Analysis of intron sequence variability of the conservative HMG-box of *Sox9* genes in allotetraploids and their original parents*

Liu Jifang, Liu Shaojun**, Tao Min, Li Wei and Liu Yun

(Key Laboratory of Protein Chemistry and Developmental Biology of State Education Ministry of China, College of Life Sciences, Hunan Normal University, Changsha 410081, China)

Accepted on August 22, 2006

Abstract The *Sox* genes of allotetraploids and their original maternal red crucian carp (*Carassius caassius* red var.) and original paternal common carp (*Cyprinus carpio* L.) were detected by PCR with the designed primers based on the conserved HMG-box sequence in different species. Sequencing of *Sox* genes indicated that two *Sox9* genes (*Atsox9a* and *Atsox9b*) existed in allotetraploids, while only one *Sox9* gene existed in red crucian carp (*Rcsox9a*) and common carp (*Ccsox9b*). All of the four *Sox9* genes contained an intron in the HMG-box, with the sizes of 413 bp, 703 bp, 401 bp and 714 bp, respectively. Moreover, the introns obeyed the rule of "GT-AG". A high similarity was observed between introns of *Atsox9a* and *Rcsox9a* (94.4%), *Atsox9b* and *Ccsox9b* (97.8%). Interestingly, the deduced amino acid sequences of their corresponding exons all shared 100% identity. Thus, introns of the HMG-domain of *Sox9s* in allotetraploids and their original parents have not only the length polymorphism but also intron variability. Our results provide significant molecular evidence for the origin and evolution of allotetraploids.

Keywords: *Sox9*, intron, genetic variability, phylogenesis.

The *Sox* gene family includes a majority of genes characterized by a HMG-box. It belongs to a super family of transcription factors, and plays important roles in various early embryo development such as sex determination, skeletogenesis, the forming of blood cells and nerve system development and so on[1,2]. Especially, the mutation of the *Sox9* gene will lead to compomedic dysplasia and sex inversion, the patients with XY chromosomes are mostly sex-reversed women[3].

Intron is one of the momentous and puzzling elements in genomes. It was also found that some introns probably contained genes encoding proteins related to their activities after the discovery of non-encoding introns in genomes[4—6]. Some studies have indicated that the splicing of introns is an important step in the gene expression of eukaryotic organism. Especially, the variable splicing can regulate spatio-temporal gene expression. For instance, more than 30% of the genes in human genome are formed by alternative splicing, which makes a gene translated into several function-related proteins, and increases the genetic complexity. Under extreme conditions, different proteins have completely different or even converse functions[7]. Two kinds of splicings of introns of HMG-box in mouse *Sox17* have been reported[8]. One produced *Sox17* mRNA, expressed abundantly in spermatogonia and decreased markedly in pachytene spermatocyte stage, the other produced *t-Sox17* mRNA which was expressed abundantly. The different splicing of *Sox17* is vital to the formation of spermatozoa. Additionally, introns have also great influences on gene expression and regulation[9,10]. Compared with the sequences of exons, introns whether in different species or in the same one have more distinctly genetic polymorphism due to the less evolutionary pressure than that of corresponding exons. Wang et al. once reported that introns could be used in evolutionary analysis of genetically close species like exons[11]. Therefore, the analysis of intron variability may provide abundant information about origin and evolution form different fishes. In this study, we cloned partial sequences of the con-

* Supported by National Natural Science Foundation of China (Grant Nos. 30330480 and 30571444), the Doctoral Station of University of the Education Ministry of China (Grant No. 200405422001), Changjiang Scholar and Innovative Team Foundation of Ministry of Education (Grant No. IRT0445)

** To whom correspondence should be addressed. E-mail: lsj@hunnu.edu.cn

served HMG-box of *Sox9s* from allotetraploids and their original parents, and analyzed the sequences of introns. The results obtained provided a line of evidence for the origin and evolution of allotetraploids and implied some functions of introns in *Sox9*.

1 Materials and methods

1.1 Genomic DNA extraction and PCR amplification

Allotetraploids and their original parents (red crucian carp and common carp) came from National Tetraploid Fish Protection Station located in Hunan Normal University. The genomic DNA extraction and PCR amplification of the HMG-box of *Sox* genes were carried out as previously described[12].

