

大鳞副泥鳅 *Sox* 基因保守区的克隆与特征分析

游翠红^{1,2} 童金苟¹ 俞小牧¹

1. 中国科学院水生生物研究所/淡水生态与生物技术国家重点实验室, 武汉 430072;

2. 汕头大学海洋生物研究所, 汕头 515063

摘要 为探讨大鳞副泥鳅(*Paramisgurnus dabryanus*)的性别决定与分化机制,采用 *Sox* 基因 HMG 盒保守区的 2 对简并引物, *SoxN* 与 *Sox9*, 用以扩增 *Sox* 基因的基因组 DNA, 并对其性腺组织进行 RT-PCR 分析。 *SoxN* 引物扩增基因组 DNA 共获得 *pdSox1*、*pdSox2*、*pdSox3*、*pdSox4*、*pdSox11*、*pdSox14*、*pdSox19*、*pdSox21* 等 13 个 *Sox* 基因, 分属 *SoxB* 与 *SoxC* 类群; *Sox9* 引物扩增基因组 DNA 共获得 *pdSox8*、*pdSox9* 等 5 个 *Sox* 基因, 属于 *SoxE* 类群, 这 5 个 *Sox* 基因均含有一个内含子, 且内含子的起始位置一致。基于 *Sox* 基因 HMG 盒蛋白质的遗传距离构建的 NJ 系统进化树表明, 大鳞副泥鳅 18 个 *Sox* 基因在小鼠或斑马鱼中均能找到相应的 *Sox* 直系同源基因, 其中, *pdSox4*、*pdSox8*、*pdSox9*、*pdSox11*、*pdSox14*、*pdSox19* 等 6 个基因存在基因加倍现象, 具有 2 个或 2 个以上的拷贝。用 *Sox9* 引物对大鳞副泥鳅的性腺进行 RT-PCR 扩增和测序分析, 结果表明, 在精巢中表达的是 *pdSox9a*, 在卵巢中表达的是 *pdSox8a*。以上结果表明, *pdSox8* 与 *pdSox9* 是大鳞副泥鳅的性别相关基因, 它们可能在大鳞副泥鳅的性别分化中发挥作用。

关键词 大鳞副泥鳅; *Sox* 基因; HMG 盒; 性别决定; 性别分化

中图分类号 S 917.4 **文献标识码** A **文章编号** 1000-2421(2012)04-0490-09

1990 年, 在人 Y 染色体上发现 SRY 雄性决定基因, 后来, 在哺乳类及非哺乳类中陆续发现众多 SRY 的同源基因 *Sry*。SRY/*Sry* 拥有一个编码 79 个氨基酸的保守序列, 即 HMG 盒, 后来将那些与 SRY/*Sry* 的 HMG 盒氨基酸序列有 60% 以上相似性的编码基因称为 *Sox* (SRY-related HMG-box genes, *Sox*) 基因^[1]。迄今为止, 在脊椎动物与果蝇等无脊椎动物中克隆到的 *Sox* 基因至少有 40 种^[2]。根据 *Sox* 基因内含子的位置、数目以及内含子存在与否等序列特征将这些 *Sox* 基因归于 10 个类群 A-J, 每个类群约有 1~5 个成员, 它们在脊椎动物的性别决定与分化、早期胚胎发育、神经系统发育、晶状体发育、软骨及多种组织器官的形成中发挥重要作用^[1-3]。

大鳞副泥鳅(*Paramisgurnus dabryanus*) 是一种小型经济鱼类, 被认为是雌性异配型(ZZ/ZW 型) 性别决定^[4], 是少数被报道具有异形性染色体的鱼类之一。目前已经有一些关于鱼类包括大鳞副泥鳅 *Sox* 基因的克隆和表达的报道^[5-8], 但是 *Sox* 基因家

族成员与大鳞副泥鳅性别决定和性别分化机制的关系都还有待探讨。笔者采用 *Sox* 基因的 HMG 盒保守区的 2 对简并引物, 扩增 *Sox* 基因的基因组 DNA, 共克隆获得 18 个大鳞副泥鳅 *Sox* 基因, 同时在性腺中进行 RT-PCR 扩增与产物测序分析, 旨在为深入探讨大鳞副泥鳅的性别决定与发育机制提供资料。

1 材料与方法

1.1 试验材料

试验用的大鳞副泥鳅购自武汉市水产批发市场。将达到性成熟的雌雄个体各取 4 尾, 剪尾鳍保存于 95% 乙醇中备用。尾鳍基因组 DNA 的提取参照萨姆布鲁克和拉舍尔^[9]的酚-氯仿-异戊醇方法稍加修改。同时取鱼的性腺组织, 用液氮速冻, 转移至 -80℃ 冰箱中备用。性腺总 RNA 的提取采用 Trizol 试剂(Invitrogen 公司)进行。

1.2 PCR 扩增与 RT-PCR 反应

扩增 *Sox* 基因基因组 DNA 的 2 对简并引物,

收稿日期: 2011-11-22

基金项目: 国家自然科学基金项目(30370225)和农业部专项课题(200903045, 2011-G12)

游翠红, 博士, 研究方向: 鱼类遗传育种, E-mail: chyong@stu.edu.cn

通讯作者: 童金苟, 博士, 研究员, 研究方向: 鱼类遗传育种, E-mail: jgtong@ihb.ac.cn

SoxN 引物与 Sox9 引物(表 1),来自欧洲黑鲈(*Dicentrarchus labrax*)^[10]。SoxN 引物用来扩增 *Sox* 基因家族的 *SoxB* 与 *SoxC* 类群中的成员,*Sox9* 引物用来特异扩增 *Sox9* 的相关基因。PCR 反应体系为 25 μ L,包括约 100 ng DNA,1 $\times$ PCR buffer,0.2 mmol/L 引物,0.2 mmol/L dNTP,2.5 U *Taq* 酶(BioStar)。PCR 反应程序如下:94 $^{\circ}$ C 变性 1 min,54 $^{\circ}$ C 退火 1 min,72 $^{\circ}$ C 延伸 1 min,35 个循环;72 $^{\circ}$ C 延伸 10 min。

