RESEARCH ARTICLE

Cloning and study of adult-tissue-specific expression of *Sox9* in *Cyprinus carpio*

DU QI-YAN, WANG FENG-YU, HUA HUI-YING and CHANG ZHONG-JIE*

College of Life Science, Henan Normal University, Xinxiang Henan, 453007, People's Republic of China

Abstract

The *Sox9* gene is one of the important transcription factors in the development of many tissues and organs, particularly in sex determination and chondrogenesis. We amplified the genomic DNA of *Cyprinus carpio* using degenerate primers, and found that there were two versions of *Sox9* in this species: *Sox9a* and *Sox9b*, that differ in having an intron of different length (704 bp and 616 bp, respectively) in the conserved HMG box region that codes for identical amino acid sequences. We used a two-phase rapid amplification of cDNA ends (RACE) for the isolation of full-length cDNA of *Sox9b*. Sequence analyses revealed a 2447-bp cDNA containing 233-bp 5′ untranslated region, a 927-bp 3′ untranslated region, including poly(A), and a 1287 bp open reading frame (ORF) encoding a protein of 428 amino acids. The HMG box of 79 amino acid motif was confirmed from positions 96–174. Sequence alignment showed that the identity of amino acids of *Sox9* among ten animal species, including *C. carpio*, is 75%, indicating that the *Sox9* gene is evolutionarily quite conserved. The expression level of *Sox9b* gene varied among several organs of adult *C. carpio*, with the level of expression being highest in the brain and testis.

[Qi-Yan D., Feng-Yu W., Hui-Ying H. and Zhong-Jie C. 2007 Cloning and study of adult-tissue-specific expression of *Sox9* in *Cyprinus carpio*. *J. Genet*. **86**, 85–91]

Introduction

There are more than 24,000 species of fish worldwide, inhabiting a wide range of aquatic habitats (Kirpichnikov 1981). A great deal is known regarding the process of sex differentiation in fish and the mechanisms involved in primary sex determination are now beginning to be defined (Devlin and Nagahama 2002). Sex determination in fish is a very plastic, primitive, labile process with respect to other vertebrates, with most of the karyotype species not showing sex chromosomes. Extensive studies on the structure and function of genes related to sex determination in fishes, followed by a comparison with results obtained from mammals, can therefore provide important insights into the evolutionary mechanism of sex determination in vertebrates (Baron *et al*. 2005).

In the twentieth century extensive studies on sex determination in mammals led to the belief that there is testis-determining factor (TDF) on the Y chromosome which plays a key role in sex determination (Kanai *et al*. 2005). Sinclair *et al*. (1990) identified the *SRY* gene (sex-determining region of Y) on the Y chromosome in humans, and it has generally been considered that the *SRY* gene is equated with TDF (Morrish and Sinclair 2002). However, the role of *SRY* in sex determination is known only in mammals, as corresponding genes have not been found in other vertebrates yet (Knower *et al*. 2003). Nevertheless, the discovery of *SRY* and *Sry* in human and mouse, respectively, rapidly led to the identification of a new gene family — the *Sox* (*SRY* related *HMG-box* gene) gene family — encoding a protein with an *SRY*-type conservative HMG-box. At present, it is believed that *Sox* genes are related to sexual development and play a diverse roles in sex determination and spermatogenesis (Bowles *et al*. 2000). Here, we report on studies of structure and tissue-specific expression patterns of the *Sox9* gene of *C. carpio*, a common and commercially valuable species of fish in China.

*For correspondence. E-mail: changzhongjie@tom.com.

Materials and methods

Materials

C. carpio were purchased from Xinxiang (Henan, China) market and were maintained at room temperature.

Keywords. gene expression; transcription factors; sex determination; *Sox9*; RACE; *Cyprinus carpio*.

The degenerate primers were designed on the basis of the HMG-box amino acid sequence of the *Sox9* genes registered in GenBank, as follows: forward 5′GG(C/T)(C/T)(G/T)(A/G)TA(C/T)TT(A/G)TA(A/G)T(C/T)(G/C)GG 3′, reverse 5′CC(A/C)ATGAA(C/T)GC(G/C)TT(C/T)AT(G/C)GT(G/C)TGG 3′.

Preparation of genomic DNA

DNA was extracted from fish blood samples by centrifuging 0.1–0.5 ml total blood, anticoagulated by ACD, at 3000 rpm for 10 min, followed by several rinses with 0.85% NaCl. The blood cells were diluted in suitable volume of lysis buffer (0.5 mol LEDTA, pH 8.0, 0.5% SDS). Proteinase K was added at the final concentration of 150 μg/ml and the solution was incubated overnight at 50°C. This was followed by three phenol-chloroform and chloroform extractions, respectively. After precipitation with ethanol, the pellet was rinsed in 70% ethanol, moderately dried, and dissolved in TE.