1.2 Reverse transcription and polymerase chain reaction (RT-PCR)

The cDNA fragments were generated from 1 μg of total RNA extracted from testis of allotetraploids, red crucian carp and common carp using AMV reverse transcriptase (Promega) in a 25 μL reaction according to the manufacturer's manual. The resultant cDNA (1 μL) was used as a template for subsequent PCR. To amplify cDNA fragments of *Atsox9b*, *Rcsox9a* and *Ccsox9b*, the specific primer sets were used separately. The PCR was performed in a 25 μL volume containing 0.3 μmol of each primer, 200 μmol/L of each dNTPs, and 1U of *Taq*-polymerase (Takara). The cycling conditions were: 5 min at 94℃, 30 cycles of 30s at 94℃, 30s at 54℃, and 1 min at 72℃ ending with 7 min of extension at 72℃. The PCR products were analyzed in 2.0 % agarose gels stained by ethidium bromide.

1.3 Cloning and sequencing of PCR products

The amplicons were separated in 1.2% agarose gels, purified by Gel Extraction Kit (Sangon), ligated into the pMD18-T vector, transferred to *E. coli* DH5α. The positive clones were identified by blue and white spots screening and PCR amplification.

1.4 Construction of a molecular phylogenetic tree

The sequences of introns in the HMG-box of *Sox9* genes in different species were aligned using the Clustal W (1.83) software, and corrected artificially. The phylogenetic tree was constructed with neighbor-joining (NJ) method using MEGA (3.0) software package[13] and the statistical reliability was tested using bootstrap. During NJ analysis, Kimura 2-parameter was chosen as the genetic distance model, and Pairwise Deletion was used for inserting/missing sites in the sequences. The confidence value of NJ subtree was tested using MEGA (3.0). The supporting value of nodes of the subtree was obtained after 1000 replicates.

2 Results

2.1 *Sox* genes analyses in allotetraploids and their original parents

In order to know the different amplicons, all bands in allotetrapoids and their original parents were cloned and sequenced. It was predicted by computers that the 600 bp and 900 bp fragments in allotetraploids, the 600 bp fragment in their original maternal (red crucian carp) and the 900 bp fragment in their original paternal (common carp) all likely had an intron which obeyed the rule of "GT-AG". The 600 bp fragment in allotetraploids was already proved to be a *Sox9a* gene with a 413 bp intron in the conserved HMG-box[12]. The sequence homology analyses of other three fragments were carried out at the following website http://www.ncbi.nlm.nih.gov/blast, which indicated that they shared the highest similarity in different species. According to the rules of nomenclature of *Sox* genes, these three fragments (900 bp in Allotetraploids, 600 bp in red crucian carp and 900 bp in common carp) were named *Atsox9b*, *Rcsox9a* and *Ccsox9b*, respectively. In order to confirm the existence of introns, RT-PCR was conducted with specific primers designed on the basis of the conservative HMG-domain of these three genes to amplify their cDNAs separately (Fig. 1). Compared with their corresponding genomic DNA, the sizes and positions of their introns were verified (Fig. 2).

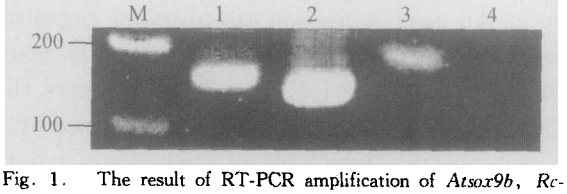

Fig. 1. The result of RT-PCR amplification of *Atsox9b*, *Rcsox9a* and *Ccsox9b* M, 100 bp ladder marker; 1, Allotetraploids *Sox9b* (*Atsox9b*); 2, Red crucian carp *Sox9a* (*Rcsox9a*); 3, Common carp *Sox9b* (*Ccsox9b*); 4, negative control.

