

# 异源四倍体鲫鲤 *Sox9a* 基因保守区序列的克隆 及内含子剪接位点分析<sup>\*</sup>

刘季芳 刘少军<sup>\*\*</sup> 张 纯 孙远东 颜金鹏 刘 筠

湖南师范大学生命科学学院, 教育部蛋白质化学和鱼类发育生物学重点实验室, 长沙 410081

**摘要** 用特异于 HMG-box 保守区的兼并引物扩增了异源四倍体鲫鲤及其原始亲本(红鲫和鲤鱼)的 *Sox* 基因. 异源四倍体鲫鲤的扩增产物共有 4 条带, 大小分别约为 200, 600, 900 和 1900 bp, 而其原始亲本的扩增产物只有 3 条带, 且在雌雄个体中未发现性别特异扩增带. 回收雄性四倍体鱼 600 bp 扩增带, 经亚克隆及测序分析得到一个新的编码 71 个氨基酸残基的基因. 该基因与虹鳉鱼 (*Oncorhynchus mykiss*)、蟾蜍 (*Xenopus laevis*)、鲤鱼 (*Cyprinus carpio*) 等的 *Sox9* 基因相似性达 97%, 据此命名为 *Atsox9a*. 通过计算机分析和 RT-PCR 方法确定该基因含有内含子, 并获得其内含子的序列和剪切位点, 此内含子剪切位点在进化上是保守的. *Atsox9a* 对于研究异源四倍体鲫鲤性别决定和分化机制及脊椎动物性别进化具有重要的理论意义.

**关键词** 异源四倍体鲫鲤 *Sox9a* 基因 内含子

1990 年 Sinclair 等<sup>[1]</sup> 分离出的人的睾丸决定基因(*SRY*)是性别决定及分化机制研究领域的一个重大突破, 使人们对性别决定及分化机制的认识有了较大发展. *SRY* 基因具有高度的进化保守性, 在许多物种中已发现 *SRY* 的同源基因, 如果蝇、两栖类、鸟类和其他哺乳动物<sup>[2,3]</sup>. 这些基因都具有 *SRY* 基因的 HMG 盒 (high mobility group box) 基序, 现已把同 *SRY* 保守区 HMG 基序具有 60% 以上相似性的基因统称为 *SRY* 基因即 *Sox* 基因 (*SRY*-box gene)<sup>[4]</sup>. 目前已在不同的物种包括哺乳类、爬行类、两栖类、鱼类以及昆虫类中克隆了 30 多个这样的基因, 而这些基因编码的氨基酸序列从人类到果蝇都是高度保守的<sup>[2,5]</sup>. *Sox* 蛋白是一个具有 HMG 盒的转录因子家族, 由 8 个亚组 (A-H) 构成<sup>[6,7]</sup>. 研究表明 *Sox* 基因在揭示脊椎动物性别决定的分子机制和胚胎发育中的基因调控具有十分

重要的意义, 如 *Sox5*, *Sox9*, *Sox17* 等与性别决定和精子发生有关; *Sox2*, *Sox3*, *Sox11* 与神经系统发育有关<sup>[8,9]</sup>. 在这些基因中 *Sox5*, *Sox9*, *Sox17*, *Sox20*, *Sox8* 等具有不同数目的内含子, 而 *Sox9* 和 *Sox17* 有一个内含子位于 HMG 盒内. *Sox* 基因盒区内含子的剪接可能具有重要的作用, 如在哺乳类, *Sox17* 基因内含子的不同剪接方式对精子的正常形成必不可少<sup>[10]</sup>, 而 *Sox9* 基因内含子也有两种切除方式, 产生的两种蛋白可能会诱发两种不同的发育分化机制<sup>[11]</sup>.