RT-PCR 反应分别以大鳞副泥鳅的精巢和卵巢 RNA 为模板,采用 9 碱基随机引物合成第一链 cDNA。反应体系为:1 μ g 总 RNA,1 $\times$ RT-PCR buffer,1 mmol/L dNTP,5 μ mol/L 随机引物,100 U M-MuLV 反转录酶(MBI 公司),20 U RNA 酶抑制剂(Promega 公司),总体积为 10 μ L。PCR 反转录程序为:30 $^{\circ}$ C 保温 10 min,42 $^{\circ}$ C 保温 1 h,70 $^{\circ}$ C 变性 15 min。获得的第一链 cDNA 经 TE 稀释 5 倍后,取 1 μ L 作为模板用 Sox9 引物进行 PCR 扩增,PCR 反应条件同上。

1.3 PCR 产物的回收与测序

采用 1.5% 的琼脂糖凝胶电泳分离 PCR 产物,切下目标电泳条带,经商品试剂盒(武汉柏傲生物工程有限公司)纯化后,用 T4 DNA 连接酶连接到 pMD18-T 质粒(TaKaRa)上,转化至大肠杆菌 *E. coli* DH5 α 感受态细胞,37 $^{\circ}$ C 培养过夜,挑单克隆菌落经 PCR 扩增鉴定阳性,将阳性单克隆菌落送上海联合基因生物技术公司测序。

1.4 序列分析及系统进化树的构建

将所得序列用 Clustal X (1.83)进行比对,非重复序列在 <http://blast.ncbi.nlm.nih.gov/> 进行 BlastX 同源搜索,序列间的相似性采用 Bioedit 软件计算,*Sox* 基因的鉴定与命名通过 BlastX 搜索 GenBank 和人类基因组中同源序列,并参照 Koopman 等^[11]脊椎动物 SOX 蛋白的标志性氨基酸残基信息进行。依据 *Sox* 基因假定编码的氨基酸序列,采用 MEGA3.1 构建系统进化树,bootstrap=1 000。

表 1 克隆大鳞副泥鳅 *Sox* 基因所用的引物

Table 1 Primers used for cloning the <i>Sox</i> genes of <i>Paramisgurnus dabryanus</i>	
引物名称 Primer name	引物序列(5'→3') Primer sequence
SoxN	ATGAAYGCNTTYATGGTNTGG
	GGNCGRtayttrtTARTCNGG
Sox9	ATGAAYGCSTTYATGGTTTGG
	GTCIGGGTGRtCYTtCTTTrtGYTG

2 结果与分析

2.1 大鳞副泥鳅基因组 DNA 与性腺 cDNA 扩增结果

应用 SoxN 引物扩增大鳞副泥鳅雌雄个体的基因组 DNA,均获得单一电泳条带,大小为 200 bp 左右;采用 Sox9 引物进行扩增,每个个体均获得 5 条电泳条带,且片段长度最大的条带在个体之间呈现多态性,但雌雄之间无明显差异(图 1);用 Sox9 引物扩增大鳞副泥鳅的精巢与卵巢 cDNA,均获得单一电泳条带,大小约为 190 bp。

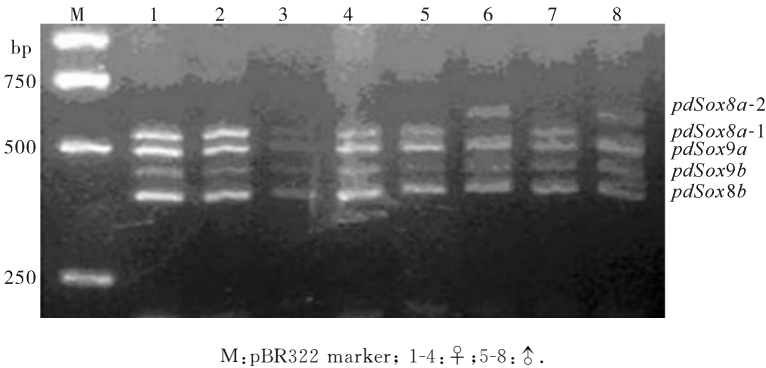


图 1 Sox9 引物扩增 8 尾大鳞副泥鳅的基因组 DNA

Fig.1 PCR amplification of the genomic DNA of eight individuals of *P. dabryanus* with the Sox9 primer

2.2 PCR 产物的测序及序列分析

用 SoxN 引物扩增基因组 DNA 获得单一的电泳条带,取其中 1 尾个体的 PCR 产物纯化后连接到载体上,挑选 34 个阳性单克隆菌落进行测序,得到的序列用 Clustal X (1.83)进行比对,共获得 13 个