Atsox9b (intron 703bp)

TGAAGCGACCCATGAATGCGTTTATGGTTTGGGCTCAAGCGGCGCGCAGGAAACTGGCGGACCAGTATCCACACCTGCAC
AACGCCGAGCTCAGCAAGACCCTCGGCAAACTCTGGAG*GTCAGAGCATTCATTGTTTATGAAGTGTAGGACAACTCCAGA*
AGCCGTAGCAACACTGATTCATTTAACTGCCCAGACAAACTCACTGTATTATTAATATTAAACTGCATTGTTTCATAGCA
TTATTATATAATGTCAGGCAATACTGATAAGCTGTTTGGGAAATAAAGGGTAAAACTATAGCTTAGGCAAATGGTGAATT
AACCCTTACTTCAGTGCAGAAAAAGTTGTGTAGTCACTAAAAACTTTGGTAACACTTTGGAAAAGGGAACACTTACTCAC
TATTAACTATGACTTTTCCCTCTATAAATTCCTAATTTGCTGCTTATTAAAAGTTAGTATGGTAGCTTTTAAGTTTAGGT
ATGAGGTAGGATTAGGGATGTAGAATAAGGTCATGTAAAATAAGACATTAATATGTGCTTAATTACTACTAATAAATGGC
TAATATTCTAGTAATATGCATGCTAATAAGAAACTAGTTAAGAGACCCTAAAATAAAGTGTTACCCAAACTGTTAGTTTT
AAGAAATTTAGTTAGATCTGGATTACGTTATTCTTTTTTTTCCTGAATAATATTTTTTCTTAAGTGAACCCAAAAAGTAA
TCATAGTTGTTTCCAAGAACAATTTATACCTTTTTTGTAGGTTTCTTTATTGTTAACTAAGGTAGATCTTAGTATTGATT
*ATTTCTCATGTGCTTTTGTAG*GTTACTGAATGAGGGCGAGAAGCGTCCATTTGTGGAGGAGGCCGAGCGTCTGAGGGTCC
AGCACAAGAAAGACCACCCCGACTACAAGTACCGACCT

Rcsox9a (intron 401bp)

TGAAGCGACCCATGAACGCGTTTATGGTGTGGGCTCAAGCCGCGCGCAGGAAACTGGCGGATCAGTATCCACACCTGCAC
AACGCCGAGCTCAGCAAGACCCTCGGGAAACTCTGGAG*GTGAGGGCTTGCATTGTTCATCGAGATTAGCACAGCTGCAGA*
AGACCATTCATGTAACTGCTCAGACAAACTTACGGTGTTATTAATATTTAATCGCATTGTTTCATAGCATTATTATATAA
TGATTAGCTTGAGAAGAGGTAACTGCTTGGTGAATTAACCCTTTGTTCACATGCACATCAGTGCAGAAACAGTTAAAAAC
ATTCCTAATTTGCTGCTTATTGATACTTAGTAAAGTAGTTTTTAAGTTTTGGTGCAAGGTTGGATTATGGGTTGTACTAT
ACTTTTTTCCCCTGAATAACCTTTTTTTTAAAGTGAACCCAAAAAGCAATAATCATAGTTGTTTCAAGAACAATTTGTAT
*CTTTTTTTGTAGGTTTCTTTATTGTGAACTGTTTTTTAG*GTTACTGAATGAGGGCGAGAAGCcGTCCGTTCGTGGAGGAG
GCCGAGCGTCTGAGGGTCCAGCACAAGAAAGACCACCCCGACTACAAGTACCGACCT

Ccsox9b (intron 714bp)

TGAAGCGACCCATGAATGCGTTTATGGTTTGGGCTCAAGCGGCGCGCAGGAAACTGGCGGACCAGTATCCACACCTGCAC
AACGCCGAGCTCAGCAAGACCCTCGGCAAACTCTGGAG*GTCAGAGCATTCATTGTTTATGAAGTGTAGGACAACTCCAGA*
AGCCGGAGCAACACTGATTCATTTAACTGCCCAGACAAACTCACTGTATTATTAATATTAAACTGCATTGTTTCATAGCA
TTATTATATAATGTCAGGCAATACTGATAAGCTGTTTGGGAAATAAAGGGTAAAACTATAGCTTAGGCAAATGGTGAATT
AACCCTTACTTCACATGCACATCATTGCAGAAAAAGTTGTGTAGTCACTAAAAACTTTGGTAACACTTTGGAAAAGGGAA
CACTTACTCACTATTAACTATGACTTTTCCCTCTATAAATTCCTAATTTGCTGCTTATTAAAAGTTAGTATGGTAGCTTT
TAAGTTTAGGTATGAGGTAGGATTAGGGATGTAGAATAAGGTCATGTAAAATAAGACATTAATATGTGCTTAATTACTAC
TAATAAATGGCTAATATTCAAGTAATATGCATGCTAATAAGAAACTAGTTAAGAGACCCTAAAATAAAGTGTTACCCAAA
CTTTTAGTTTTAAGAAATTTAGTTAGATCTGGATTACGTTATTCTTTTTTTTCCTGAATAATATTTTTTCTTAAGTGAAC
CCAAAAAGTAATCATAGTTGTTTCCAAGAACAATTTATACCTTTTTTGTAGGTTTCTTTATTGTTAACTAAGGTAGATCT
*TAGTATTGATTATTTCTCATGTGCTTTTGTAG*GTTACTGAATGAGGGCGAGAAGCGTCCATTTGTGGAGGAGGCCGAGCG
TCTGAGGGTCCAGCACAAGAAAGACCACCCCGACTACAAGTACCGACCT