PCR amplification of the genomic DNA

The degenerate PCR was performed in a volume of 20 μl with about 200 ng genomic DNA, 1 μmol/l of each primer, 200 μmol/l dNTP, and 2U of *Taq* DNA polymerase. 35 cycles were performed with annealing temperature 52°C for 40 s and extension at 72°C for 1min. The amplification products were electrophorosed on 1.5% agarose gels.

Preparation of total RNA and RT-PCR

For the isolation of total RNA, approximately 20 mg of adult tissues were homogenized using a glass tissue homogenizer. Homogenization of these tissues was carried out in 1 ml of trizol reagent. The RNA was precipitated from the trizol reagent with isopropyl alcohol and was resuspended in nuclease-free water. RNA concentration and purity was determined by spectrophotometry and the total RNA was stored at −80°C.

To produce cDNA, reverse transcriptase (RT) reactions were performed with RevertAid^{TM} FIRST Strand cDNA Synthesis Kit (Fermentas, Lithuania, USA) and total RNA isolated from adult brain. Approximately 5μl of total RNA was used in each 20μl reaction along with oligo(dT) primers. The RT reactions were conducted according to manufacturer's protocol.

Cloning and sequencing of PCR products

PCR products were cloned into pGEM®-T vector and sequenced by the dideoxy-chain-termination method using Sp6 and T7 promoter sequences as primer. The nucleotide sequences were analysed by the BLASTx program on the NCBI Blast Server (http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi).

5′ and 3′ RACE analysis of Sox9

According to the HMG-box sequences of the *Sox9* gene that have been identified, the gene specific and nested primers were designed for the 5′ and 3′ RACE, respectively. The 5′ and 3′ RACE was performed using a SMART RACE cDNA Amplification Kit (Takara, Shiga, Japan), according to the manufacturer's instructions. PCR was performed with gene-specific primers and universal primers, with the primer sequences as follows:

5′ RACE Gsp1 5′ TCTTGTGCTGGACCCTCAGACGC 3′
Ngsp1 5′ CACAAATGGACGCTTCTCGCCCT 3′
3′ RACE Gsp2 5′ TTTATGGTCTGGGCTCAAGCGGC 3′
Ngsp2 5′ GTATCCACACCTGCACAACGCCG 3′

The PCR products were cloned and sequenced as described above.

Semiquantitative RT-PCR

We designed a pair of specific primers on the basis of the full-length cDNA sequence of the *Sox9* gene and analysed the expression in different tissues of adult *C. carpio* by using semi-quantitative RT-PCR method as described in Chiang *et al.* (2001). The total RNA samples were isolated directly from the brain, liver, heart, kidney, testis and ovary, and their concentration determined by spectrophotometric methods. The quantification of total RNA in each tissue was precisely 1μg in the reverse transcription reaction. The standard PCR reactions were performed at 94°C for 30 s, 57°C for 30 s, 72°C for 60 s, for 30 cycles with 2 μl cDNA template and the final extension reaction was carried out at 72°C for 10 min. The referential gene was GAPDH. *Sox9b* primers were: forward 5′ AGGAGGGCATTGACTTCG 3′, reverse 5′ TGTACATGGGCCTCTGGT 3′. The GAPDHP primers were: forward 5′ GCCTCCTGCACCACCAACTG 3′, reverse 5′ CGGAAGGCCATGCCGGTCAG 3′.

Homology search and dendrogram analysis

The amino acid sequences of *C. carpio Sox9b* were aligned with other full-length *Sox9b* gene sequences taken from GenBank, using ClustalW multiple alignment software (http://www.bioinformatics.nl/tools/clustalw.html). The aligned sequences were used to construct a phylogenetic tree by the neighbor-joining method, using the same software.

Results

Cloning and structural analysis of the Sox9 HMG-box in C. carpio

Using the degenerate primer set specific to the HMG-box motif, designed on the basis of published *Sox* gene sequences from different species, we observed three bands of lengths 210 bp, 850 bp and 900 bp, respectively, with genomic DNA of *C. carpio* as template (figure 1). The results were same between the male and female individuals. No sex-specific bands were observed.

The amplified fragments were purified and cloned, finally yielding two different fragments of 830 bp and 910 bp. Using PCGENE program (Intelligenetics, Mountainview, CA,

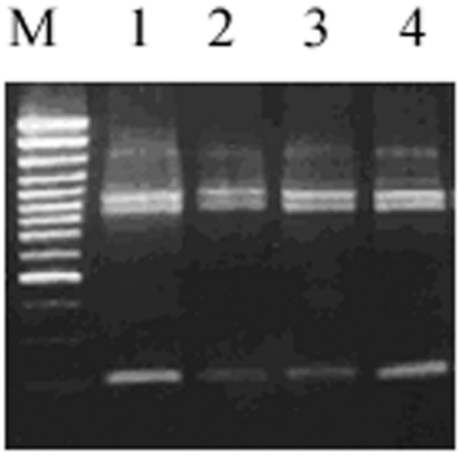

Figure 1. The result of TD-PCR on *C. carpio* genome. M, 100-bp DNA ladder; 1,2 males; 3,4 females.

