

DOI: 10.3724/SP.J.1118.2015.14527

中国大鲵 *Sox19* 基因全长 cDNA 序列的克隆及组织表达分析

刘静, 何青, 刘丽丽, 李成磊, 王勤

四川农业大学 生命科学院, 四川 雅安 625014

摘要: 采用逆转录聚合酶链反应(RT-PCR)及 cDNA 末端快速扩增(RACE)技术克隆了中国大鲵(*Andrias davidianus*) *Sox19* 基因全长 cDNA 序列, 并进行各组织间的表达分析。中国大鲵 *Sox19* 基因全长 1290 bp, ORF 长 858 bp, 编码 285 个氨基酸, 位于 39~109 位的 HMG-box 是其主要功能区。氨基酸序列分析表明, 其与舌齿鲈(*Dicentrarchus labrax*)、斑马鱼(*Danio rerio*)及红鳍东方鲀(*Takifugu rubripes*)的相似度分别为 64%、58%、55%。采用邻接法对不同物种的 *Sox19* 编码区序列构建分子系统树发现: 中国大鲵 *Sox19* 基因属于较早从鱼类分化出来的基因, 表明 *Sox19* 基因从中国大鲵到鱼类的进化速度比较快。荧光定量 PCR 结果显示, *Sox19* 在所有检测的组织中均有表达, 且在心脏中的表达量最高, 卵巢中次之, 肾脏中的表达量略高于肠道。中国大鲵 *Sox19* 基因的发现打破了 *Sox19* 基因一直被认为是鱼类所特有基因的观点, 填补了两栖类有关此基因的空白, 为进一步了解 *Sox19* 基因在中国大鲵中的重要功能奠定了基础, 同时 *Sox19* 基因还可作为研究中国大鲵早期胚胎神经系统发育情况的分子标记, 为其他两栖类动物的研究提供借鉴。

关键词: 中国大鲵; *Sox19* 基因; 基因克隆; 基因组织表达

中图分类号: S917

文献标志码: A

文章编号: 1005-8737-(2015)06-1142-08

Sox 家族是一类与哺乳动物性别决定基因 *SRY* (sex determination region of Y chromosome) 相关基因构成的基因家族, 编码一组进化上高度保守、结构上与 *SRY* 相关的转录因子, 其家族成员均含有一个高度保守的 HMG-box DNA 结合域, HMG-box 由 70 多个氨基酸组成, 通过与 DNA 小沟的 WACAAW 模体结合, 从而导致 DNA 构象发生变化^[1]。*Sox* 基因家族在动物界广泛存在并在个体发育过程中发挥至关重要的作用, 尤其在性别决定、神经嵴及神经发育过程等多个方面发挥重要的生物学功能^[2-3], 例如, *Sox9*、*Sox5*、*Sox6*、*Sox3*、*Sox8* 等对性别决定与分化起重要作用, *Sox1*、*Sox2*、*Sox3*、*Sox13*、*Sox21* 等参与早期胚胎发生, *Sox3*、*Sox10*、*Sox19*、*Sox21* 和 *Sox22* 等参与神经系统的发育、眼的发育、骨形成和生血

细胞的发生等^[4-5]。Bowles 等^[6]根据 HMG 盒的相似性, 将 *Sox* 基因分成 A–J 共 10 个亚族, 其中 *Sox19* 基因是 *Sox* 基因家族 B1 亚族成员之一, 最初由 Vríz 等^[7-8]和 Winata 等^[9]从斑马鱼(*Danio rerio*)中分离, 并通过原位杂交技术在原肠胚早期中枢神经系统的圆形区域内检测到 *Sox19* 基因的表达, 并发现 *Sox19* 基因不仅在端脑, 在间脑腹侧区、中脑和后脑都能检测到其表达, 同时又进一步证明了 *Sox19* 基因主要参与早期胚胎神经系统的发育。随后在红鳍东方鲀(*Takifugu rubripes*)^[10]及舌齿鲈(*Dicentrarchus labrax*)^[11]中得到了 *Sox19* 全长序列, Navarro-Martín 等^[11]还发现其在舌齿鲈的多个组织尤其在大脑、性腺及皮肤有较高的表达量。*Sox19* 基因一直被认为是鱼类所特有的基因^[10-12], 目前对其研究较少, 除了这 3 种鱼类得到全长

收稿日期: 2015-01-08; 修订日期: 2015-03-03.

基金项目: 四川省科技厅应用基础项目(2011JY0103).