湖南师范大学生命科学学院与其协作单位合作, 研制出两性可育、能自然繁殖的异源四倍体鲫鲤群体<sup>[12]</sup>, 其原始母本为红鲫 (*Carassius auratus red variety*), 原始父本为湘江野鲤 (*Cyprinus carpio L. Xiangjiang*). 经染色体、DNA 含量等方法检测证实鲫鲤 F<sub>1</sub> ~ F<sub>2</sub> 代为二倍体 (2n=100), F<sub>3</sub> ~ F<sub>11</sub> 代

2003-09-15 收稿, 2003-11-10 收修改稿

<sup>\*</sup> 国家自然科学基金重点(批准号: 30330480, 30170733)、国家基础研究发展规划(批准号: 2001CB109006)和教育部跨世纪人才培养基金(批准号: 200248)资助项目

<sup>\*\*</sup> 通讯作者, E-mail: lsj@hunnu.edu.cn

为四倍体 ( $4n=200$ )<sup>[12,13]</sup>. 异源四倍体鲫鲤雌、雄两性可育并能自然繁殖, 形成了一个遗传性状稳定的四倍体鱼新种群. 利用四倍体鱼与二倍体鲫、鲤鱼杂交, 获得了具有生长快、肉质鲜美、抗病力强等优良性状的三倍体湘云鲫、湘云鲤, 并在全国 28 个省市推广养殖, 取得了显著的经济和社会效益. 以前我们对四倍体鲫鲤和三倍体湘云鲫的染色体、DNA 含量、性腺发育、配子的特征、胚胎发育、mtDNA 等生物学特征进行了较系统的研究<sup>[12~19]</sup>, 但是至今还没有关于它们与性别相关基因的报道. 本文采用 PCR 亚克隆技术及 RT-PCR, 对异源四倍体鲫鲤 *Sox9a* 基因进行研究, 旨在为异源四倍体鲫鲤性别决定的分子机制及 *Sox9* 基因的进化等研究提供分子生物学证据.

## 1 材料与方法

### 1.1 材料

异源四倍体鲫鲤及原始亲本红鲫和鲤鱼均取自“湖南师范大学”国家四倍体鱼种质资源保护基地”, 性别按活体解剖确定, 均为性成熟个体. 兼引物参照 *Sox* 基因保守区序列<sup>[20]</sup>, 由 Sangon 公司合成. 上游引物: 5' TGAAGCGACCCATGAA (C/T) G3', 下游引物: 5' AGGTCG (A/G) TACTT (A/G) TA (A/G) TT3'.

### 1.2 基因组 DNA 的提取及 PCR 扩增 *Sox* 基因保守区

取四倍体鱼及原始亲本的血液提取基因组 DNA, 实验方法参照文献[21]进行. PCR 扩增反应总体积为 25  $\mu$ L, 其中基因组 DNA 80  $\mu$ g, 每种 dNTP 200  $\mu$ mol  $\cdot$  L<sup>-1</sup>, 每种引物各 0.3  $\mu$ mol  $\cdot$  L<sup>-1</sup>, TaqDNA 聚合酶 0.9U, 反应液在 94  $^{\circ}$ C 变性 1 min, 52  $^{\circ}$ C 退火 1 min, 72  $^{\circ}$ C 延伸 1 min, 共 35 个循环, 最后在 72  $^{\circ}$ C 延伸 7 min. PCR 产物在 1.5% 琼脂糖凝胶电泳中分离后, 选取雄性四倍体鱼 600 bp 左右扩增片段, 用 Sangon 公司胶回收试剂盒纯化.

### 1.3 PCR 产物的克隆及序列分析

将回收的 DNA 片段克隆于载体 pUCm-T 中, 转化到大肠杆菌 DH5a, 经培养后, 采用 PCR 方法进一步筛选阳性菌株; 对阳性菌株扩大培养后提取

质粒并用 *Pst*I 酶切鉴定. 对筛选到的不同阳性克隆进行测序并推测编码的氨基酸序列, 在 NCBI 数据库查找斑马鱼、黄鳝、鼠、鲤等相应的氨基酸序列, 用 Jellyfish (2.1) 软件进行同源性分析, 根据其查询序列的最高相似性而命名所得克隆.