Fig. 2. Analyses of sizes and positions of introns in *Atsox9b*, *Rcsox9a* and *Ccsox9b*. The underline indicates the splicing sites of introns, the italic shows the nucleic acid sequences of introns and the boxed are the primers for RT-PCR.

2.2 Analyses of intron genetic variability in *Sox9* HMG-domain in allotetraploids and their original parents

Atsox9a and *Atsox9b* in allotetraploids include a 413 bp and 703 bp intron in the HMG-box respectively, while their original maternal (red crucian carp) *Sox9a* (*Rcsox9a*) and their original paternal (common carp) *Sox9b* (*Ccsox9b*) had a 401 bp and 714 bp intron respectively. The nucleic acid sequences of introns were aligned by Clustal W (1.83) program (Fig.3). As shown in Fig. 3(a), the comparison of *Atsox9a* with *Rcsox9a* reveals 7 variable nucleic acids and 390 conserved nucleic acids and some insertions and deletions. There was an insertion of the 14 bp (AAGGAAATGCTGAT) at the site of 127—140 nt in *Atsox9a* and an insertion of the 2 bp (TA) at the position of 311—312 nt in *Rcsox9a*. The nucleic acid sequences of introns between *Atsox9a* and *Rcsox9a* showed 94.4% identity. In Fig. 3(b), the alignment of *Atsox9b* and *Ccsox9b* reveals 3 variable nucleic acids and 689 conservative nucleic acids and some insertions or gaps. There was a 11 bp (ATGCACATCAT) deletion at the site of 217—227 nt in the intron of *Atsox9b*, and a high similarity (97.8%) was observed between introns of *Atsox9b* and *Ccsox9b*. The content of G + C in introns and exons of *Atsox9a*, *Atsox9b*, *Rcsox9a* and *Ccsox9b* was calculated using the method of nucleotide composition in MEGA (3.0) software package. As shown in Table 1, the content of G + C in introns was lower than that in the corresponding exons.

Table 1. The G + C content of exons and introns in different genes

Gene[a)]	Exon(%)	Intron(%)
Atsox9a	58.8	33.4
Atsox9b	57.7	30.6
Rcsox9a	59.3	32.9
Ccsox9b	57.7	30.6

a) At, allotetraploids; Rc, red crucian carp; Cc, common carp.