作者简介: 刘静(1987–), 女, 硕士研究生, 从事动物分子生物学研究. E-mail: liujingabcd123@163.com

通信作者: 王勤(1968–), 教授. E-mail: wangqin3269@163.com

cDNA 序列外, 仅在大鳞副泥鳅(*Paramisgurnus dabryanus*)、尼罗罗非鱼(*Oreochromis niloticus*)、黄鳝(*Monopterus albus*)^[12]、比目鱼(*Hippoglossus*)、斑点鲟(*Acipenser sturio*)^[13]及鳙鱼(*Aristichthys nobilis*)^[14]中得到部分序列, 此外, Hiraoka 等^[15]在小鼠中也分离得到类似 *Sox19* 基因的部分序列, 但和其他 *Sox19* 基因相似性极低。

中国大鲵(*Andrias davidianus*)属两栖纲(Amphibia), 有尾目(Caudata), 隐鳃鲵科(Cryptobranchidae), 大鲵属, 是中国特有的大型珍稀二级保护物种, 也是中国目前唯一尚存的隐鳃鲵科两栖动物。但是由于生态破坏和过度捕捞, 自然种群正在逐渐减少, 而在人工养殖中, 存在产卵率低、受精率低、孵化率低的养殖障碍^[16-17]。本研究利用 RT-PCR 和 RACE 方法首次从两栖动物中国大鲵中克隆 *Sox19* 基因的全长 cDNA 序列, 并根据推导的氨基酸序列与其他物种相应序列进行同源性比较, 构建中国大鲵 *Sox19* 基因的系统进化树, 同时进行该基因在中国大鲵不同组织中的 mRNA 表达分析, 为进一步研究该基因的功能提供基础资料。

1 材料与方法

1.1 实验材料

实验用 1 尾已经性成熟的健康无病雌性大鲵,

来源于四川省乐山市大鲵养殖户, 体重为 2.13 kg, 体长 56 cm、头宽 9 cm、体宽 9.5 cm、尾长 26 cm, 确保大鲵健康无病后开始实验。解剖后取卵巢、心脏、肾、肠等内脏组织, 经液氮速冻后冻存于 -80℃ 超低温冰柜备用。

1.2 方法

1.2.1 总 RNA 的提取及 cDNA 第一链的合成 取冻存的中国大鲵各组织, 参照 RNAisoTM Plus (TaKaRa)操作说明提取总 RNA, 经紫外分光光度计测定纯度并用琼脂糖凝胶电泳检测完整度。以符合要求的总 RNA 为模板, 根据 RevertAidTM First Strand cDNA Synthesis Kit (Fermentas)说明书反转录合成 cDNA 第一链, 作为保守片段的扩增及荧光定量 PCR 的模板。

1.2.2 保守片段的克隆 根据不同物种 *Sox19* 基因保守的 HMG-box 区设计简并引物 Sox19-BS-F/Sox19-BS-R, 引物由上海 Invitrogen 生物技术有限公司合成(表 1)。反应体系: ddH₂O 20 μL, 10×Buffer 3 μL, 25 mmol/L MgCl₂ 1.5 μL, dNTP mix 2 μL, 引物 Sox19-BS-F 和 Sox19-BS-R 各 1 μL, cDNA 模板 1 μL, rTaq DNA 聚合酶 0.5 μL (5 U/μL), 总体积 30 μL。反应参数: 94℃ 5 min; 94℃ 30 s, 54℃ 50 s, 72℃ 1 min, 30 个循环; 72℃ 10 min。PCR 产物回收、纯化, 克隆和挑选阳性克隆送测。

表 1 扩增完整中国大鲵的 *Sox19* 基因 cDNA 所用引物
Tab. 1 Primer pairs used for amplification and sequencing of the complete *Sox19* cDNA in *Andrias davidianus*

引物名称 primer name	引物序列(5'-3') prime sequence(5'-3')	扩增目标 amplification target
Sox19-BS-F	5'-AATGCCTTCATGGTTTGGT-3'	保守片段的扩增
Sox19-BS-R	5'-GTAATCGGGGTACTCCTTC-3'	
Sox19-3-F1	5'-CTTCATGGTTTGGTCCAGTACCCAGCGC-3'	
Sox19-3-F2	5'-AACCCCAAAATGCACAACCTCCGAGATCAGC-3'	3'RACE
Sox19-3-F3	5'-GCTGAGTGGAAGCTCCTGAGTGACGCGG-3'	
Sox19-5-R1	5'-TAATCGGGGTACTCCTTCATGTGGGTGG-3'	
Sox19-5-R2	5'-CTTGGCCTCGTCAATAAAGGGGCGCTTCTC-3'	5'RACE
Sox19-5-R3	5'-CGCTTGCTGATCTCGGAGTTGTGCATTTTG-3'	
Sox19-F	5'-GCGTCTGCACCGAGTAAGCCAT-3'	
Sox19-R	5'-TGATCCAGTCGTTCCGAGTCGC-3'	ORF
β-F	GCCGTGACCTGACAGACTACCT	
β-R	AGTCCAGGGCGACATAGCAGAG	
Sox19-DL-F	GACATTCTTGGGCGTCTGCACCGAG	RT-PCR
Sox19-DL-R	GGACGTGTGCTGAGGCATCTGGACCT	