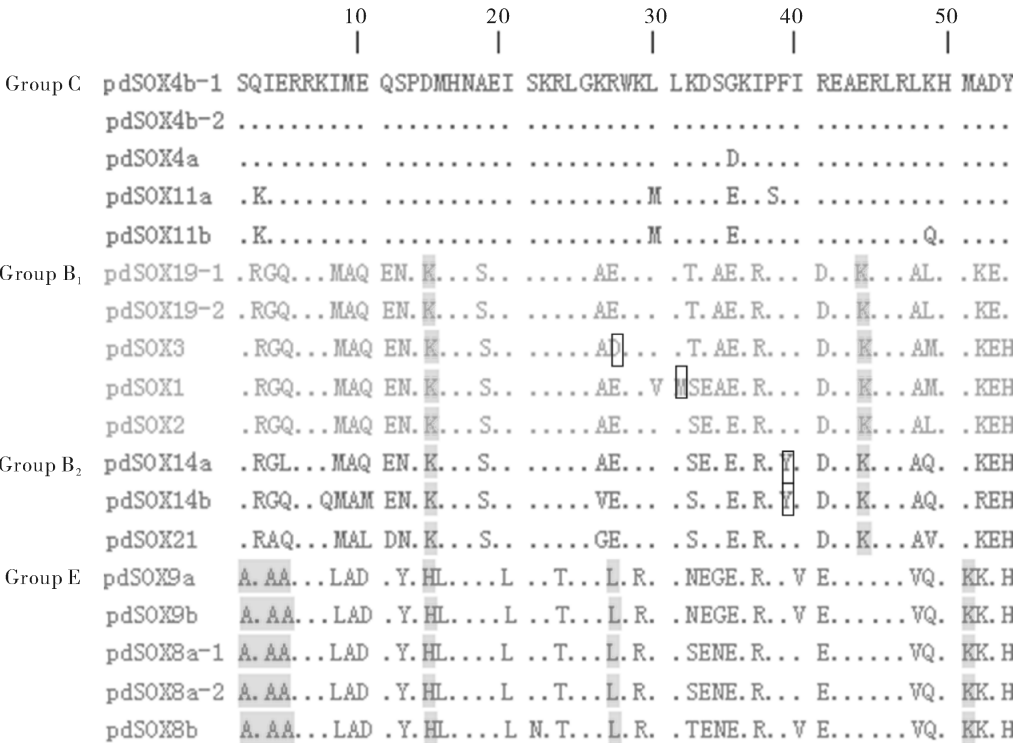
不同的序列,片段大小均为 204 bp。去除引物后的核苷酸序列见图 2,假定编码的氨基酸序列见图 3。依据 BlastX 同源搜索结果,结合 SOX 蛋白的标志性氨基酸残基分析结果(图 3)及聚类结果(图 4)等,将这 13 个基因分别命名为 *pdSox1*, *pdSox2*,

pdSox2 TCG AGG GGA CAG AGG AGG AAA ATG GCA CAG GAG AAC CCA AAA ATG CAC AAT TCG GAA ATA AGC AAA CGA CTC GGG GCC GAG
pdSox14a ...A..A..C... ..T..A..A..T..C... ..A... ..C... ..C... ..
PdSox14b ...C..A..T... ..A..A..C..G... ..T... ..G... ..C... ..C... ..G... ..G... ..A... ..C... ..T... ..
pdSox21 ...A..CT... ..A..C..A..G... ..T... ..T... ..G... ..C... ..T... ..G... ..C... ..G... ..A... ..GA... ..
pdSox3 ...T..C..T... ..C... ..G... ..A... ..T... ..T... ..C... ..C... ..C... ..G... ..T... ..T... ..C... ..
pdSox1 ...T..C... ..T... ..C..C..A..G... ..C... ..T... ..C... ..C... ..G... ..C... ..G... ..C... ..A... ..A... ..
pdSox19-1 ...T..C... ..T... ..C..C... ..G... ..A... ..C... ..C... ..C... ..T... ..C... ..T... ..G... ..T... ..G... ..C... ..A... ..
pdSox19-2 ...T..C... ..T... ..C..C... ..G... ..C... ..C... ..C... ..C... ..T... ..C... ..T... ..G... ..T... ..G... ..C... ..A... ..
pdSox4b-1 ...CAA ATC G.A... ..C..C... ..T ATG G... ..C... ..TCG... ..G G.C... ..C G... ..G G.C TCG... ..G A... ..T... ..C AAG CGC
pdSox4b-2 ...CAA ATC G.A... ..C..C... ..T ATG G... ..C... ..TCG... ..G G.C... ..C G... ..G G.C TCG... ..G A... ..T... ..C AAG CGC
pdSox4a ...CA. AT. G.A... ..C..C... ..G T ATG G... ..C... ..TCG... ..G G.C... ..C G... ..G G.C TC... ..G A... ..C AAA CGC
pdSox11b ...A.A. ATC G... ..C... ..C ATG G... ..C... ..TCT... ..G.T... ..C G.C... ..G G.C TC... ..A... ..G... ..A AAG CGA
pdSox11a ...T.AA ATC G... ..C... ..A... ..C ATG G... ..C... ..TCG... ..G G.C... ..C G.C... ..G G.C TC... ..C... ..G... ..C AAG CGC
pdSox2 TGG AAA CTT CTC TCC GAA AGC GAG AAG CGA CCT TTT ATC GAC GAA GCC AAG CGT CTT CGG GCG CTC CAC ATG AAG GAA CAC
pdSox14aG... ..T..A... ..TC... ..A... ..A... ..A..AC...G...A.G... ..C... ..T..AA...A... ..
pdSox14bT.G...T TCT... ..A..A... ..G..AC...G...A.A... ..G... ..C... ..A...GA...G... ..T
pdSox21G... ..G... ..C TCT... ..A...A...G...C... ..C... ..A...C... ..A...G... ..T
pdSox3T..G..A..T... ..T GCG...A...C... ..T...G...G.T..A...A... ..C..A..G...G... ..
pdSox1G..A..A..G...G GCT...A...C...G... ..G... ..A... ..A... ..A... ..T... ..A... ..G... ..T...G... ..T
pdSox19-1G... ..G... ..T..A... ..C GCG...A... ..G...G...A... ..A... ..T... ..C... ..G...G... ..T... ..
pdSox19-2G... ..G... ..T..A... ..C GCG...A... ..G...G...A... ..A... ..T... ..C... ..G...G... ..T... ..
pdSox4b-1G... ..G... ..G AAA... ..C... ..C... ..ATT... ..C... ..C... ..CGA... ..G... ..A... ..G... ..G... ..C..A... ..CTC AAG...GCA... ..C..T... ..
pdSox4b-2G... ..G... ..G AAA... ..C... ..GC... ..ATT... ..C... ..C... ..CGA... ..G... ..A... ..G... ..G... ..C..A... ..CTC AAG...GCA... ..C..T... ..
pdSox4aT..A AAA... ..C... ..C... ..ATC... ..C... ..C... ..CG...G...G... ..G... ..CTC AAG...GCA... ..C..T... ..
pdSox11bA.G... ..AAG... ..C... ..T... ..A... ..ATC...CGG...G... ..A... ..G... ..CTA... ..AA...GCC... ..T... ..T... ..
pdSox11aA.G.T.A AAG... ..C... ..T... ..A... ..AT... ..T... ..C... ..CGG...G... ..A... ..A... ..G... ..TT... ..AAA...GCT... ..T... ..T... ..