### 1.4 总 RNA 的提取及 RT-PCR

取雄性四倍体鱼精巢提取总 RNA. RT-PCR 反应总体积为 25  $\mu$ L, 含 25 pmol  $\cdot$  L<sup>-1</sup> 每引物, 1.5 mmol  $\cdot$  L<sup>-1</sup> MgCl<sub>2</sub>, AMV 反转录酶和 Taq 聚合酶各 1 U. 反转录条件为: 42  $^{\circ}$ C 热循环 60 min 进行 cDNA 第一链的合成, 接下来的 PCR 反应程序为: 94  $^{\circ}$ C 预变性 4 min, 然后进行 30 个热循环, 每一个热循环包括 94  $^{\circ}$ C 变性 1 min, 55  $^{\circ}$ C 退火 40 s, 72  $^{\circ}$ C 延伸 1 min, 最后在 72  $^{\circ}$ C 延伸 10 min. 取 PCR 产物用 1.5% 的琼脂糖凝胶电泳进行检测. 对 PCR 产物进行克隆及测序, 方法同 1.3 节所述.

## 2 结果

### 2.1 四倍体鱼及其原始亲本 *Sox* 基因的扩增

PCR 扩增产物琼脂糖凝胶电泳结果如图 1. 在雌雄四倍体鱼中有 4 条明显的扩增带, 分子大小分别约为 200, 600, 900 和 1900 bp, 而在其原始母本红鲫和原始父本鲤鱼中均只有 3 条扩增带, 大小分别约为 200, 600 和 1900 及 200, 500 和 900 bp. 这些结果表明在四倍体鱼及其原始亲本中存在 *SRY* 的同源基因, 但在雌雄四倍体鱼中均未发现性别特异扩增带. 选取雄性四倍体鱼 600 bp 左右的扩增带回收后亚克隆.

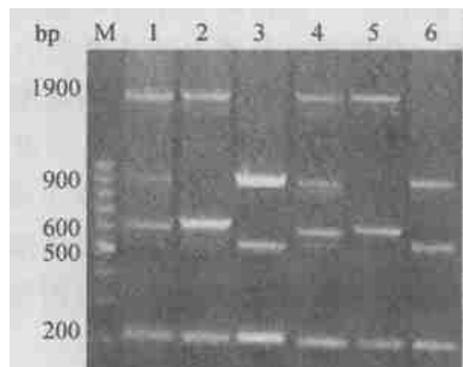


图 1 异源四倍体鱼及其原始亲本 *Sox* 基因的扩增图谱  
M, 100bp DNA 分子标记; 1, 雌性四倍体鱼; 2, 雌性红鲫  
3, 雌性鲤鱼; 4, 雄性四倍体鱼; 5, 雄性红鲫; 6, 雄性鲤鱼

```

A   AAG CGA CCC ATG AAT GCG TTT ATG GTG TGG GCT CAA GCC GCG CGC
B   K   R   P   M   N   A   F   M   V   W   A   Q   A   A   R
A   AGG AAA CTG GCG GAT CAG TAT CCA CAC CTG CAC AAC GCC GAG CTC
B   R   K   L   A   D   Q   Y   P   H   L   H   N   A   E   L
A   AGC AAG ACC CTC GGG AAA CTC TGG AG gt gag... (内含子413bp) ... tttagg
B   S   K   T   L   G   K   L   W   R
A   TTA CTG AAT GAG GGC GAG AAG CCG TCC GTT CGT GGA GGA GGC CGA
B   L   L   N   E   G   E   K   P   S   V   R   G   G   G   R
A   GCG TCT GAG GGT CCA GCA CAA GAA AGA CCA CCC CGA CTA CAA GTA
B   A   S   E   G   P   A   Q   E   R   P   P   R   L   Q   V
A   CCG ACC
B   P   T
    
```