1.2.3 中国大鲵 *Sox19* 基因 cDNA 全长克隆 根据已获得的 *Sox19* 基因的保守片段, 分别设计 3 条特异引物 F3-1/F3-2/F3-3 进行 3'-RACE 反应, 设计 R5-1/R5-2/R5-3 进行 5'-RACE 反应(表 1), 参照 SMARTerTM RACE cDNA Amplification Kit 说明书进行 RACE 反应, 分别将第三轮的 PCR 产物回收、纯化、克隆和测序。将得到的测序结果拼接后设计一对特异引物 Sox19-F/Sox19-R (表 1), 以 cDNA 为模板扩增 *Sox19* 基因的 ORF 序列, PCR 产物回收、纯化、克隆和测序, 获得完整的大鲵 *Sox19* 的 cDNA 序列。

1.3 序列分析方法

使用 MultAlin (<http://multalin.toulouse.inra.fr/multalin/multalin.html>)进行氨基酸序列的多重序列比对, 用 PSORT (<http://psort.hgc.jp/form.html>)分析亚细胞定位, 用 ExPASy (<http://ca.expasy.org/tools/protscale.html>)分析其亲疏水性, 用 TMPred (http://www.ch.embnet.org/software/TMPRED_form.html)分析其跨膜区和跨膜方向, 用 SignalP 4.1 Server (<http://www.cbs.dtu.dk/services/SignalP/>)对其蛋白质序列进行信号肽分析, 用 SOPMA (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)及 PredictProtein (<https://www.predictprotein.org/>)分析蛋白质二级结构, 用 SMART(<http://smart.embl-heidelberg.de/>)分析其结构域, 用 SWISS-MODEL (<http://swissmodel.expasy.org/>)预测其蛋白质三级结构, 用 MEGA 5.0 构建系统发育树。

1.4 荧光定量 PCR 检测 *Sox19* 基因在各组织中的表达

按照前述方法对健康的中国大鲵心脏、肾、卵巢和肠分别提取总 RNA, 并反转录合成 cDNA 作为荧光定量 PCR 的模板。根据获得的 *Sox19* 基因序列, 设计定量 PCR 的特异性引物 Sox19-DL-F/Sox19-DL-R (表 1), 并以大鲵 β -actin 基因作为内参基因(表 1), 10 倍梯度稀释各个组织的 cDNA 模板制作荧光定量 PCR 标准曲线后, 对各组织中 *Sox19* 基因的表达进行检测。反应体系为 12.5 μ L: 6.25 μ L SYBR Premix Ex TaqTM, 0.5 μ L PCR For-

ward Primer (10 μ mol/L), 0.5 μ L PCR Reverse Primer (10 μ mol/L), 1 μ L cDNA; 反应参数: 95 $^{\circ}$ C 30 s; 接着 95 $^{\circ}$ C 5 s, 66 $^{\circ}$ C 30 s, 44 个循环; 熔解曲线 60~95 $^{\circ}$ C。样品和内参分别设 3 个重复, 采用 CT 值法分析数据, 比较中国大鲵 *Sox19* 基因在卵巢、心脏、肾和肠的表达差异。

2 结果与分析

2.1 中国大鲵 *Sox19* 编码的蛋白质序列分析

本实验获得的中国大鲵 *Sox19* 基因 cDNA 全长为 1533 bp, 经在 NCBI 中进行序列比对, 证明该基因为 *Sox19* 基因, 已被 GeneBank 收录, 序列号为 KJ623265。该基因 cDNA 序列开放阅读框 (Open Reading Fragment, ORF) 包含 858 个碱基序列, 5'-UTR 和 3'-UTR 分别为 270 bp 和 162 bp, 其氨基酸序列共编码 285 个氨基酸。应用生物信息学对中国大鲵 *Sox19* 编码的蛋白进行分析预测表明: 中国大鲵 *Sox19* 蛋白为亲水性蛋白; 亚细胞定位显示其 96% 都存在于细胞核中; 没有跨膜区, 不含信号肽; 二级结构中 α -螺旋有 24.91%, 延伸链 3.16%, β -转角 2.11%, 无规则卷曲 69.82%; 三级结构显示此基因有明显的 α -螺旋结构, 这可能是 *Sox19* 基因的功能位点; 位于 39~109 位的 HMG-box 结构域是唯一的结构域, 是其主要功能区。

与 NCBI 已记录的所有物种, 即 8 种鱼类的 *Sox19* 编码的氨基酸序列进行比对(图 1), 其与舌齿鲈、斑马鱼和红鳍东方鲀相似度分别为 64%、58%、55%; 此外, 图 1 也显示出中国大鲵 *Sox19* 蛋白的 HMG-box 区域相对比较保守, 仅有小部分氨基酸与其他物种有差别。