USA), the two sequences were analysed and predicted to encode the same 68 amino acid protein, with the possibility that the two fragments differ in possessing introns of 704 bp and 616 bp, respectively. Homologous search with BLAST in GenBank showed that the sequences were similar to *Sox* genes in different animals, with the highest similarity being with the *Sox9* genes. Consequently, these two fragments were named *CcSox9a* and *CcSox9b*.

To verify the difference in the introns between the two fragments, two pairs of specific primers were designed accurately on the HMG-box of *CcSox9a* and *CcSox9b*, and used to amplify the total RNA from brain tissue by RT-PCR method (figure 2). The products were cloned and sequenced. The cDNA sequences of *CcSox9a* and *CcSox9b* were then compared with the sequences of the genomic DNA. The results showed that the prediction of the splicing sites of the introns were correct and also consistent with the "GT-AG" rule (figures 3,4).

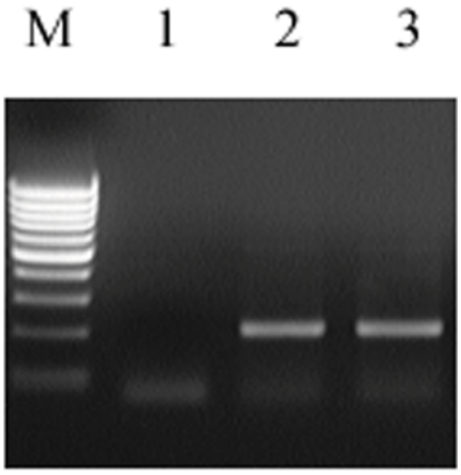

Figure 2. The results of RT-PCR using specific primers of *CcSox9a* and *CcSox9b*. M, 100-bp DNA ladder; 1, negative control; 2, *CcSox9a*; 3, *CcSox9b*.

Cloning and analysis of the full-length CcSox9b cDNA sequence

The 5′ and 3′ RACE results (figures 5,6) were joined to obtain the full length of *CcSox9b* cDNA sequence using the DNAMAN software (figure 7). The fragment of the 5′ RACE

```
CCAATGAATGCCTTTATGGTCTGGGCTCAAGCGGCGCGCAGGAAACTGGCGGACCAGTATCCACACCTGCACAACGCCG
 P  M  N  A  F  M  V  W  A  Q  A  A  R  R  K  L  A  D  Q  Y  P  H  L  H  N  A
AGCTCAGCAAGACCCTCGGCAAACTCTGGAGgtcagagcattcattgtttatgaagtgtaggacaactccagaagccgg
E  L  S  K  T  L  G  K  L  W  R
agcaacactgattcatttaactgcccagacaaactcactgtattattaatattaaactgcattgtttcatagcattatt
atataatgtcaggcaatactgataagctgtttgggaaataaagggtaaaactatagcttaggcaaatggtgaattaacc
cttacttcagtgcagaaaaagttgtgtagtcactaaaaactttggtaacactttggaaaagggaacacttactcactat
taactatgacttttccctctataaattcctaatttgctgcttattaatagttagtatggnagctttaagtttaggtat
gaggtaggattagggatgtagaataagggcatggaaaaaaaagacattaatatgtgcttaactactactaataaatggc
taatattctagtaatatgcatgctaataagaaactagttaagagaccctaaaataaagtgttacccaaacttttagttt
taagaaatttagttcgatctggattactttattctttttttttcctgaataatatttttttcttaagtgaacccaaaaagt
aataatcatagtttccaagaacaatttatacctttttgtaggtttctttattgttaactaaggtagatcttagtattg
attatttctcatgtgcttttgtagGTTACTGAATGAGGGCGAGAAGCGTCCATTTGTGGAGGAGGCCGAGCGTCTGAGG
                        L  L  N  E  G  E  K  R  P  F  V  E  E  A  E  R  L  R
GTCCAGCACAAGAAAGACCACCCCAACTACAAGTACCAGCC
  V  Q  H  K  K  D  H  P  N  Y  K  Y  Q
```