图 2 计算机推测的 *Atsox9a* 内含子剪接位点

A, *Atsox9a* 基因的 DNA 序列, 其中下划线表示内含子的剪接位点; B, 根据 *Atsox9a* 基因序列推导的氨基酸序列

	1	11	21	31	41	51	61
A	TGAAGCGACCCATGAATGCGTTTATGGTGGGGCTCAAGCCGCGCCAGGAACTGGCCGATCAGTATCC						
B	TGAAGCGACCCATGAATGCGTTTATGGTGGGGCTCAAGCCGCGCCAGGAACTGGCCGATCAGTATCC						
Consensus	tgaagcgaccatgaa gc tttatggt tgg c a cg g g ag aaa t gga cagt tcc						
	71	81	91	101	111	121	131
A	ACACCTGCACAAACCCGAGCTCAGCAAGACCCCTCGGAAACTCTGGAGGTGAGGGCTTGCACTGTTTCATC						
B	ACACATGCACAAACCCGAGATCTCCAAAAGACTCGGGGAAGGCTTGGAG-----						
Consensus	acac tgcacaacgc gag tc caa a ct gggaa c tggag						
	141	151	161	171	181	191	201
A	GAGATTAGCACAGCTGCAGAAGACCATTTCATTCAACTGCTCAGACAACTTACTGTGTTATTATAATTTA						
B	-----						
Consensus	-----						
	211	221	231	241	251	261	271
A	ATCCGATTGTTTCATAGCATTATTATATAATGATAAGGAAATGCTGATTACCTTGAGAAGAGGTAACCTGC						
B	-----						
Consensus	-----						
	281	291	301	311	321	331	341
A	TTGGTGAATTAACCCCTTTGTTCCATGCACATCACTGCAGAAACAGTTAAAAACATTCTCAATTTGCTGC						
B	-----						
Consensus	-----						
	351	361	371	381	391	401	411
A	TTAATGATACTTAGTAGAGTAGTTTTTAAGTTTTGGTGGAGGTTGGATTATGGGTTGACTATACTITTT						
B	-----						
Consensus	-----						
	421	431	441	451	461	471	481
A	TTCCCTGAATAACATTTTTTAAGTGAACCCAAAAAGCAATAATCATAGTTGTTTCAAGAACAAITTTGT						
B	-----						
Consensus	-----						
	491	501	511	521	531	541	551
A	ATCTTTTTTTGTAGGTTCTTTTATTGTGAAGTGTTTTTTAGGTTACTGAATGAGGGCGAGAAGCCGTCGG						
B	-----GTTACTGAAAGCAAGTGAAGAGATTCC						
Consensus	gttac a g aa tcc						
	561	571	581	591	601	611	621
A	TTCGTGGAGGAGCCGAGCGTCTGAGGTCACAGCAAGAAAGACCCCGACTACAGTACCAGCCTA						
B	TTTATCCGCGAAGCCGAGCGGCTCCGGCTCCAAGACATGGCTGAETADCCCTGACTACAAGTACCAGCCTA						
Consensus	tt t ga gcggagcg ct gg tcca caca g gac accc gactacaagtaccgaccta						
	631	641	651	661	671	681	691
A	-----						
B	-----						
Consensus	-----						

图 3 异源四倍体鱼 *Sox9a* 基因组序列与 cDNA 序列的比较

21994-2(3669)基因组序列dbi:Sox9a cDNA 序列; 阴影部分表示基因组序列和 cDNA 序列相同的部分; 虚线表示内含子部分

## 2.2 *Atsox9a* 基因保守区的克隆及测序

为了确定这些扩增产物的性质, 对 600 bp 左右扩增带进行了克隆及序列分析. 将测序结果输入 NCBI 数据库与青鳉鱼 (*Oryzias Latipes*) *Sox9*、黄鳝 (*Monopterus albus*) *Sox9*、鲤鱼 (*Cyprinus carpio*) *Sox9b*、蛙 (*Rana rugosa*) *Sox9*、鼠 (*Mus musculus*) *Sox10* 氨基酸序列进行比较, 同源性分别为 97%, 97%, 97%, 95% 和 93%, 由此命名为 *Atsox9a*. 该基因大小为 630 bp, 属于 HMG-box 基序的一部分, 通过计算机分析推测可能具有一个内含子, 该内含子的位置及大小与同属于 E 亚组的人的 *Sox8* 基因有较大的区别, 但插入位点却完全一致, 符合 GT-AG 规则(图 2).