2.2 中国大鲵 *Sox19* 编码的氨基酸进化分析

选取不同物种 *Sox19* 基因编码的蛋白质序列, 利用 MEGA 5.0 软件, 采用邻接法构建系统发育树(图 2)。结果显示, 中国大鲵 *Sox19* 基因属于较早从鱼类分化出来的基因, 显示其较原始的进化地位, 基本符合物种进化的历程。

2.3 荧光定量 PT-PCR 检测 *Sox19* 基因在组织中的表达

以 Sox19-DL-F/Sox19-DL-R 为荧光定量 PCR

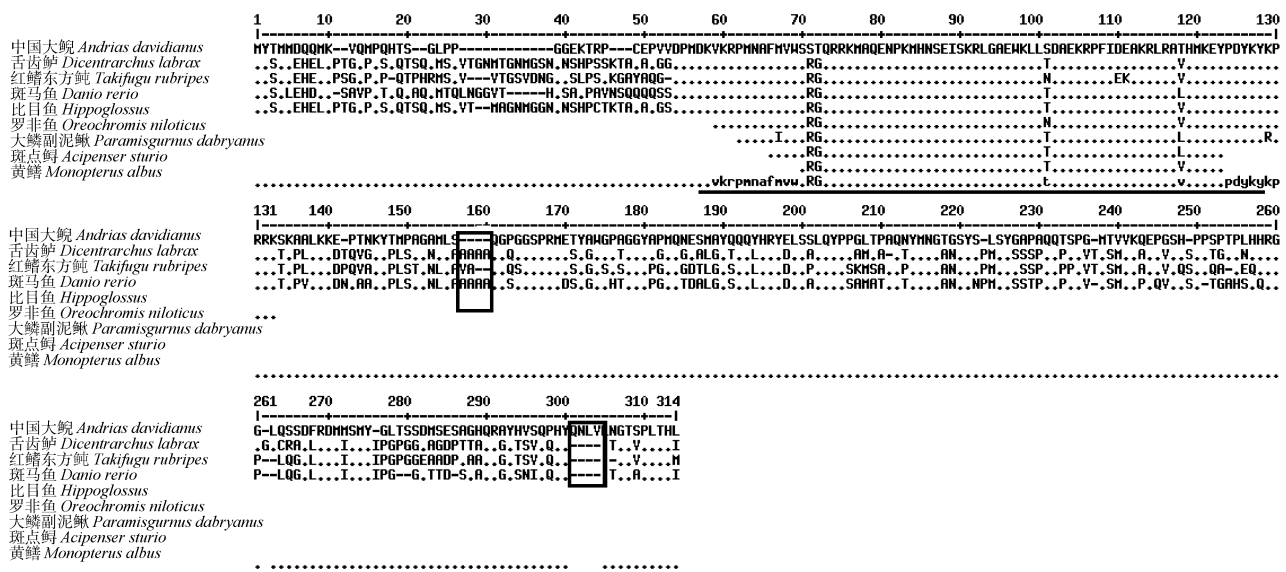


图 1 中国大鲵 *Sox19* 基因预测氨基酸序列与其他鱼类的比较

“.”代表相同氨基酸, HMG 盒用黑线标出, 鱼类比中国大鲵多出的“AAAA”以及中国大鲵比鱼类多出的“QNLV”用方框表示。
Fig. 1 Alignment of deduced amino acid sequences of *Andrias davidianus* *Sox19* gene with those of other fishes
“.”and the black line indicate identical amino acids and the HMG box, respectively. The boxes indicate the “AAAA” sequence only existing in fishes and the “QNLV” sequence only existing in *Andrias davidianus*, respectively.

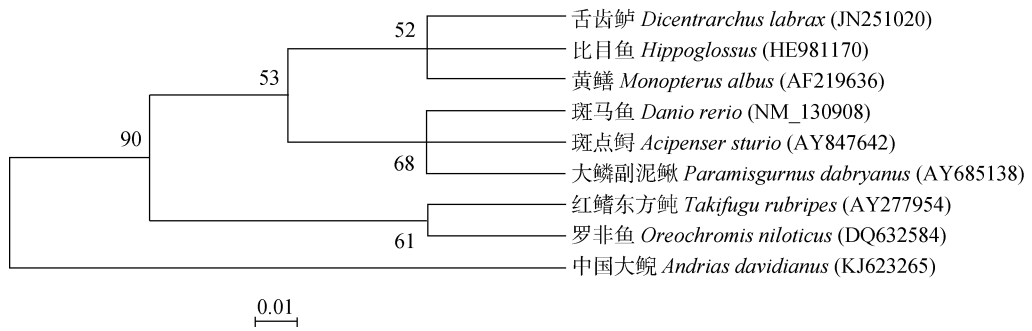


图 2 中国大鲵与其他物种 *Sox19* 基因的系统进化树

分支上的数字表示 1000 次重复抽样得到所得到的置信度百分比, 标尺长度表明每个位点发生 0.01 次置换; 物种名后为各序列 GenBank 登录号。

Fig. 2 Phylogenetic tree of *Sox19* gene from *Andrias davidianus* and other species
The values at the forks indicate the percentage of trees in which this grouping occurred after bootstrapping the data (1000 replicates). The scale bar shows the number of substitutions per site. The GenBank accession numbers of the sequences are behind the species names.