同首行等同的核苷酸用“-”表示 A base identical to the first line in each position is indicated by a dash.

图 2 SoxN 引物扩增大鳞副泥鳅基因组 DNA 获得的 Sox 基因的 HMG 盒核苷酸序列

Fig.2 Alignment of the nucleotide sequences encoding the HMG domains of Sox genes of *P. dabryanus* which were cloned from the genomic DNA with the SoxN primer



类群的标志性氨基酸残基被涂上阴影,用来确定 Sox 基因直系同源的标志性氨基酸残基被框住 Residues acting as signatures for different Sox groups are shaded and those used for assigning gene orthology are boxed.

图 3 大鳞副泥鳅 SOX 蛋白的 HMG 盒氨基酸序列

Fig.3 Alignment of the inferred amino acid sequences of SOX HMG domain of *P. dabryanus*

表 2 大鳞副泥鳅 Sox 基因的命名及直系同源证据¹⁾

Table 2 Summary of evidence for the Sox gene orthology and their nomination in *P. dabryanus*

命名 Nomination	与 Hu SRY 相似性/% Sequence identity to Hu SRY	与 Mo Sry 相似性/% Sequence identity to Mo Sry	相似性最高的基因(序列相似性/%) Gene with the highest identity (sequence identity)	HMG 蛋白残基 HMG signature residues	HMG 蛋白遗传距离 HMG protein genetic distance	GenBank 登录号 GenBank accession
<i>pdSox1</i>	67.6	63.2	<i>Carassius carassius red</i> var $\times$ <i>Megalobrama am- blycephala Sox1</i> (100)	SOX1	SOX1	JN700995
<i>pdSox2</i>	64.7	63.2	<i>Oreochromis niloticus Sox2</i> (100)	SOX2/19/21	SOX2	JN701001
<i>pdSox3</i>	69.1	64.7	<i>Oreochromis niloticus Sox3</i> (98)	SOX3	SOX3	JN700994
<i>pdSox4a</i>	52.9	57.3	<i>Oreochromis niloticus Sox4</i> (98)	SOX4	SOX4	JN701006
<i>pdSox4b-1</i>	52.9	57.3	<i>Oreochromis niloticus Sox4</i> (98)	SOX4	SOX4	JN701004
<i>pdSox4b-2</i>	51.4	55.8	<i>Oreochromis niloticus Sox4</i> (97)	SOX4	SOX4	JN701005
<i>pdSox8a-1</i>	32.3	35.2	<i>Oncorhynchus keta Sox8</i> (98)	SOX8/9	SOX8/9	JN701010
<i>pdSox8a-2</i>	32.3	35.2	<i>Oncorhynchus keta Sox8</i> (98)	SOX8/9	SOX8/9	JN701009
<i>pdSox8b</i>	25.0	29.4	<i>Oncorhynchus keta Sox8</i> (96)	SOX8/9	SOX8/9	JN701011
<i>pdSox9a</i>	26.4	30.8	<i>Cyprinus carpio Sox9c</i> (100)	SOX8/9	SOX9	JN701007
<i>pdSox9b</i>	26.4	30.8	<i>Cyprinus carpio Sox9c</i> (100)	SOX8/9	SOX9	JN701008
<i>pdSox11a</i>	54.4	57.3	<i>Oreochromis niloticus Sox11a</i> (98)	SOX11	SOX11	JN701003
<i>pdSox11b</i>	55.8	58.8	<i>Carassius carassius</i> $\times$ <i>Cyprinus carpio Sox11</i> (100)	SOX11	SOX11	JN701002
<i>pdSox14a</i>	63.2	61.7	<i>Andrias davidianus Sox14</i> (98)	SOX14	SOX14	JN700999
<i>pdSox14b</i>	61.7	57.3	<i>Oreochromis niloticus Sox14b</i> (92)	SOX14	Sox14	JN701000
<i>pdSox19-1</i>	69.1	67.6	<i>Ornithorhynchus anatinus Sox2</i> (98)	SOX2/19/21	SOX19	JN700997
<i>pdSox19-2</i>	69.1	67.6	<i>Ornithorhynchus anatinus Sox2</i> (98)	SOX2/19/21	SOX19	JN700996
<i>pdSox21</i>	64.7	58.8	<i>Cyprinus carpio Sox21b</i> (100)	SOX2/19/21	SOX21	JN700998

1) Hu: 人 Human; Mo: 小鼠 Mouse.

表 3 大鳞副泥鳅 18 个 Sox 基因之间氨基酸序列的相似性
Table 3 Amino acid sequence identities of pairs of eighteen Sox genes in *P. dabryanus*