## 2.3 *Atsox9a* 基因组序列与 cDNA 序列的比较

经 RT-PCR 及 cDNA 序列的测定, *Atsox9a* 的 cDNA 序列大小为 217 bp, 与基因组序列比较发现 *Atsox9a* 确实含有一个内含子, 其大小为 413 bp, 插入位点符合 GT-AG 规则(图 3).

## 3 讨论

鱼类的性别决定机制是多样的和易变的, 大部分鱼无异形性染色体, 仅少数鱼类存在异形性染色体, 并且染色体的数目差异很大, 其中鲤形目的鲫鱼、鲶形目的胡鲶鱼以及鲈形目的刺鲈具有 XX/XY 型性染色体, 鲱形目的短颌鲚具有 ZZ/ZO 性染色体. 大鳞副泥鳅为 ZZ/ZW 型. 我们以前对异源四倍体鱼核型及染色体显带的研究结果没有发现异源四倍体鱼有明显性染色体, 但我们对异源四倍体鱼雌核发育的研究表明其雌性为 XXXX, 属雌性配子同型, 这可能是四倍体鱼性染色体尚处于进化的原始状态, 性染色体与常染色体在形态上无法用传统的细胞学方法区分所致. 而雄性四倍体鱼是否为 XXYY 或 XXXY, 目前我们正在用雄核发育、性反转和杂交的方法做进一步研究. 因此, 关于四倍体鱼性别决定机制的研究是一个非常复杂而有意义的话题. 本文通过兼并 PCR 技术克隆了一个新的 *Sox* 基因即 *Atsox9a*, 它的主要特征是具有 HMG 结构域, 可以和 DNA 进行序列特异性结合, 该区域的突变将会使其失去与 DNA 的结合能力从而导致性反转<sup>[22]</sup>. 所以, *Atsox9a* 基因在四倍体鱼的性别决

定中可能具有重要的作用. 其次, 令人注目的是 *Atsox9a* 在 HMG 盒区有一个内含子, 而在目前已发现的 30 多个 *Sox* 基因中, 仅 *Sox5*, *Sox9*, *Sox17* 和 *Sox8* 等具有数目不等的内含子, 其中 *Sox9* 基因与性别决定直接相关, 且在进化上是十分保守的, 其保守基序内的内含子可能具有重要的作用. Takase 等曾报道蛙 *Sox9* 的内含子有两种不同的剪切方式, 产生两种不同的蛋白, 在性腺发育中起重要的作用. 而且, 目前在与 *Sox9* 同属于 E 亚组的人的 *Sox8* 基因 HMG 盒区也发现一个 1.39 kb 大小的内含子, 该内含子的大小和位置虽然与四倍体鱼 413 bp 的内含子不同, 但其插入位点却是一致的<sup>[23]</sup>. 因此, *Atsox9a* 基因 HMG 盒区内含子剪切位点是非常保守的. 在高等真核生物中, 绝大多数基因都具有内含子, 内含子与外显子重新组合对促进基因的大幅度进化起主要作用<sup>[24]</sup>. 所以, 深入研究 *Atsox9a* HMG-box 基序中内含子的拼接方式是十分有意义的. 此外, 从图 1 中可知, 四倍体鱼兼有其原始父母本的扩增带, 但也有变异, 如四倍体鱼中没有 500 bp 的扩增带, 这也说明四倍体鱼是异源四倍体而不是同源四倍体, 它们是通过杂交而来, 不是雌核发育或雄核发育的后代. 在四倍体鱼 *Sox* 基因的扩增中发现 4 条不同的扩增带, 但未发现性别特异带, 这可能是因为鱼类的性别决定机制较原始, 与性别决定有关的 *Sox* 基因尚未集中到某一特定染色体上. 异源四倍体鱼具有 4 套染色体组, 两套来自鲫鱼, 两套来自鲤鱼, 经过 9 代 ( $F_3 \sim F_{11}$ ) 的四倍体性繁殖已形成了一个数目庞大、遗传性状稳定的群体. 而鱼类在脊椎动物的系统进化中处于承前启后的关键地位, 研究四倍体鱼性别决定及分化机制不仅可以用来指导生产实践, 而且可以帮助人们更好地理解性染色体的进化问题 and 高等脊椎动物的性别决定.