引物, 中国大鲵心脏、肾、卵巢和肠 cDNA 为模板, β -actin 为内参基因, 利用荧光定量 PCR 方法分析中国大鲵 *Sox19* 基因 mRNA 的组织表达特征。由图 3 可知, *Sox19* 基因在所有检测的组织中均有表达, 且在心脏中的表达量最高, 卵巢中次之, 肾脏中的表达量略高于肠道。

3 讨论

Sox19 基因在早期胚胎神经系统的发育过程

中发挥重要作用, 可以作为一个新型的工具去探测卵黄对于卵裂球分子组成的影响, 以及更加准确地追踪体内或体外移植中特定卵裂球的状态^[18]。然而到目前为止 *Sox19* 基因仅在鱼类当中发现, 其他种类动物的研究少之甚少, 两栖类动物研究更是空白。本研究首次在中国大鲵中克隆了 *Sox19* 基因, 经 PredictProtein 在线软件分析其编码的蛋白质的结构, 发现其唯一的结构域 HMG-box 处

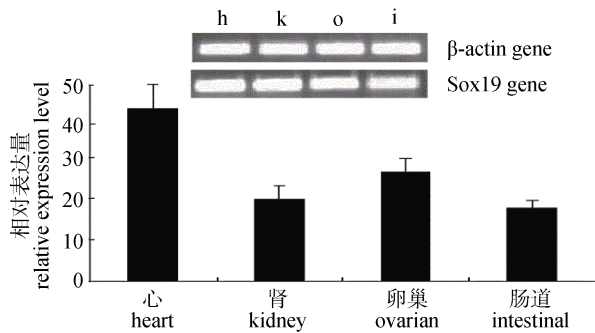


图 3 中国大鲵 *Sox19* 基因在心脏、肾、卵巢和肠的相对表达分析

Fig. 3 Relative expression analysis of heart, kidney, ovary and intestine of *Andrias davidianus* *Sox19* gene

于 N 端 39~109 位, α -螺旋也主要集中在 HMG-box 结构域, DNA 结合位点分别集中于 N 端 41~52 位以及 112~114 位, 说明 HMG-box 通过识别和结合 DNA 从而导致 DNA 构象发生变化, 也进一步验证了 HMG-box 结构域是其主要功能区。中国大鲵 *Sox19* 基因编码 285 个氨基酸, 比目前得到全长序列的舌齿鲈、斑马鱼及红鳍东方鲀 *Sox19* 编码的氨基酸数量要少(分别为 309 个、297 个、299 个)(图 1); 同时, 在 150 位氨基酸处, 鱼类比中国大鲵多出“AAAA”丙氨酸同类多聚体, 研究表明在脊椎动物中一些转录因子的丙氨酸、甘氨酸等氨基酸构成的同类多聚体具有转录调控功能, 这些同类多聚体是生物基因组进化的一个重要来源^[19]; 然而, 中国大鲵比鱼类在 C-末端多出“谷氨酰胺-天冬氨酸-亮氨酸-缬氨酸(QNLV)”序列, 这可能是中国大鲵 *Sox19* 基因在进化过程中丢失同源多聚体并替换为更加复杂的氨基酸编码序列, 但这种变化蕴藏的功能还有待进一步研究。中国大鲵 *Sox19* 蛋白与已知的鱼类进行同源性比较, 发现与红鳍东方鲀、斑马鱼和舌齿鲈相似度不是很高, 系统进化树也显示, 中国大鲵 *Sox19* 基因属于较早从鱼类分化出来的基因(图 2), 表明 *Sox19* 基因从鱼类到中国大鲵的进化速度比较快。

在鱼类中, *Sox19* 基因在多个组织中均表达, 例如, Vríz 等^[7-8]从斑马鱼中克隆发现 *Sox19* 基因, 并进一步证实其在间脑腹侧区、中脑和后脑都有表达; Santos 等^[20]以 *Sox19* 基因进行微阵列分析在斑马