(a)

```
Atsox9a-intron      GTGAGGGCTTGCACTGTTCATCGAGATTAGCACAGCTGCAGAAGACCATTCATTCAACTG 60
Rcsox9a-intron      GTGAGGGCTTGCATTGTTCATCGAGATTAGCACAGCTGCAGAAGACCATTCATGTAACTG 60
                    ************ ************************************  *****

Atsox9a-intron      CTCAGACAAACTTACTGTGTTATTAATATTTAATCGCATTGTTTCATAGCATTATTATAT 120
Rcsox9a-intron      CTCAGACAAACTTACGGTGTTATTAATATTTAATCGCATTGTTTCATAGCATTATTATAT 120
                    ************** ********************************************

Atsox9a-intron      AATGATAAGGAAATGCTGATTAGCTTGAGAAGAGGTAACTGCTTGGTGAATTAACCCTTT 180
Rcsox9a-intron      AATGAT--------------TAGCTTGAGAAGAGGTAACTGCTTGGTGAATTAACCCTTT 166
                    ******              ****************************************

Atsox9a-intron      GTTCACATGCACATCAGTGCAGAAACAGTTAAAAACATTCCTAATTTGCTGCTTATTGAT 240
Rcsox9a-intron      GTTCACATGCACATCAGTGCAGAAACAGTTAAAAACATTCCTAATTTGCTGCTTATTGAT 226
                    ************************************************************

Atsox9a-intron      ACTTAGTAGAGTAGTTTTTAAGTTTTGGTGCGAGGTTGGATTATGGGTTGTACTATACTT 300
Rcsox9a-intron      ACTTAGTAAAGTAGTTTTTAAGTTTTGGTGCAAGGTTGGATTATGGGTTGTACTATACTT 286
                    ******** ********************** ****************************

Atsox9a-intron      TTTTCCCCTGAATAACATTTTTTT--AAGTGAACCCAAAAAGCAATAATCATAGTTGTTT 358
Rcsox9a-intron      TTTTCCCCTGAATAACCTTTTTTTTAAAGTGAACCCAAAAAGCAATAATCATAGTTGTTT 346
                    **************** *******  ***********************************

Atsox9a-intron      CAAGAACAATTTGTATCTTTTTTTGTAGGTTTCTTTATTGTGAACTGTTTTTTAG 413
Rcsox9a-intron      CAAGAACAATTTGTATCTTTTTTTGTAGGTTTCTTTATTGTGAACTGTTTTTTAG 401
                    ****************************************************
```

(b)

```
Atsox9b-intron      GTCAGAGCATTCATTGTTTATGAAGTGTAGGACAACTCCAGAAGCCGTAGCAACACTGAT 60
Ccsox9b-intron      GTCAGAGCATTCATTGTTTATGAAGTGTAGGACAACTCCAGAAGCCGGAGCAACACTGAT 60
                    ********************************************** ***********

Atsox9b-intron      TCATTTAACTGCCCAGACAAACTCACTGTATTATTAATATTAAACTGCATTGTTTCATAG 120
Ccsox9b-intron      TCATTTAACTGCCCAGACAAACTCACTGTATTATTAATATTAAACTGCATTGTTTCATAG 120
                    ************************************************************

Atsox9b-intron      CATTATTATATAATGTCAGGCAATACTGATAAGCTGTTTGGGAAATAAAGGGTAAAACTA 180
Ccsox9b-intron      CATTATTATATAATGTCAGGCAATACTGATAAGCTGTTTGGGAAATAAAGGGTAAAACTA 180
                    ************************************************************

Atsox9b-intron      TAGCTTAGGCAAATGGTGAATTAACCCTTACTTCAG----------TGCAGAAAAAGTT 229
Ccsox9b-intron      TAGCTTAGGCAAATGGTGAATTAACCCTTACTTCACATGCACATCATTGCAGAAAAAGTT 240
                    ********************************** ************

Atsox9b-intron      GTGTAGTCACTAAAAACTTTGGTAACACTTTGGAAAAGGGAACACTTACTCACTATTAAC 289
Ccsox9b-intron      GTGTAGTCACTAAAAACTTTGGTAACACTTTGGAAAAGGGAACACTTACTCACTATTAAC 300
                    ************************************************************

Atsox9b-intron      TATGACTTTTCCCTCTATAAATTCCTAATTTGCTGCTTATTAAAAGTTAGTATGGTAGCT 349
Ccsox9b-intron      TATGACTTTTCCCTCTATAAATTCCTAATTTGCTGCTTATTAAAAGTTAGTATGGTAGCT 360
                    ************************************************************

Atsox9b-intron      TTTAAGTTTAGGTATGAGGTAGGATTAGGGATGTAGAATAAGGTCATGTAAAATAAGACA 409
Ccsox9b-intron      TTTAAGTTTAGGTATGAGGTAGGATTAGGGATGTAGAATAAGGTCATGTAAAATAAGACA 420
                    ************************************************************

Atsox9b-intron      TTAATATGTGCTTAATTACTACTAATAAATGGCTAATATTCTAGTAATATGCATGCTAAT 469
Ccsox9b-intron      TTAATATGTGCTTAATTACTACTAATAAATGGCTAATATTCAAGTAATATGCATGCTAAT 480
                    ****************************************** ****************

Atsox9b-intron      AAGAAACTAGTTAAGAGACCCTAAAATAAAGTGTTACCCAAACTGTTAGTTTTAAGAAAT 529
Ccsox9b-intron      AAGAAACTAGTTAAGAGACCCTAAAATAAAGTGTTACCCAAACTTTTAGTTTTAAGAAAT 540
                    ****************************************** ****************

Atsox9b-intron      TTAGTTAGATCTGGATTACGTTATTCTTTTTTTTCCTGAATAATATTTTTTCTTAAGTGA 589
Ccsox9b-intron      TTAGTTAGATCTGGATTACGTTATTCTTTTTTTTCCTGAATAATATTTTTTCTTAAGTGA 600
                    ************************************************************

Atsox9b-intron      ACCCAAAAAGTAATCATAGTTGTTTCCAAGAACAATTTATACCTTTTTTGTAGGTTTCTT 649
Ccsox9b-intron      ACCCAAAAAGTAATCATAGTTGTTTCCAAGAACAATTTATACCTTTTTTGTAGGTTTCTT 660
                    ************************************************************

Atsox9b-intron      TATTGTTAACTAAGGTAGATCTTAGTATTGATTATTTCTCATGTGCTTTTGTAG 703
Ccsox9b-intron      TATTGTTAACTAAGGTAGATCTTAGTATTGATTATTTCTCATGTGCTTTTGTAG 714
                    ******************************************************
```

