBI 登录号 M24113),结合在线预测分析,本实验克隆的准噶尔雅罗鱼 *-actin* 基因启动子 SZ21,具有近端启动元件 CCAAT 框、CC(A/T)₆GG 框、TATA 框和包含在 *-actin* 基因第 1 个内含子中的增强子序列,不包含 3 非翻译区下调序列;近端启动元件 CC(A/T)₆GG 序列除了在转录起点上游 -59 ~ -50 处含有一个以外,在第 1 个内含子的反义序列中也有一个 (1 176 ~ 1 194), TATA box 除了在转录起点上游 -26 ~ -22 处含有一个以外,在第 1 个内含子的正义和反义序列中还含有多个位点,这可能也预示着准噶尔雅罗鱼 *-actin* 基因的第 1 个内含子作为 *-actin* 基因启动子的一部分在起始转录表达中的重要性。除了位于转录起始位点上游 100 bp 左右的 CAAT、TATA 框等基本调控元件外, *-actin* 基因可能还有更多的正调控或负调控元件^[9],对于有些特定基因,其转录区内部也存在着转录因子的结合位点,能够对不同的环境条件作出应答。相关实验已证实启动子 SZ21 具有启动绿色荧光蛋白基因在真核细胞中转录表达的活性(另文发表)。克隆的该启动子虽然经软件分析,除了在 5 侧翼区含有 CCAAT 框、TATA 框、CC(A/T)₆GG 序列外,还在第 1 个内含子中含有多个 TATA box,一个 CC(A/T)₆GG 序列、核心启动子起始因子元件、核心启动子基序十元件、Ccaat/增强子结合蛋白等多种重要调控元件,但是在对真核细胞基因表达的实际启动转录中,这些调控元件对转录启动活性的影响如何,还需要深入的实验加以验证。

内含子序列包含相对较快的进化速率,通常用于分析亲缘关系相对较近的物种间的系统发育,内含子的系统发育分析成为基于线粒体

的系统发育分析的一个很好的补充^[16]。王宁等^[17]在灵长目动物中分离出 7 个断裂基因的内含子和外显子序列,分别进行系统发育分析; Prychitko 等^[18]研究发现,内含子对于揭示啄木鸟近缘物种间的进化关系有很大的作用; Lee 等^[19]用 *-actin* 基因内含子 2 构建分子系统发育进化树,分析了不同鱼类物种的系统发育关系,并且在家族水平上评估了来自进化的基因流。我们用克隆到的准噶尔雅罗鱼 *-actin* 基因的第 1、第 2、第 3 内含子与草鱼、青鱼、团头鲂、鲤鱼和泥鳅的相应序列进行系统发育分析,内含子 1、内含子 2 和内含子 3 的系统发育分析都显示出与传统分类不一致的结果,即团头鲂与草鱼和青鱼的亲缘关系要比准噶尔雅罗鱼与草鱼和青鱼的亲缘关系更近一些,这与准噶尔雅罗鱼与草鱼、青鱼同属雅罗鱼亚科,而团头鲂属于鲃亚科的传统分类不一致。当然该结果还需要用线粒体 DNA 的序列做互补验证。不过,这个结果与我们分析的编码蛋白质的氨基酸序列也是一致的,在编码蛋白质的氨基酸序列中,雅罗鱼亚科的草鱼、青鱼与鲤亚科的鲤鱼及鲃亚科的团头鲂的 155 个氨基酸完全一致,而同属雅罗鱼亚科的准噶尔雅罗鱼第 128 个氨基酸由苯丙氨酸变成了亮氨酸,也显现出与传统分类的不一致。