	<i>pdSox1</i>	<i>pdSox2</i>	<i>pdSox3</i>	<i>pdSox4a</i>	<i>pdSox4b-1</i>	<i>pdSox4b-2</i>	<i>pdSox8a-1</i>	<i>pdSox8a-2</i>	<i>pdSox8b</i>	<i>pdSox9a</i>	<i>pdSox9b</i>	<i>pdSox11a</i>	<i>pdSox11b</i>	<i>pdSox14a</i>	<i>pdSox14b</i>	<i>pdSox19-1</i>	<i>pdSox19-2</i>	<i>pdSox21</i>
<i>pdSox1</i>	∖																	
<i>pdSox2</i>	92.6	∖																
<i>pdSox3</i>	92.6	91.1	∖															
<i>pdSox4a</i>	61.7	64.7	66.1	∖														
<i>pdSox4b-1</i>	61.7	64.7	66.1	98.5	∖													
<i>pdSox4b-2</i>	60.2	65.6	64.7	97	98.5	∖												
<i>pdSox8a-1</i>	32.3	34.3	33.8	36.7	36.7	37.3	∖											
<i>pdSox8a-2</i>	32.3	34.3	33.8	36.7	36.7	37.3	100.0	∖										
<i>pdSox8b</i>	29.4	32.8	30.8	36.7	36.7	37.3	57.5	57.5	∖									
<i>pdSox9a</i>	30.8	32.8	30.8	36.7	36.7	37.3	60.6	60.6	92.4	∖								
<i>pdSox9b</i>	30.8	32.8	30.8	36.7	36.7	37.3	60.6	60.6	92.4	100.0	∖							
<i>pdSox11a</i>	61.7	63.2	64.7	94.1	94.1	92.6	35.2	35.2	33.8	35.2	35.2	∖						
<i>pdSox11b</i>	63.2	64.7	66.1	94.1	94.1	92.6	35.2	35.2	33.8	35.2	35.2	97.0	∖					
<i>pdSox14a</i>	91.1	94.1	89.7	64.7	64.7	63.2	33.8	33.8	30.8	30.8	30.8	63.2	66.1	∖				
<i>pdSox14b</i>	85.2	88.2	86.7	64.7	64.7	63.2	30.8	30.8	29.4	29.4	29.4	63.2	66.1	91.1	∖			
<i>pdSox19-1</i>	91.1	92.6	95.5	67.6	67.6	66.1	33.8	33.8	32.3	32.3	32.3	66.1	67.6	89.7	86.7	∖		
<i>pdSox19-2</i>	91.1	92.6	95.5	67.6	67.6	66.1	33.8	33.8	32.3	32.3	32.3	66.1	67.6	89.7	86.7	100.0	∖	
<i>pdSox21</i>	86.7	89.7	88.2	67.6	67.6	66.1	36.7	36.7	32.3	32.3	32.3	66.1	67.6	88.2	88.2	88.2	88.2	∖

pdSox3, *pdSox4a*, *pdSox4b-1*, *pdSox4b-2*, *pdSox11a*, *pdSox11b*, *pdSox14a*, *pdSox14b*, *pdSox19-1*, *pdSox19-2*, *pdSox21*(表2)。其中,除 *pdSox11a* 及 *pdSox14b* 各只有1个克隆外,其他基因均有2个以上的克隆,且 *pdSox11a* 与它的拷贝 *pdSox11b*、*pdSox14b* 与它的拷贝 *pdSox14a* 的核苷酸序列相差超过20个核苷酸,不可能是测序错误。13个基因中, *pdSox11*、*pdSox14* 及 *pdSox19* 各有2个拷贝, *pdSox4* 有3个拷贝,其中, *pdSox4b-1* 与 *pdSox4b-2* 只相差1个第二位置的核苷酸, *pdSox19-1* 与 *pdSox19-2* 只相差2个第三位置的核苷酸(图2)。将这13个基因与 GenBank 中已有的大鳞副泥鳅 *Sox* 基因进行比对,发现 *pdSox11b* 与 GenBank 中大鳞副泥鳅的 *Sox11a* (AY685137)、*Sox11b* (AY685136) 及 *Sox11c* (AY685140) 去除引物后的核苷酸序列完全一致,应是同一基因; *pdSox21* 与 GenBank 中大鳞副泥鳅的 *Sox21a* (AY685135)、*Sox21b* (AY685139) 的核苷酸序列相差超过10个核苷酸,是不同的拷贝; *pdSox19-1*、*pdSox19-2* 与 GenBank 中大鳞副泥鳅的 *Sox19* (AY685138) 的核苷酸序列相差超过20个核苷酸,是不同的拷贝,而 *pdSox1*、*pdSox2*、*pdSox3*、*pdSox4* 及 *pdSox14* 是本研究最新发现的大鳞副泥鳅 *Sox* 基因家族的成员。依据 Koopman 等^[10] 的分类标准,将 *pdSox1*、*pdSox2*、*pdSox3* 与 *pdSox19* 等5个基因归于 B₁ 类群, *pdSox14* 与 *pdSox21* 等3个基因归于 B₂ 类群, *pdSox4* 与 *pdSox11* 等5个基因归于 C 类群(图3)。

将 *Sox9* 引物扩增的大鳞副泥鳅基因组 DNA 获得的5条电泳带逐一测序,得到的最小的电泳条带为190 bp,与应用 *SoxN* 引物扩增获得的 *pdSox14a* 的核苷酸序列完全一致,应是同一基因;其他4个片段按照片段大小依次命名为 *pdSox8a-1/pdSox8a-2* (531 bp/589 bp), *pdSox9a* (487 bp), *pdSox8b* (431 bp) 与 *pdSox9b* (381 bp) (图1),其中, *pdSox8a-1* 与 *pdSox8a-2* 的外显子序列几乎完全一致,仅相差1个第三位置的核苷酸,二者编码相同的氨基酸(图3),它们分别包含339、397 bp 的内含子序列, *pdSox8a-1* 的内含子比 *pdSox8a-2* 小58 bp,并且,二者的内含子序列还相差19个核苷酸。

pdSox9 和 *pdSox8* 等5个基因属于 *SoxE* 类

群,将它们与 GenBank 中已有的大鳞副泥鳅 *SoxE* 类群基因进行比对,发现 *pdSox8a-1* (531 bp) 与 Xia 等^[8] 克隆的 *pdSox8b* 的外显子序列完全一致,在内含子区二者仅相差7个核苷酸(包括4个核苷酸置换,3个核苷酸缺失),它们应是同一基因; *pdSox8b* 与 Xia 等^[8] 克隆的 *pdSox8a* 去除引物后的核苷酸序列在外显子区仅相差1个第二位置的核苷酸,导致编码的氨基酸序列相差1个,可能是 PCR 扩增、测序错误所致,在内含子区也仅相差1个核苷酸,它们应是同一基因。Xia 等^[8] 克隆的 *pdSox8b* 与早期克隆的 *Sox8* (AF116250, 500 bp)^[5-6] 应是同一基因;常重杰等^[5,7] 克隆的 *Sox9* (AF116249) 与本研究 *pdSoxE* 类群的所有基因的核苷酸序列相差均很大,且其编码的氨基酸如 ARAA 残基与 *SoxE* 类群的标志性氨基酸残基 AQAA (图3) 不符,因此未采用。