Figure 3. The DNA sequence and the predicted amino acid sequence of *CcSox9a*. Underlined sequences were used for primers of PCR reactions to confirm the splicing sites. Lower case letters indicate introns.

```
CCAATGAACGCGTTTATGGTGTGGGCTCAAGCCGCGCGCAGGAAACTGGCGGATCAGTATCCGCACCTGCACAACGCCG
  P  M  N  A  F  M  V  W  A  Q  A  A  R  R  K  L  A  D  Q  Y  P  H  L  H  N  A
AGCTCAGCAAGACCCTCGGAAAACTCTGGAGgtgagagagagagacaattcattttactttgcttcgtttttgtttgtt
E  L  S  K  T  L  G  K  L  W  R
caggactgctccagaagtcaggaaactttgtgtagtactattaaattgcaatgatttatagcattattatgcaatggtt
ataaaacaatattgataagcttgagaaaatgtaactatttgggtagtaaggtaaaactttagtttagtcaaatagtaaa
ttaaccctaacttcacatgcacttcaacacagaaaaagttgtggagtcggtaaaaacctccagttttaagaaaagatgc
tgattactttattacttttgggtttatgtaagaaaagtagtattcacttaaaaaaaaaaaataattataatagctgnt
tcaagaacaatctatacctttttttaaagnttatttattattaattggacaagataaagaaagtngattacttttctg
aattactttattgctttgcgggtttacttattaagaaaaaaatagtataaaaaaaaaaaaaagaatagtcatatttcaaa
aacaatttgtaccttaatgaaatgtctgttttgttaattgaagaatgcaaaggaagtaatctgagcattggttaattct
tgtggatatgtgtagGTTACTGAATGAGGGCGAGAAGCGTCCGTTCGTGGAGGAGGCCGAGCGTCTGAGGGTGCAGCAC
                  L  L  N  E  G  E  K  R  P  F  V  E  E  A  E  R  L  R  V  Q  H
AAGAAAGACCACCCCGACTACAAGTACAGACC
  K  K  D  H  P  D  Y  K  Y  R
```

Figure 4. The DNA sequence and the predicted amino acid sequence of *CcSox9b*. Underlined sequences were used for primers of PCR reactions to confirm the splicing sites. Lower case letters indicate introns.

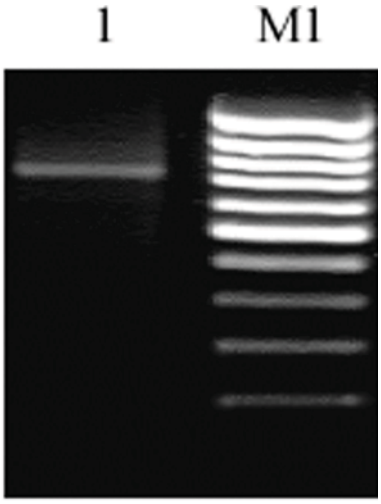

Figure 5. The result of 5′ RACE on *C. carpio Sox9b*. M1, 100-bp DNA ladder; 1, the result of 5′ RACE.

includes the cDNA sequence from nucleotide positions −233 to +443, whereas the fragment of the 3′ RACE extends from nucleotide position +356 to the poly(A). Sequence analyses indicate a 2447-bp cDNA sequence containing the 233-bp 5′ untranslated region, 927-bp 3′ untranslated region, including poly(A), and a 1287-bp ORF that encodes a protein of 428 amino acids. The HMG box of 79 amino acid motif was confirmed between nucleotide positions 96–174 (figure 8).

Comparison between CcSox9b and Sox9 in other species

The deduced amino acid sequence of CcSox9b protein was compared with the GenBank sequences of different Sox9 proteins of nine other species: *Gasterosteus aculeatus, Takiugu rubripes, Danio rerio, Xenopus laevis, Alligator mississippiensis, Canis familliaris, Mus musculus, Macaca mulatta* and *Homo sapiens*. The similarity between CcSox9b protein sequence and the sequences from these nine species was 72%, 69%, 69%, 71%, 68%, 65%, 66%, 66% and 66%, respectively (figure 9). The high sequence similarity seen in this comparison indicates that the *Sox9* gene is quite highly conserved across species separated by vast lengths of evolutionary time, suggesting that perhaps the Sox9 protein plays some fundamental role in vertebrate development. Moreover, the HMG-box motif is even more highly conserved, with only two amino acids differing among these species. Another highly conserved region is in the C-terminal, which is the *trans*-activation domain of *Sox9*.

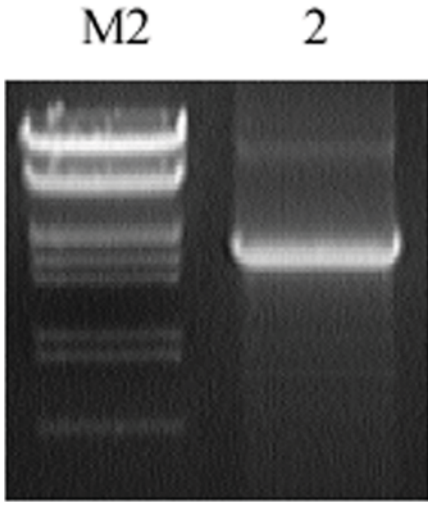

Figure 6. The second result of 3′ RACE on *C. carpio Sox9b*. M2, lambda DNA/*EcoR*□ + *Hind*□; 2, the result of 3′ RACE.