参 考 文 献

- [1] Pabo C O, Sauer R T. Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem*, 1992, 61: 1053 - 1095.
- [2] Pederson T. Half a century of "the nuclear matrix". *Mol Biol Cell*, 2000, 11: 799 - 805.
- [3] Pederson T, Aebi U. Actin in the nucleus: what form and what for. *Struct Biol*, 2002, 140 (1/2/3): 3 - 9.
- [4] Robert M F, Monette R M, Aikichi I, et al. 5' flanking and first intron sequences of the human *-actin* gene required for efficient promoter activity. *Nucleic Acids Research*, 1989, 17 (1): 254 - 270.
- [5] Liu Z J, Moav B, Faras A J, et al. Functional analysis of elements affecting expression of the *-actin* gene of carp. *Mol Cell Biol*, 1990, 10: 3432 - 3440.
- [6] Williams D W, Muller F, Lavender F L, et al. High

transgene activity in the yolk syncytial layer affects quantitative transient expression assays in zebrafish (*Danio rerio*) embryos Transgenic Res, 1996, 5: 433 - 442

[7] Feng H, Cheng J, Luo J, et al Cloning of black cap *-actin* gene and primarily detecting of its promoter region the function Acta Genetica Sinica, 2006, 33 (2): 133 - 140.

[8] 郁二蒙,叶星,王海英,等. 罗非鱼 肌动蛋白基因启动子的分离及其启动子活性的初步检测. 农业生物技术学报, 2008, 16(2): 242 - 247.

[9] 王海英,叶星,劳海华,等. 唐鱼 *-actin* 基因近端和远端启动子的鉴定及其启动活性分析. 生物工程学报, 2008, 24(10): 1768 - 1775.

[10] 胡文革,段子渊,王金富,等. 新疆 3种雅罗鱼线粒体 DNA 控制区序列的差异和系统进化关系. 遗传学报, 2004, 31(9): 970 - 975.

[11] 胡文革,段子渊,王金富,等. 新疆 3种雅罗鱼线粒体 DNA 细胞色素 *b* 序列的差异与系统进化. 动物学杂志, 2005, 40(3): 6 - 11.

[12] Liu Z J, Zhu Z, Roberg K, et al The *-actin* gene of carp (*Ctenopharyngodon idella*). Nucleic Acids Res, 1989, 17 (4): 5850.

[13] Hwang G L, Rahman M A, Razak S A, et al Isolation and characterisation of tilapia *-actin* promoter and comparison of its activity with carp *-actin* promoter Biochimica et Biophysica Acta, 2003, 1625: 11 - 18.

[14] Li W H, Graur D. Fundamentals of Molecular Evolution. Massachusetts: AC Publishers, 1991.

[15] Liliana D Z, Anne S T, Bruce M P. A 40-base-pair sequence in the 3' end of the *-actin* gene regulates *-actin* mRNA transcription during myogenesis PNAS, 1998, 85: 1389 - 1393.

[16] Manuel I, Scott W R. Spliceosomal introns as tools for genomic and evolutionary analysis Nucleic Acids Research, 2008, 36(5): 1703 - 1712.

[17] 王宁,陈润生. 基于内含子和外显子的系统发育分析的比较. 科学通报, 1999, 44(19): 2095 - 2102.

[18] Prychitko T M, Moore W S. Comparative evolution of the mitochondrial cytochrome *b* gene and nuclear beta-Fibrinogen gene intron 7 in woodpeckers Mol Biol Evol, 2000, 17 (7): 1101 - 1111.

[19] Lee J S, Gye M C. Use of beta-actin gene intron 2 as a phylogenetic marker in fish taxonomy DNA Seq, 2001, 12 (1): 71 - 76.

《动物学杂志》第十届编辑委员会

主 编:马 勇

副主编:宋延龄 赵 勇 彭景榎 徐延恭 顾亦农(常务)

编 委:(以姓氏笔画为序)

- 马 勇 马建章 王祖望 王跃招 王德华 方盛国 计 翔 孙青原 孙悦华
 刘遵发 许木启 李 宁 李 明 李进华 李枢强 李新正 张正旺 张春光
 张树义 张瑾峰 吴孝兵 陈佩惠 **宋大祥** 宋延龄 宋林生 杨 光 杨增明
 孟安明 宛新荣 郑光美 赵 勇 费 梁 钟文勤 桂建芳 夏国良 顾亦农
 徐存拴 徐宏发 徐延恭 曹 焯 彭贤锦 彭景榎 蒋志刚 魏辅文

责任编辑:顾亦农 梁 冰