鱼的精巢与卵巢中的表达; Navarro-Martín 等^[11]发现舌齿鲈的 *Sox19* 基因在性腺及皮肤中的表达量都比较高, 且随着卵巢的分化而表达上调。而中国大鲵与鱼类 *Sox19* 基因的组织表达情况并不完全相同。本研究采用荧光定量 PCR 法对 *Sox19* 基因在中国大鲵卵巢、肠、心脏及肾组织中的分布进行检测, 显示 *Sox19* 基因在所有检测的组织中均有表达, 其在心脏中的表达量最高, 卵巢中次之(图 3), 暗示大鲵 *Sox19* 基因对大鲵心脏的功能维持比性腺的作用更加重要, 而在肾和肠中的表达量相差不大, 在肾中的表达量略高于肠, 表明此基因对大鲵肾和肠功能的维持也起到一定的作用。这不仅进一步证实 *Sox19* 基因在鱼类卵巢中高表达的现象在中国大鲵中也是存在的, 而且表明了 *Sox19* 基因在不同物种中存在表达差异性。

目前有研究者认为 *Sox* 基因家族起源于一个共同的祖先基因, 在进化过程中通过基因的重复、散布突变以及新功能的获得等扩展其家族成员, 如 *SRY* 就是 *Sox* 基因家族进化过程中旁系同源的一个衍生基因^[21], 可能作为 *Sox3* 基因的变异形式出现并获得在睾丸发育过程中雄性特异的作用^[22]。Koopman 等^[10]、郭稳杰等^[14]认为 *Sox19* 基因很可能是由 *Sox3* 基因的一个复制子进化而来; Hett 等^[13]又进一步提出 *Sox19* 基因在斑马鱼、红鳍东方鲀和舌齿鲈等鱼类中存在直系同源基因, 而在哺乳动物中没有相应的直系同源基因, 且 *Sox19* 在斑点鲟的存在暗示它在较早的辐鳍鱼纲 (Actinopterygii) 就已存在。这表明了 *Sox19* 基因存在于鱼类与哺乳动物共同祖先中, 可能由 *Sox3* 基因复制进化产生, 即在进化过程中哺乳动物可能丢失了 *Sox19* 基因, 或在鱼类进化历程中逐渐由某一基因的复制子演化成 *Sox19* 基因。此外, 虽然序列差异较大, 但是 *Sox19* 与哺乳动物特异性基因 *Sox15*、非洲爪蟾 *SoxD* 是直系同源基因^[23-24]。有趣的是, 只存在于哺乳动物的 *Sox15* 基因一直都被认为是一个假基因^[25], 在哺乳动物进化过程中经历了快速的演变; Lee 等^[26]也发现缺 *Sox15* 基因的转基因小鼠除了在修复肌肉损伤方面有缺陷外, 完全可以健康生长。 *Sox19* 基因目前只在少数的低

等脊椎动物中存在, 除了中国大鲵, 仅在舌齿鲈、斑马鱼、红鳍东方鲀得到了 cDNA 全长, 它是否和同源的 *Sox15* 一样在哺乳动物快速演变中作为假基因存在, 还需要进一步的研究来证实。

Sox19 基因一直被认为是鱼类所特有的基因, 因此, 中国大鲵 *Sox19* 基因 cDNA 全长的成功克隆不仅弥补了两栖类中这一基因空缺, 证明 *Sox19* 基因不是鱼类所特有的基因, 哺乳动物可能在进化过程中丢失了该基因或者进化为另一个新的基因, 而且中国大鲵 *Sox19* 基因的成功克隆有助于进一步在 *Sox* 基因家族的进化方面做深入研究。此外, *Sox19* 转录因子在原肠胚早期的中枢神经系统的圆形区域内也能检测到其存在, 可能是脊椎动物中枢神经系统最早的分子标记^[9], *Sox19* 基因序列的克隆能为以后更好地研究大鲵早期胚胎神经系统发育情况打下基础。

参考文献:

- [1] Degani G. Expression of *SOX3* and *SOX9* genes in gonads of blue gourami[J]. *Advances in Biological Chemistry*, 2014, 4(5): 322–330.
- [2] Hong C S, Saint-Jeannet J P. Sox proteins and neural crest development[J]. *Semin Cell Dev Biol*, 2005, 16(6): 694–703.
- [3] Betancur P, Bronner-Fraser M, Sauka-Spengler T. Genomic code for *Sox10* activation reveals a key regulatory enhancer for cranial neural crest[J]. *Proc Natl Acad Sci USA*, 2010, 107(8): 3570–3575.
- [4] Cheng H H, Zhou R J. Alternative development of early embryogenesis: sex determination[J]. *Hereditas*, 2007, 29(2): 145–149. [程汉华, 周荣家. 早期胚胎的发育选择: 性别决定[J]. *遗传*, 2007, 29(2): 145–149.]
- [5] Li K L, Hu J Z, Yan B, et al. Isolation and expression analysis of *Sox10* gene relating body color variation in Oujiang color common carp[J]. *Journal of Fishery Sciences of China*, 2013, 20(6): 1139–1147. [李康乐, 胡建尊, 颜标, 等. 瓯江彩鲤体色相关基因 *Sox10* 的分离与表达分析[J]. *中国水产科学*, 2013, 20(6): 1139–1147.]
- [6] Bowles J, Schepers G, Koopman P. Phylogeny of the Sox family of developmental transcription factors based on sequence and structural indicators[J]. *Dev Biol*, 2000, 227(2): 239–255.
- [7] Vriz S, Lovell-Badge R. The zebrafish Zf-Sox 19 protein: a novel member of the Sox family which reveals highly conserved motifs outside of the DNA-binding domain[J]. *Gene*, 1995, 153(2): 275–276.
- [8] Vriz S, Joly C, Boulekbache H, et al. Zygotic expression of the zebrafish Sox-19, an HMG box-containing gene, suggests an involvement in central nervous system development[J]. *Mol Brain Res*, 1996, 40(2): 221–228.
- [9] Winata C L, Kondrychyn I, Kumar V, et al. Genome wide analysis reveals Zic3 interaction with distal regulatory elements of stage specific developmental genes in zebrafish[J]. *PLoS Genet*, 2013, 9(10): e1003852.
- [10] 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.
- [11] Navarro-Martín L, Galay-Burgos M, Piferrer F, et al. Characterisation and expression during sex differentiation of *Sox19* from the sea bass *Dicentrarchus labrax*[J]. *Comp Biochem Physiol B: Biochem Mol Biol*, 2012, 163(3–4): 316–323.
- [12] Liu L, Zhou R J. Isolation of *Sox11a*, *Sox11b* and *Sox19* genes from Rice field eel (*Monopterus albus*) using degenerate primers and nested PCR[J]. *Aquat Sci*, 2001, 63(2): 191–195.
- [13] Hett A K, Ludwig A. SRY-related (*Sox*) genes in the genome of European Atlantic sturgeon (*Acipenser sturio*)[J]. *Genome*, 2005, 48(2): 181–186.
- [14] Guo W J, Yu X M, Tong J G. Cloning and sequence evolution analysis of Sox genes in bighead carp (*Aristichthys nobilis*)[J]. *Acta Hydrobiologica Sinica*, 2014, 38(4): 664–668. [郭稳杰, 俞小牧, 童金苟. 鳙 *Sox* 基因克隆及序列进化分析[J]. *水生生物学报*, 2014, 38(4): 664–668.]
- [15] Hiraoka Y, Ogawa M, Sakai Y, et al. Isolation and expression of a human SRY-related cDNA, *hSOX20*[J]. *Biochim Biophys Acta—Gene Struct Expression*, 1998, 1396(2): 132–137.
- [16] Yin M G, Cao Y, Li C. Current status and protection countermeasures for Chinese giant salamander *Andrias davidianus*[J]. *Guizhou Agricultural Sciences*, 2014, 42(11): 197–202. [殷梦光, 曹宇, 李灿. 中国大鲵资源现状及保护对策[J]. *贵州农业科学*, 2014, 42(11): 197–202.]
- [17] Luo L T, Wan H L, Lan X P, et al. Research advances in resource and conservation genetics of Chinese giant salamander[J]. *Guangdong Agricultural Sciences*, 2011(17): 100–103. [雒林通, 万红玲, 兰小平, 等. 中国大鲵资源现状及保护遗传学研究进展[J]. *广东农业科学*, 2011(17): 100–103.]
- [18] Santacruz H, Vriz S. Regional distribution of maternal mRNA sox-19 in the zebrafish early embryo[J]. *Biol Cell*, 1996, 88(3): 153–155.
- [19] Wu J H. Expression pattern and function analysis of *Hoxc13* gene in skin of cashmere goat[D]. Hohhot: Inner Mongolia

- Agricultural University, 2011. [吴江鸿. Hoxc13 基因在绒山羊皮肤中的表达规律及体外功能分析[D]. 呼和浩特: 内蒙古农业大学, 2011.]
- [20] Santos E M, Workman V L, Paull G C, et al. Molecular basis of sex and reproductive status in breeding zebrafish[J]. *Physiol Genomics*, 2007, 30(2): 111–122.
- [21] Ohno S. The one-to-four rule and paralogues of sex-determining genes[J]. *Cell Mol Life Sci*, 1999, 55(6–7): 824–830.
- [22] Graves J A M. From brain determination to testis determination: evolution of the mammalian sex-determining gene[J]. *Reprod Fertil Dev*, 2001, 13(8): 665–672.
- [23] Okuda Y, Yoda H, Uchikawa M, et al. Comparative genomic and expression analysis of group B1 *sox* genes in zebrafish indicates their diversification during vertebrate evolution[J]. *Dev Dyn*, 2006, 235(3): 811–825.
- [24] Ito M. Function and molecular evolution of mammalian Sox15, a singleton in the SoxG group of transcription factors[J]. *Int J Biochem Cell Biol*, 2010, 42(3): 449–452.
- [25] Yamada K, Kanda H, Aihara T, et al. Mammalian *Sox15* gene: promoter analysis and implications for placental evolution[J]. *Zool Sci*, 2008, 25(3): 313–320.
- [26] Lee H J, Göring W, Ochs M, et al. Sox15 is required for skeletal muscle regeneration[J]. *Mol Cell Biol*, 2004, 24(19): 8428–8436.