ACGCGGGGGC AGAACCTCTC CACGTCTCCA CACACATACC TGCACTCAGT CAGGATTTGT CCTCGTCTCA 70
GAGATACATG TCGTGCTGTG TGGTGCGAGT TTATTTTCCC GATCTGAGCA GTTCTGGGGA CTTTGGACAC 140
TCGGAGCCCC TTTGCGCACG GCCGCGCGCG CGTTGACGCA CTCCTCTGAG GACTTTACAG CTGCGCTAGA 210
GTTTCCAGCA CTCACCCTTC TGCATGAATC TCTTCGACTC CTACCTGAAA ATGAGCGACG AGCAGGACAA 280
GGGTCTCTCC GACGCGCCCA GCCCGAGCAT GTCCGAGGAT TCCGCCGGGT CTCCGTGTCC GTCCGGATCG 350
GGCTCGGACA CCGAGAACAC CCGCCCGGAG GAGCACCTGG GAGAGTTTAA GAAGGACGAG GACAAGTTCC 420
CCGTGTGTAT CAGGGACGCG GTGTCTCAGG TGCTGAAGGG CTACGACTGG ACCCTGGTGC CCATGCCTGT 490
GCGAGTGAAC GGCGCGCACA AGAGCAAGCC GCACGTCAAG AGACCCATGA ACGCGTTTAT GGTTTGGGCT 560
CAAGCGGCGC GCAGGAAACT GGCGGACCAG TATCCACACC TGCACAACGC CGAGCTCAGC AAGACCCTCG 630
GAAAACTCTG GAGGTTACTG AACGAGGGCG AGAAGCGTCC GTTCGTGGAG GAGGCCGAGC GTCTGAGGGT 700
GCAGCACAAG AAAGACCACC CCGACTACAA GTACCGGCCC AGACGGAGAA AATCAGTGAA GAACGGCCAG 770
AGCGAGAGCG AGGACGGCGA GCAGACCCAC ATCTCACCCA ACGCCATCTT CAAAGCCCTG CAGCAGGCCG 840
ACTCCCCCGC GTCCAGCATG GGCGAAGTGC ACTCTCCGGG AGACCACTCA GGTCAGTCCC AGGGTCCCCC 910
GACTCCTCCC ACGACCCCCA AAACGGACCT GCCGTCCAGC AAAGCGGATC TGAAGCGCGA GGGCCGTCCG 980
CTGCAGGAGG GCATTGACTT CGGCGCCGTA GACATCGGCG AGCTGAGCAG CGACGTCATC TCCAACATGG 1050
AGCCATTCGA CGTCAACGAG TTCGACCAGT ACCTGCCCCC TCGCGGACAC CCGGGGGTCA GCGGCGGCAT 1120
GCAGGCGTAT CCCGCCGGGT ACGGTGGCAC CTGGATGTCC AAACAGCACT CGATGGCCAG CGGTGGCGAG 1190
CAGAGCCAAG GCCAGCAGCG GACGCAGATC AAGACGGAGC GGCTGAGCCC CAGCCACTAC AGCGAGCAGC 1260
AGCGGCAGGG CTCTCCGCAG CATGTGGCCT ACGGCTGCTT CAACCTGCAG CACTACAGCA GCAGCGGCGG 1330
CAGCAGTGGC AGCGGCTCGT ACTACAACTA CACCGAGCAC CAGAGCCCCG CCGGCTCCTA CTACAGCCAG 1400
TACCCCGCCT TCAGCTACCA GAGGCCCATG TACACCCCGA TCGCCGACAC GCTGCCGCAG GCGCACAGTC 1470
CGCAGCACTG GGAGCAGCAG CCCGTCTACA CCCAGCTGTC CAGGCCCTGA GAGGCCATCC CGGCTGGAGA 1540
CGAACTCAAA CTGAGCGAAT GGCTCCCGAC TGTGCCTTTG GAGTTGTGAT TATATTTTTC TAGATATAAT 1610
GAGACAAGAG AATCCTCTGT GAGGACAGAT TCCAGATACT TATTTTTAGT ATGTACCGTG TATGTGTCAC 1680
GTTTCTCAGT TGTGTCGGGG GATTCGTACG CAGCTGTCTG ATGATATTTT TTTGTAAAGC CTGATATTTT 1750
CCTCTGTATA TATTTGTGGG GGTTTCTTAC AACTCGGACG AAGTGTTGAA TTGTCAGATG TACAGTCGAC 1820
TGCTTCTAGA GTCCGTGCCT ATCTCTCGTT TCTGTGAAAA CTGCCTCTAG AGTCCGTCTG TGCCTATAGA 1890
GATCTCGCTG CTGGAGGGGG GACGGGAAGT GCCCTTTGAC CCGAGGCTCC TCCGCAGCTC GCCATTGGTC 1960
AGATCGTCAT TCCTCCAGTC TTAACCTCTT TAATTTATTC ATTAGCGTTC ATTTGACTTG AAAGTAATAC 2030
TATTTGCGTT TTATCATGTA CTCGGGCTTC GTACGACTAA TTTGTGATGA AACGTTTCCG GAGACGATTA 2100
CGATTTAGAT GTAGTGCAAA CTTTAACCAT CTCATTTATT TATCGTTTCT TAATGTTTCC ATTTAAAATG 2170
TCTAGAAATG GTACTCTAGT TGACTTCCGA TTGTGAAACT TGATCAAGCC TTTCATAGAA ACAACTGGAA 2240
CTTGTTTTTT TTGTTGTTTT TTTTGTTCTC AAACATGTTG ACTTCTTCAG GCTGTTGGCT ACATCTCTAC 2310
AAAAACACAC CAGCATTTAT TTCTTTTCCC TTTGTTTTTT TTAAATGTCA TAATGAAAAA CTGTAAAACA 2380
TGCCACCATC TTAAAAGAAG CTTATCTTTT GTATCATATT GTTTGCAAAA AAAAAAAAAA AAAAAAA 2447