Fig. 3. The nucleic acid sequences comparison of introns of *Sox*9 HMG-box in allotetraploids and their original parents. (a) The comparison between allotetraploids and their original mother (red crucian carp); (b) the comparison between Allotetraploids and their original father (common carp). At, allotetraploids; Rc, red crucian carp; Cc, common carp. " * " indicates the same nucleotide sites; "." implies the different nucleic acid sites; " – " shows the missing nucleic acid sites.

2.3 Phylogenetic analysis

Using NJ method, a phylogenetic tree was constructed based on the introns of HMG-box in *Sox9* genes from allotetraploids and their original maternal red crucian carp (*Carassius carassius* red var.) and paternal common carp (*Cyprinus carpio* L.), zebrafish (*Brachydanio rerio*) and salmon (*Oncorhynchus keta*). Fig. 4 illustrates that allotetraploids and their original parents are clustered together, separating from zebrafish and salmon. *Rcsox9a* and *Atsox9a*, *Ccsox9b* and *Atsox9b* formed a sister group respectively, demonstrating that the relationship between allotetraploids and their original parents is closer than that in other fishes. Meanwhile, it also indicates that allotetraploids are the hybrid offspring of red crucian carp and common carp, possessing the genetic characteristics of their parents.

3 Discussion

With the accumulated knowledge of the gene functions implemented by encoding regions, more and more attention is now paid to the functions of introns and sequences between genes. It has been well recognized that there are a majority of introns in pre-mRNAs in eukaryotic cells, but their functions remain unveiled. Obviously, the existence of introns in high organism has increased the mutation endurance greatly, but most random mutations happening in introns do not have serious influences on organisms, because the sizes of introns are much larger than those of exons in high organism. The positions of introns were found to be considerably conserved during the evolution of most genes, while their sizes and sequences had no distinct conservation. For instance, no remarkable similarity was found in introns of HMG-box in salmon *Sox9b* (GenBank accession number, AY573260) and zebrafish *Sox9a* (GenBank accession number, AY090035) with the lengths of 303 bp and 480 bp respectively. However, we cannot explain the reasons and significance of the conservation of introns yet. Additionally, it was discovered that the insertion sites commonly follow the rule of "GT-AG", the splicing sites of introns of the four *Sox9* genes in this report all obeyed the rule of "GT-AG". Another type of introns obeying the rule of "AT-AC" also existed in quite a few high plants, which starts from two nucleic acids "AT" and ends as "AC"[14]. From the analyses of the *Sox9* sequences in different species we could know that the similarity of introns in *Sox9s* HMG-box is high in these species, despite of a little lower than that of exons. For example, the identity of introns between *Atsox9a* and *Rcsox9a* was 94.4%, but 99.5% for their exons. The phylogenetic analysis based on the introns of *Sox9* HMG-domain from different fishes indicated that the evolutionary relationships among these fishes were consistent with those obtained from the traditional taxology. Namely, allotetraploids, red crucian carp, common carp and zebrafish clustered together and formed the group of Cyprinidae, apart from salmon which belongs to the family Salmonidae (Fig. 4). In addition, the G+C content of introns of *Sox9* HMG-box in allotetraploids and their original parents was lower than that of exons, which was consistent with the result reported previously that a high A+T content existed in non-encoding regions[15]. From our results we may say that introns could be used as genetic markers for the studies on the evolutionary relationships in relatively close species.