Sox9 引物扩增的大鳞副泥鳅精巢 cDNA 得到的序列与 *pdSox9a* 共有的核苷酸序列完全一致,表明精巢表达的是 *pdSox9a*, *pdSox9a* 含有295 bp 内含子;扩增卵巢 cDNA 得到的序列与 *pdSox8a-2* 共有的核苷酸序列完全一致,与 *pdSox8a-1* 仅相差1个第三位置的核苷酸,表明卵巢表达的是 *pdSox8a*。结合 BlastX 同源搜索,发现大鳞副泥鳅 *pdSoxE* 类群中5个基因片段均含有1个内含子,且内含子的起始位置一致(从编码的第29位精氨酸 R 开始), *pdSox9b*、*pdSox8b* 各自含有189 bp、239 bp 的内含子。

本研究共获得18个大鳞副泥鳅 *Sox* 基因,其中, *pdSox4b-1* 与 *pdSox4b-2*、*pdSox19-1* 与 *pdSox19-2*、*pdSox8a-1* 与 *pdSox8a-2*、*pdSox9a* 与 *pdSox9b*,各自编码相同的氨基酸。18个 *Sox* 基因之间氨基酸序列的相似性见表3。

2.3 系统聚类分析

将大鳞副泥鳅 *Sox* 基因与小鼠相应的 *Sox* 基因以及斑马鱼的 *Sox19*、*Sox21* 基因,依照 HMG 盒蛋白质的遗传距离构建系统进化树。3种聚类方法(最大简约法 MP、最小进化法 ME 与邻接法 NJ)构建的进化树几乎完全一致。18个大鳞副泥鳅 *Sox* 基因分别归于 B、C、E 3个类群,其中, E 类群与 C 类群距离较近,与 B 类群距离较远,而 B 类群又分为2小类群,其中, *pdSox21* 并未与 *pdSox14* 聚在一起,而是单独聚为一小类(图4)。

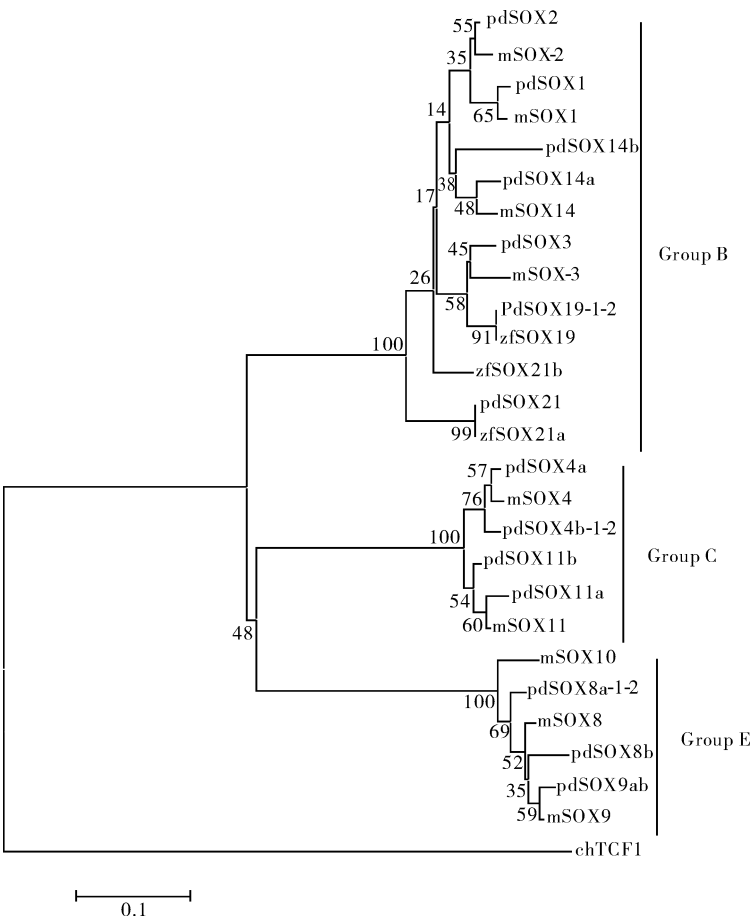


图 4 基于大鳞副泥鳅与小鼠 Sox 基因的 HMG 盒蛋白遗传距离构建的 NJ 系统进化树

Fig. 4 Phylogenetic tree based on the genetic distance of SOX HMG box proteins of *P. dabryanus* and *Mus musculus* by using the neighbor algorithm

3 讨 论

自 1997 年有学者认为大鳞副泥鳅具有异形性染色体^[4]之后,相继有许多学者对其产生兴趣,并采用多种方式、从多个角度探讨大鳞副泥鳅的性别决定和性别分化机制^[12-14]。*Sox* 基因家族的一些成员如 *Sox3*、*Sox5*、*Sox6*、*Sox8*、*Sox9*、*Sox17* 等被认为与动物的性别决定或性别分化相关^[2]。本研究采用 HMG 盒简并引物进行扩增,共克隆到 18 个大鳞副泥鳅 *Sox* 基因,分别归于 *SoxB*、*SoxC* 及 *SoxE* 3 个类群,其中 *pdSoxE* 类群包含内含子。通过构建 *Sox* 基因的 HMG 盒蛋白遗传距离的系统进化树,

每个大鳞副泥鳅 *Sox* 基因在哺乳动物或斑马鱼中均能找到相应的 *Sox* 直系同源基因(图 4),并且在 *pdSoxB*、*pdSoxC* 类群中各大鳞副泥鳅 *Sox* 基因与哺乳动物人、小鼠 *SRY/Sry* 的 HMG 盒氨基酸序列的相似性均超过 50%(表 2),符合 *Sox* 基因的命名法则。在 *pdSoxE* 类群中,*pdSox8* 与 *pdSox9* 等 5 个基因与人、小鼠 *SRY/Sry* 的氨基酸序列的相似性低于 35.2%(表 2),与大鳞副泥鳅其他 *Sox* 基因的氨基酸序列的相似性也不超过 35.2%(表 3),但是它们与一些已经报道的鱼类 *Sox8*、*Sox9* 基因的同源性很高(高于 96%;表 2),因此,笔者认定本研究中克隆的这批 *Sox* 基因是大鳞副泥鳅 *Sox* 基因