Figure 7. The cDNA sequences of the *C. carpio Sox9b*. --shows 5′-untranslated region; ww shows 3′ untranslated region; ATG shows initiation codon; TGA shows the termination codon.

MNLFDSYLKM SDEQDKGLSD APSPSMSEDS AGSPCPSGSG SDTENTRPEE HLGEFKKDED KFPVCIRDAV
SQVLKGYDWT LVPMPVRVNG AHKSKPHVKR PMNAFMVWAQ AARRKLADQY PHLHNAELSK TLGKLWRLLN
EGEKRPFVEE AERLRVQHKK DHPDYKYRPR RRKSVKNGQS ESEDGEQTHI SPNAIFKALQ QADSPASSMG
EVHSPGDHSG QSQGPPTPPT TPKTDLPSSK ADLKREGRPL QEGIDFGAVD IGELSSDVIS NMEPFDVNEF
DQYLPPRGHP GVSGGMQAYP AGYGGTWMSK QHSMASGGEQ SQGQQRTQIK TERLSPSHYS EQQRQGSPQH
VAYGCFNLQH YSSSGGSSGS GSYYNYTEHQ SPAGSYYSQY PAFSYQRPMY TPIADTLPQA HSPQHWEQQP
VYTQLSRP

Figure 8. The deduced amino acid sequence of the *C. carpio Sox9b*. The sequences underlined are the HMG-box motif.

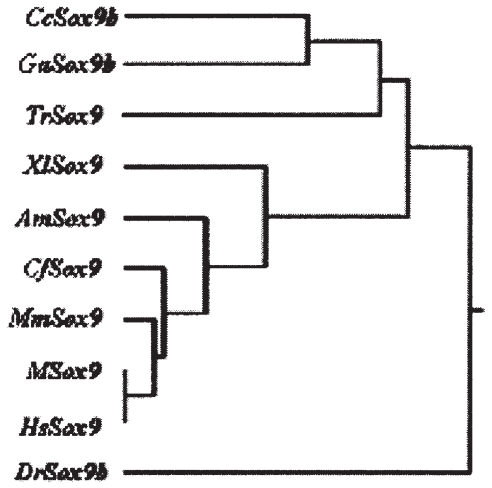

Figure 9. Dendrogram of *Sox9* proteins from various species, based on homology. *Cc*, *Cyprinus carpio*; *Ga*, *Gasterosteus aculeatus* (AAQ62979); Tr, *Takiugu rubripes* (AAQ18508); Dr, *Danio rerio* (AAH67133); Xl, *Xenopus laevis* (AAK61366); Am, *Alligator mississippiensis* (AAD17974); Cf, *Canis familliaris* (AAP69840); Mm, *Mus musculus* (AAH24958); M, *Macaca mulatta* (AAK01651); Hs, *Homo sapiens* (AAP36231).

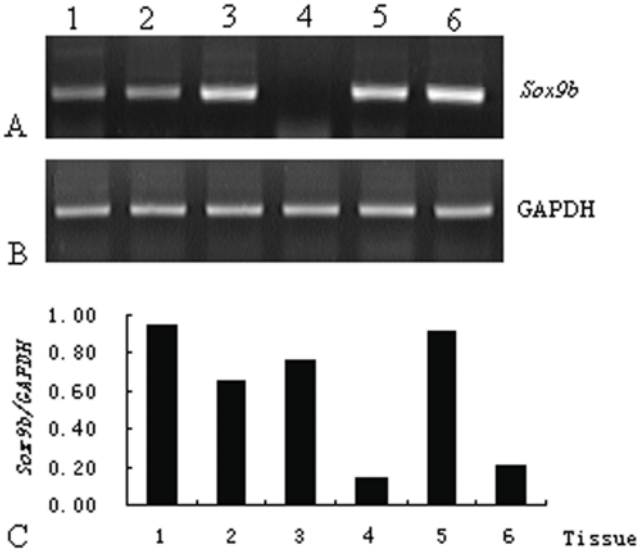

Figure 10. Expression profile of *Sox9b* in several organs of adult *C. carpio*, (A) The result of the *CcSox9b* expression, (B) the result of GAPDH expression, and (C) the *Sox9b*/GAPDH ratio. 1, kidney; 2, liver; 3, testis; 4, ovary; 5, heart; 6, brain.