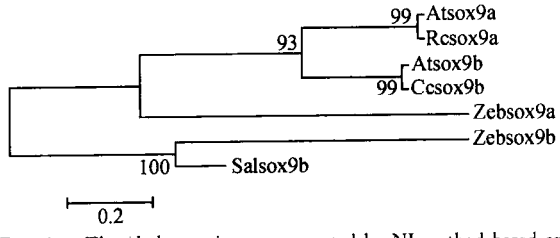

Fig. 4. The phylogenetic tree generated by NJ method based on introns of *Sox9* HMG-box. At, allotetraploids; Rc, red crucian carp; Cc, common carp; Zeb, zebrafish; Sal, salmon; Numbers on each branch are the bootstrap values in thousand runs.

For the moment among more than 40 *SOX/Sox* genes known[8,16,17], only *Sox5*, *Sox9*, *Sox17* and *Sox20* were found to have different introns. Takase et al. reported that there were two kinds of splicing of the intron in frog *Sox9*, which produced two types of proteins, probably inducing two differently developmental and differential mechanisms[18]. From the knowledge that *Sox9* plays an important role in sex determination and differentiation, we assume that the splicing of introns in *Sox9* HMG-box may be important for the functions of *Sox9* in vertebrate development.

Acknowledgements We like to thank Dr. YAO Zhanzhou and LUO Kaikun as well as the faculty in our laboratory for their zealous assistance in this work.

References

1 Wegner M. From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res, 1999, 27(6): 1409—1420

2 Wilson M and Koopman P. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. Curr Opin Genet Dev, 2002, 12(4): 441—446
3 Schafer AJ, Foster JW, Kwok C, et al. Campomelic dysplasia with XY sex reversal: diverse phenotypes resulting from mutations in a single gene. Ann N Y Acad Sci, 1996, 785: 137—149
4 Mount S and Henikoff S. Nested geness take flight. Curr Biol, 1993, 3: 372—374
5 Mikol DD, Rongnoparut P, Allwardt BA, et al. The oligodendrocyte-myelin glycoprotein of mouse: primary structure and gene structure. Genomics, 1993, 17(3): 604—610
6 Higuchi M, Single FN, Kohler M, et al. RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. Cell, 1993, 75(7): 1361—1370
7 Howe K and Ares M. Intron self-complementarity enforces exon inclusion in a yeast pre-mRNA. Proc Natl Acad Sci USA, 1997, 94(23): 12467—12472
8 Kanai Y, Kanai-Azuma M, Noce T, 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—681
9 Tee MK, Babalola GO, Aza-Blanc P, et al. A promoter within intron 35 of the human C4A gene initiates abundant adrenal-specific transcription of a 1 kb RNA: location of a cryptic CYP21 promoter element? Hum Mol Genet, 1995, 4(11): 2109—2116
10 Kiss T and Filipowicz W. Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. Genes Dev, 1995, 9(11): 1411—1424
11 Wang N and Chen RS. Comparison based on phylogenetic analyses of introns and exons. Chin Sci Bull, 1999, 44(19): 2095—2102
12 Liu JF, Liu SJ, Zhang C, et al. The conserved region sequence cloning of *Sox9a* and analysis of the splicing site of the intron in Allotetraploid fish. Prog Nat Sci, 2004, 14(6): 641—645
13 Kumar S, Tamura K and Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform, 2004, 5(2): 150—163
14 Tarn WY and Steitz JA. Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. Trends Biochem Sci, 1997, 22(4): 132—137
15 Li WH and Graur D. Fundamentals of Molecular Evolution. Massachusetts: IAC Publishers, 1991
16 Wright EM, Snopek B and Koopman P. Seven new members of the Sox gene family expressed during mouse development. Nucleic Acids Res, 1993, 21(3): 744
17 Foster JW, Dominguez-Steglich MA, Guioli S, et al. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature, 1994, 372(6506): 525—530
18 Takase M, Noguchi S and Nakamura M. 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—254