家族成员。

在所克隆的 18 个大鳞副泥鳅 *Sox* 基因中,多个基因存在基因复制现象。例如,大鳞副泥鳅的 *Sox4*、*Sox8*、*Sox9*、*Sox11*、*Sox14*、*Sox19* 等 6 个家族成员均存在 2~3 个拷贝,这提示脊椎动物在硬骨鱼类分化之前曾发生过一次全基因组复制事件^[15]。其中,*Sox19* 基因为鱼类所特有^[15]。大鳞副泥鳅 *Sox19* 基因存在基因复制现象,进一步证实鱼类在进化过程中曾发生过另外的部分或全基因组复制事件^[16],例如自然界中的四倍体泥鳅(*Misgurnus anguillicaudatus*)可能就是由二倍体泥鳅自然进化而来^[17]。将大鳞副泥鳅 *Sox* 基因与小鼠相应的 SOX 蛋白进行聚类,*pdSox* 基因的各拷贝间显示一种“外类群”拓扑结构而不是预期的“姊妹群”拓扑结构(图 4),暗示大鳞副泥鳅 *Sox* 基因的各拷贝间存在不同的进化速率^[18]。

小鼠的 *Sox9* 已被证实在雄性的性别分化过程中发挥重要作用,它的表达激活了下游一系列与雄性性别发育相关的基因的表达,最终促使哺乳动物性腺分化成精巢^[19];而 *Sox8* 则起着强化 *Sox9* 功能的作用,甚至能部分替代 *Sox9* 的功能^[20-21]。迄今,在一些鱼类中已经克隆到 *Sox9* 基因,并且多数鱼类的 *Sox9* 基因存在基因复制现象,它们主要在精巢或卵巢中表达^[22-23]。斑马鱼 *Sox9a* 主要在精巢中表达,但在其成鱼的第 II 期卵母细胞的卵浆中却同时检测到 *Sox8* 和 *Sox9b* 的表达^[24],这暗示 *Sox9*、*Sox8* 可能在硬骨鱼类的性别分化过程中发挥一定作用。本研究克隆到大鳞副泥鳅的 2 个 *pd-Sox9* 和 3 个 *pdSox8* 基因,并且发现 *pdSox8a* 在大鳞副泥鳅的不同个体中呈现多态性,分别表现为 *pdSox8a-1* 或 *pdSox8a-2*,这是硬骨鱼类中 *Sox8* 基因存在个体多态性的首次报道。本文还在大鳞副泥鳅的精巢与卵巢中分别检测到 *pdSox9a*、*pdSox8a* 的表达,这与常重杰等^[7]发现 *pdSox9* 在精巢、脑中宏量表达的结果一致。Xia 等^[8]发现 *pdSox8a* 与 *pdSox8b* 主要在大鳞副泥鳅的脑与性腺中表达,并且,在精巢中的表达量均高于卵巢,认为 *pdSox8* 和 *pdSox9* 可能与大鳞副泥鳅的性腺分化有关^[7-8]。本研究在大鳞副泥鳅的卵巢中仅克隆到 *pdSox8a* 基因,未能克隆到 *pdSox8b* 基因。笔者的结果以及他人的证据表明,*pdSox9*、*pdSox8* 是大鳞副泥鳅的性别相关基因,它们可能在其性别分化过程中发挥作用。至于这 2 种 *Sox* 基因是否与大鳞副泥鳅的

性别决定机制相关,尚需开展深入的基因功能研究才能提供线索。

常重杰等^[6]报道大鳞副泥鳅 *Sox8* 基因具有内含子。本研究进一步发现大鳞副泥鳅 *Sox8a* 基因存在 2 个拷贝(*pdSox8a-1* 与 *pdSox8a-2*),它们的外显子序列几乎完全一致,其内含子序列存在很大的差异性,并且长度也不相同。此外,本研究还发现 *pdSox9a* 与 *pdSox9b* 的内含子序列也存在多态性。其他鱼类的 *Sox9* 基因也存在类似的现象,例如不同品系或群体的鲤至少存在 5 个有活性的不同的 *Sox9* 基因,这些基因的内含子中存在大量的序列多态,导致它们在鱼体组织中的表达强度或表达场所、丰度有差异^[25]。大鳞副泥鳅的 *Sox9* 和 *Sox8* 基因的内含子序列的多态性是否与 *pdSox9*、*pdSox8* 在性别分化过程中发挥的功能相关,有待进一步的研究证实。

参 考 文 献

- [1] BOWLES J, SCHEPERS G, KOOPMAN P. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators[J]. *Developmental Biology*, 2000, 227: 239-255.
- [2] 曹慧慧,王昌留. *Sox* 基因家族特点及其功能[J]. 山东大学学报:自然科学版, 2011, 27(1): 58-63.
- [3] 江玲霞,李纪委,贺彧,等. *Sox* 基因家族功能的研究进展[J]. 生物技术通报, 2008(6): 44-48.
- [4] 常重杰,于其兴. 大鳞副泥鳅 ZZ/ZW 型性别决定的细胞遗传学证据[J]. 遗传, 1997(3): 17-19.
- [5] 常重杰,周荣家,余其兴. 两种泥鳅中 *PdSox8* 和 *PdSox9* 基因的染色体定位[J]. 遗传学报, 2000, 27(5): 377-382.
- [6] 常重杰,杜启艳,周荣家,等. 具有内含子的大鳞副泥鳅 *Sox8* 基因[J]. 发育与繁殖生物学, 2000, 9(2): 15-22.
- [7] 常重杰,周荣家,余其兴. 大鳞副泥鳅中 *Sox9* 基因保守区的序列分析[J]. 遗传学报, 2000, 27(2): 121-126.
- [8] XIA X H, ZHAO J, DU Q, et al. cDNA cloning and expression analysis of two distinct *Sox8* genes in *Paramisgurnus dabryanus* (Cypriniformes)[J]. *Journal of Genetics*, 2010, 89: 183-192.
- [9] 萨姆布鲁克 J, 拉舍尔 D W. 分子克隆实验指南 [M]. 3 版. 黄培堂,译. 北京:科学出版社, 2002: 461-471.
- [10] GALAY-BURGOS M, LLEWELLYN L, MYLONAS C C, et al. Analysis of the *Sox* gene family in the European sea bass (*Dicentrarchus labrax*) [J]. *Comparative Biochemistry and Physiology*, 2004, 137: 279-284.
- [11] KOOPMAN P, SCHEPERS G, BRENNER S, et al. Origin and diversity of the *Sox* transcription factor gene family: genome-wide analysis in *Fugu rubripes* [J]. *Gene*, 2004, 328: 177-186.
- [12] YOU C, YU X, TAN D Q, et al. Gynogenesis and sex determination in large scale loach, *Paramisgurnus dabryanus* (Sau-