Expression of CcSox9b in adult tissues

The adult-tissue-specific expression patterns of *CcSox9b* were examined by semiquantitative RT-PCR with the specific primers outside the HMG-box motif. The referential gene is GAPDH. The results indicate that *CcSox9b* is widely expressed in various adult tissues, such as brain, liver, heart and testis, especially in brain and testis, but the expression is quite low in kidney and ovary (figure 10).

Discussion

We have found that the *C. carpio* genome (tetraploid, $2n$ = 100) contains two duplicate orthologs of the *Sox9* gene, termed *Sox9a* and *Sox9b*, respectively. Some other recent studies on fish have suggested the existence of two versions of *Sox9*. Chiang *et al*. (2001) cloned two *Sox9* genes from zebrafish, and gene phylogenies showed that both genes were orthologous to tetrapod *Sox9* genes. The expression patterns of these two genes are distinct, but overlap in some regions during embryogenesis. Further genetic mapping showed that these two loci reside on chromosome segments which were apparently duplicated in a large-scale genomic duplication event in the ray-fin fish lineage (Chiang *et al*. 2001). It appears that duplication of genomic segments during the evolution of teleost fishes and mutations of duplicated *Sox9* genes, has led to the appearance of different copies of this gene (Cresko *et al*. 2003). Zhou *et al*. (2003) even reported that the *Sox9a* gene was duplicated during evolution of the rice-field eel, *Monopterus albus*. The duplicated copies of the gene (named *Sox9a1* and *Sox9a2*) have same expression patterns (Zhou *et al*. 2003; Lu *et al*. 2003).

Sox9 is a transcription factor required for sex determination and differentiation in vertebrates (Yokoi *et al*. 2002). Mutations in human *SOX9* cause campomelic dysplasia, a dominant skeletal dysmorphology syndrome often associated with male to female sex reversal (Foster *et al*. 1994; Wagner *et al*. 1994). In fishes, it has been suggested that the initial events in sex determination occur in the brain, primarily in the preoptic-anterior-hypothalamic areas. These initial events lead to the local production or conversion of sex steroids in brain cells, and different ratios of sex steroids play an important role in the control of gonadal development (Devlin and Nagahama 2002). *CcSox9b* overexpresses in the adult brain and testis, but is weakly expressed in the ovary. This pattern indicates that *CcSox9* is an upstream gene in the male sexual differentiation pathway, and possibly may play a role in the initial events in *C. carpio* sex differentiation. Expression of *CcSox9* in adult liver, heart and kidney was detected in both sexes, suggesting that this gene may have a wider role in the development of the genitourinary system in fishes.

Zhan and Song (1980) regarded the sex chromosomes of *C. carpio* as being XX/XY, based on the analyses of karyotypes and banding patterns. However, most studies have not shown sex chromosomes in these species (Zhan and Song 1980). We found that there is a pair of chromosomes in the B-group that have obvious differences in size (Chang Zhong-Jie et al. unpublished data), but whether or not these sex chromosomes remains to be seen. In *C. carpio*, as in many other fishes, it is known that sex reversal can occur if a certain concentration of sex steroids were added into the water, or as a result of other environmental changes (Devlin and Nagahama 2002). In consequence, it seems to explain that in fishes the

genes related to sex-determination function is so limited that the genes in sexual differentiation pathway cannot entirely and independently control the initiation of the direction of sex development.

Members of the Sox family of transcription factors are found throughout the animal kingdom, and are involved in a diverse range of developmental processes (Chang *et al*. 2002), many of them related to sex determination and differentiation (Chaboissier *et al*. 2004). *Sox9* expresses shortly after *Sry* in the presertoli cells and is a strong candidate for activation by Sry protein because it is upregulated in Sertoli cells just after *Sry* expression. *Sox9* is necessary for the activation of the gene encoding antimullerian hormone (AMH), also known as Mullerian inhibiting substance (Sekido *et al*. 2004). AMH is a transforming growth factor β-like glycoprotein hormone, which is secreted by Sertoli cells and which causes regression of the female mullerian ducts (Lasala *et al*. 2004). Other factors, including Sf1 and Wt1, are also involved in the activation of the *Amh* gene. Thus, *Sox9* can act as a classic transcriptional activator; however, its role in sex determination must be more extensive than just activation of AMH, because AMH-null mice are not sex reversed (Rodriguez-Mari *et al*. 2005). Although several of the key processes and genes involved in male sexual determination have been identified, sex determination in vertebrates involves an intricately regulated network of gene expression and it will be a challenging task to characterize the role of all of these genes in the network.

References

Baron D., Houlgatte R., Fostier A. and Guiguen Y. 2005 Large-scale temporal gene expression profiling during gonadal differentiation and early gametogenesis in rainbow trout. *Biol. Reprod*. **73**, 959–966.