Cloning and tissue expression analysis of the full length cDNA encoding the *Sox19* gene in Chinese giant salamander (*Andrias davidianus*)

LIU Jing, HE Qing, LIU Lili, LI Chenglei, WANG Qin

College of Life Sciences, Sichuan Agricultural University, Ya'an 625014, China

Abstract: The *Sox* family of transcription factors is involved in a variety of developmental processes, including sex determination and gonadal differentiation. The high-mobility-group (HMG) domains of all *Sox* proteins have highly conserved primary structures, and all appear to be capable of binding to the same target DNA sequence of AACAA(A/T)G. The *Sox* gene family is divided into 10 subgroups (designated A–J) based on sequence similarity and gene organization. *Sox19* is a member of the group B *Sox* genes, though it differs from other members in possessing an intron. The *Sox* B group is further divided into the B1 subgroup, which encodes transcriptional activators, and the B2 subgroup, which encodes repressors. Based on sequence and phylogenetic analyses, *Sox19* has been assigned to the B1 subgroup. Interestingly, the *Sox19* gene has always been considered fish specific but the full length cDNA encoding *Sox19* has only been cloned in *Danio rerio*, *Takifugu rubripes*, and *Dicentrarchus labrax*, and little is known about the expression patterns in other species. In this study, we successfully cloned the full-length *Sox19* cDNA sequence of the Chinese giant salamander (*Andrias davidianus*), which is the largest extant amphibian, and used quantitative real-time polymerase chain reaction (PCR) analysis to detect *Sox19* gene expression in different tissues. We cloned a 1290-bp *Sox19* gene cDNA, including 270 bp of the 5'-untranslated region (UTR) sequence, 858 bp of the open reading frame (ORF), and 162 bp of the 3'-UTR sequence. The ORF encoded 285 amino acids, of which a 71 amino acid protein was from the N-region and a 39–109 amino acid region was the highly conserved HMG-box. Bioinformatics software was applied to analyze the *A. davidianus Sox19* protein, which is hydrophilic. A subcellular localization analysis showed that 96% of the protein was present in the nucleus, possibly without the transmembrane region or signal peptide, 24.91% was in the α -helix, 3.16% was in the extended strand, 2.11% was in the β -corner, and 69.82% was in the random coil. The tertiary structural prediction revealed a α -helical structure as the functional site. The *A. davidianus Sox19* protein encoded 285 amino acids, which is less than that in fish. Therefore, fish have more similar “AAAA” polymers than those of the *A. davidianus* (150 amino acids), but *A. davidianus* has an additional “QNLV” at the C-terminal end, suggesting that the *Sox19* gene may have lost the lactamine oligomer during evolution and mutated to a more elaborate and complex transcriptional regulatory region. The amino acid sequence analysis revealed that the putative *A. davidianus-Sox19* protein had 55%, 58%, and 64% identity with *T. rubripes*, *D. rerio*, and *D. labrax*, respectively. The phylogenetic analysis further demonstrated that *A. davidianus* split off earlier based on the *Sox19* gene data, revealing a more primitive evolutionary status. The *Sox19* gene could have evolved through ancient fish-specific duplication of *Sox3*, although *Sox19* is expressed in the developing lens and central nervous system. We analyzed *Sox19* expression in the gonad, heart, kidney, and bowel of adult *A. davidianus* to confer the functions of the protein by fluorescence quantitative PCR. The results show the highest *Sox19* gene expression level in the heart, suggesting that *Sox19* helps maintain heart function. The high expression level in the gonads shows that *Sox19* may be involved in gonadal development and differentiation. *Sox19* was also expressed in the kidney and intestine, indicating that *Sox19* plays a role maintaining kidney and bowel functions in *A. davidianus*. The *Sox19* gene expression levels in *A. davidianus* were not the same as those in the fish species investigated. Our results provide insight into the role of the *Sox19* gene in *A. davidianus* and further our understanding of *Sox19* gene expression in amphibians. The *Sox19* gene could be the earliest molecular marker of embryogenesis in *A. davidianus*.

Key words: *Andrias davidianus*; *Sox19* gene; gene clone; gene tissue distribution

Corresponding author: WANG Qin. E-mail: wangqin3269@163.com