- vage)[J]. *Aquaculture International*, 2008, 16: 203-214.
- [13] 徐玲花, 严镇钧, 曾庆韬. 大鳞副泥鳅 3 个 *Dmrt* 基因 DM 保守区的序列分析[J]. *遗传*, 2008, 30(11): 1448-1452.
- [14] 南平, 杜启艳, 燕帅国, 等. 温度对泥鳅和大鳞副泥鳅性腺分化的影响和 *CYP19a* 基因的克隆与时空表达[J]. *中国水产科学*, 2005, 12(4): 407-413.
- [15] MAZZUCHELLI J, YANG F, KOCHER T D, et al. Comparative cytogenetic mapping of *Sox2* and *Sox14* in cichlid fishes and inferences on the genomic organization of both genes in vertebrates[J]. *Chromosome Research*, 2011, 19: 657-667.
- [16] CHRISTOFFELS A, KOH E G, CHIA J M, et al. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes[J]. *Molecular Biology and Evolution*, 2004, 21(6): 1146-1151.
- [17] 方礼豹, 周小云, 崔蕾, 等. 二倍体、四倍体泥鳅与大鳞副泥鳅杂交子代 DNA 相对含量与染色体组型的比较[J]. *华中农业大学学报*, 2011, 30(4): 500-505.
- [18] ZHONG L, YU X, TONG J. *Sox* genes in grass carp (*Ctenopharyngodon idella*) with their implications for genome duplication and evolution[J]. *Genetics Selection Evolution*, 2006, 38(6): 673-687.
- [19] CLARKSON M J, HARLEY V R. Sex with two SOX on; SRY and SOX9 in testis development[J]. *Trends in Endocrinology & Metabolism*, 2002, 13(3): 106-111.
- [20] KOOPMAN P. Sex determination: a tale of two *Sox* genes[J]. *Trends in Genetics*, 2005, 21(7): 367-370.
- [21] CHABOISSIER M C, KOBAYASHI A, VIDAL V I P, et al. Functional analysis of *Sox8* and *Sox9* during sex determination in the mouse[J]. *Development*, 2004, 131: 1891-1901.
- [22] CHIANG E F, PAI C I, WYATT M, et al. Two *Sox9* genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites[J]. *Developmental Biology*, 2001, 231: 149-163.
- [23] KLUVER N, KONDO M, HERPIN A, et al. Divergent expression patterns of *Sox9* duplicates in teleosts indicate a lineage specific subfunctionalization[J]. *Development Genes and Evolution*, 2005, 215: 297-305.
- [24] RODRIGUEZ-MARI A, YAN Y L, BREMILLER R A, et al. Characterization and expression pattern of zebrafish anti-müllerian hormone (*Amh*) relative to *sox9a*, *sox9b*, and *cyp19a1a*, during gonad development [J]. *Gene Expression Patterns*, 2005, 5: 655-667.
- [25] 华慧颖, 王芳, 常重杰, 等. 鲤鱼 *Sox9* 基因 HMG 保守区的克隆及序列分析[J]. *河南农业科学*, 2011, 40(7): 147-151.

Cloning and characterization of conservative region sequences of *Sox* genes in *Paramisgurnus dabryanus*

YOU Cui-hong^{1,2} TONG Jin-gou¹ YU Xiao-mu¹

1. State Key Laboratory of Freshwater Ecology and Biotechnology/
Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China;
2. Marine Biology Institute of Shantou University, Shantou 515063, China

Abstract In order to further investigate the sex determination and sex differentiation mechanism of large scale loach *Paramisgurnus dabryanus*, two degenerate primers, SoxN and Sox9, were used to carry out PCR amplification of the genomic DNA and RT-PCR analysis of the gonad tissues. Thirteen genes of the *Sox* family including *pdSox1*, *pdSox2*, *pdSox3*, *pdSox4*, *pdSox11*, *pdSox14*, *pdSox19* and *pdSox21* were cloned from the genomic DNA of *P. dabryanus* with the SoxN primer and classified as the *SoxB* and *SoxC* groups. Another five *Sox* genes including *pdSox8* and *pdSox9*, with one intron starting at the same position, were obtained by using the Sox9 primer and classified as the *SoxE* group. Based on the genetic distance of the HMG-box amino acid, all the eighteen *pdSox* were orthologous to the *Sox* genes of mouse or zebrafish on the NJ phylogenetic tree. Six of the 13 *Sox* genes including *pdSox4*, *pdSox8*, *pdSox9*, *pdSox11*, *pdSox14* and *pdSox19* had two or more copies. Moreover, the results of RT-PCR and sequence analyses with the Sox9 primer showed that *pdSox9a* was expressed in the testis whereas *pd-sox8a* was expressed in the ovary. The present results suggested that *pdSox8* and *pdSox9* were sex-related genes and they might play a role in the sex differentiation of *P. dabryanus*.

Key words *Paramisgurnus dabryanus*; *Sox* gene; HMG box; sex determination; sex differentiation
(责任编辑: 边书京)