Bowles J., Schepers G. and Koopman P. 2000 Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev. Biol*. **227**, 239–255.

Chaboissier M. C., Kobayashi A., Vidal V. I., Lutzkendorf S., van de Kant H. J., Wegner M. *et al*. 2004 Functional analysis of *Sox8* and *Sox9* during sex determination in the mouse. *Development* **131**, 1891–1901.

Chang Z., Du Q. and Shao H. 2002 New advances in *Sox* gene family. *Hereditas (Beijing)* **24**, 470–476.

Chiang E. F., Pai C. I., Wyatt M., Yan Y. L., Postlethwait J. and Chung B. 2001 Two *Sox9* genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites. *Dev. Biol*. **231**, 149–63.

Cresko W. A., Yan Y. L., Baltrus D. A., Amores A., Singer A., Rodriguez-Mari A. and Postlethwait J. H. 2003 Genome duplication, sub-function partitioning, and lineage divergence: *Sox9* in stickleback and zebrafish. *Dev. Dynam*. **228**, 480–489.

Devlin R. H. and Nagahama Y. 2002 Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* **208**, 191–364.

Foster J. W., Dominguez-Steglich M. A., Guioli S., Kowk G., Weller P. A., Stevanovic M. *et al*. 1994 Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* **372**, 525–530.

Kanai Y., Hiramatsu R., Matoba S. and Kidokoro T. 2005 From *SRY* to *SOX9*: mammalian testis differentiation. *J. Biochem. (Tokyo)* **138**, 13–19.

Kirpichnikov V. S. 1981 *Genetic bases of fish selection*. pp. 40–42 Springer, Berlin.

Knower K. C., Kelly S. and Harley V. R. 2003 Turning on the male: *SRY*, *SOX9* and sex determination in mammals. *Cytogenet. Genome Res*. **101**, 185–198.

Koopman P., Gubbay J., Vivian N., Goodfellow P. and Lovell-Badge R. 1991 Male development of chromosomally female mice transgenic for *SRY. Nature* **35l**, 117–121.

Lasala C., Carre-Eusebe D., Picard J. Y. and Rey R. 2004 Subcellular and molecular mechanisms regulating anti-Mullerian hormone gene expression in mammalian and non-mammalian species. *DNA Cell Biol*. **23**, 572–585.

Lu H., Cheng H., Guo Y. and Zhou R. 2003 Two alleles of the *Sox9a2* in the rice field eel. *J. Exp. Zool*. B: *Mol. Dev. Evol*. **299**, 36–40.

Morrish B. C. and Sinclair A. H. 2002 Vertebrate sex determination: many means to an end. *Reproduction* **124**, 447–457.

Rodriguez-Mari A., Yan Y. L., Bremiller R. A., Wilson C., Canestro C. and Postlethwait J. H. 2005 Characterization and expression pattern of zebrafish anti-Mullerian hormone (*Amh*) relative to *Sox9a, Sox9b*, and *cyp19a1a*, during gonad development. *Gene Expr. Patterns* **5**, 655–667.

Sekido R., Bar I., Narvaez V., Penny G. and Lovell-Badge R. 2004 *SOX9* is up-regulated by the transient expression of *SRY* specifically in Sertoli cell precursors. *Dev. Biol*. **274**, 271–279.

Sinclair A. H., Berta P., Palmer M. S., Hawkins J. R., Griffths B. L., Smith M. J. *et al*. 1990 A gene from the human sex determining region encodes a protein with homology to a conserved DNA binding motif. *Nature* **346**, 240–244.

Wagner T., Wirth J., Meyer J., Zabel B., Held M., Zimmer J. *et al*. 1994 Autosomal sex reversal and *campomelic dysplasia* are caused by mutations in and around the SRY-related gene *SOX9. Cell* **79**, 1111–1120.

Wilson M. J., Jeyasuria P., Parker K. L. and Koopman P. 2005 The transcription factors steroidogenic factor-1 and *Sox9* regulate expression of vanin-1 during mouse testis development. *J. Biol. Chem*. **280**, 5917–5923.

Yokoi H., Kobayashi T., Tanaka M., Nagahama Y., Wakamatsu Y., Takeda H. *et al*. 2002 *Sox9* in a teleost fish, medaka (*Oryzias latipes*): evidence for diversified function of *Sox9* in gonad differentiation. *Mol. Repr. Dev*. **63**, 5–16.

Zhan R. and Song Z. 1980 Analysis and comparison between the karyotypes of *Cyprinus carpio* and *Carassius auratus*, as well as *Aristichthya nobilis* and *Hypophthatmichthys molitrix. Acta Genet. Sinica* **7**, 72–77.

Zhou R., Liu L., Guo Y., Yu H., Cheng H., Huang X. *et al*. 2003 Similar gene structure of two *Sox9a* genes and their expression patterns during gonadal differentiation in a teleost fish, rice field eel (*Monopterus albus*). *Mol. Repr. Dev*. **66**, 211–217.

Received 19 April 2006