## 参 考 文 献

- 1 Sinclair A H, et al. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, 1990, 346(6281): 240
- 2 Denny P, et al. A conserved family of genes related to the testis-determining gene, *SRY*. *Nucleic Acids Research*, 1992, 20(11): 2887
- 3 Tiersch T R, et al. Studies on the phylogenetic conservation of the

- SRY* gene. *Human Genet*, 1991, 87(5): 571
- 4 Foster J W, et al. An *SRY*-related sequence on the marsupial X chromosome: Implications for the evolution of the mammalian testis-determining gene. *Proc Natl Acad Sci USA*, 1994, 91(5): 1927
  - 5 Paul D. A conserved family of genes related to the testis determining gene, *SRY*. *Nucleic Acids Research*, 1992, 20(11): 2887
  - 6 Wright E M, et al. Seven new members of the *Sox* gene family expressed during mouse development. *Nucleic Acids Research*, 1993, 21: 744
  - 7 Osaki E, et al. Identification of a novel *Sry*-related gene and its germ cell-specific expression. *Nucleic Acids Research*, 1999, 27(12): 2503
  - 8 Goze G, et al. Partial cloning of *Sox11* and *Sox12* two new human *Sox* genes. *Nucleic Acids Research*, 1993, 21(12): 2943
  - 9 Uwauogho D, et al. Embryonic expression of the chicken *Sox2*, *Sox3* and *Sox11* genes suggests an interactive role in neuronal development. *Mech Dev*, 1995, 49(1~2): 23
  - 10 Kanai Y, et al. Identification of two *Sox17* messenger RNA isoforms with and without the high mobility group box region, and their differential expression in mouse spermatogenesis. *J Cell Biol*, 1996, 133(3): 667
  - 11 Takase M, et al. Two *Sox9* messenger RNA isoforms: Isolation of cDNAs and their expression during gonadal development in the frog *Rana rugosa*. *FEBS Lett*, 2000, 466(2~3): 249
  - 12 Liu S J, et al. The formation of tetraploid stocks of red crucian carp  $\times$  common carp hybrids as an effect of interspecific hybridization. *Aquaculture*, 2001, 192(2~4): 171
  - 13 孙远东, 等. 异源四倍体鲫鱼  $F_9$ - $F_{11}$  染色体和性腺观察. *遗传学报*, 2003, 30(5): 414
  - 14 刘少军, 等. 三倍体湘云鲫繁殖季节的性腺结构观察. *水生生物学报*, 2000, 24(4): 301
  - 15 刘少军, 等. 异源四倍体鲫鱼成熟性腺和红细胞超微结构观察. *自然科学进展*, 2003, 13(2): 194
  - 16 刘少军, 等. 三倍体湘云鲫性腺指数分析. *水产学报*, 2002, 26(2): 111
  - 17 李建中, 等. 异源四倍体鲫鱼的受精细胞学. *动物学报*, 2002, 48(2): 233
  - 18 李建中, 等. 异源四倍体鲫鱼的性腺发育研究. *水生生物学报*, 2002, 26(2): 116
  - 19 黎双飞, 等. 异源四倍体鲫鱼肝组织线粒体 DNA 的研究. *水产学报*, 2000, 24(6): 301
  - 20 Bowles J, et al. Phylogeny of the *Sox* family of developmental transcription factors based on sequence and structural indicators. *Dev Biol*, 2000, 227: 239
  - 21 Sambrook J, et al. *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press, 1989
  - 22 Foster J W, et al. Campomelic dysplasia and autosomal sex reversal caused by mutations in an *SRY*-related gene. *Nature*, 1994, 372(6506): 525
  - 23 Dietmar P. The *Sox8* gene is located within 700kb of tip of chromosome 16p and is deleted in a patient with ATR-16 syndrome. *Genomics*, 2000, 63: 108
  - 24 Maeda N, et al. The evolution of multigene families: Human hemoglobin genes. *Annual Review Genetics*, 1986, 20: 81