Isolation and sequencing of seven Sox genes from the lacertid lizard *Eremias breuchleyi*

Jie Li, Ping-Ping Zheng, Jiao-Lian Song, Jin-Long Rui and Liu-Wang Nie

*Life Science College, Anhui Normal University, Wuhu, Anhui province, China*

Abstract

The Sox family of genes shares a high sequence similarity with the HMG box region of the human Y chromosomal gene, *SRY*. We used highly degenerate primers to clone and sequence seven *Eremias breuchleyi* Sox genes (*EbSox2, EbSox3, EbSox4, EbSox11, EbSox12, EbSox14 and EbSox21*). A database search for the cloned sequences revealed the following percentage identity with the homologous human SOX genes: *EbSox2 = 96%, EbSox3 = 88%, EbSox4 = 94%, EbSox11 = 99%, EbSox12 = 96%, EbSox14 = 98%, EbSox21 = 97%*. Cluster analysis indicates that they seem to belong to group B and group C of Sox gene family, respectively.

Key words: *Eremias breuchleyi*, PCR, sequence analysis, Sox genes, SSCP.

Received: October 10, 2005; Accepted: December 21, 2005.

The *Sox* (*SRY* related HMG-box gene, *Sox*) genes form a large family which is characterized by a highly conserved DNA-binding and share a high sequence similarity with the HMG (high mobility group, HMG) box region of the human Y chromosomal gene, *SRY* (*Sex-determining region of Y chromosome, SRY*) (Hawkins JR, 1994; Pevny LH, 1997). More than 30 *Sox* genes have been identified in mammals and their orthologues have been found in a wide range of other metazoans (Hagiuda et al., 2003). The *Sox* genes are highly conserved and are known to play important roles in embryonic development including roles in gonad, central nervous system, neural crest and skeletal development (Nagai, 2001). For instance, mutation in the *SOX9* gene has been associated with sex reversal in men (Foster et al., 1994; Wagner et al., 1994), while targeted mutagenesis in mice has shown that *Sox4* is essential for heart and lymphocyte development (Schilham et al., 1996). In addition, tissue culture experiments have shown that mouse *Sox1, Sox2* and *Sox3* genes are expressed mainly in nervous system development and are involved in determining the fate of neuronal cells (Collignon et al., 1996; Pevny et al., 1998; Li et al., 1998). However, the role of these genes in the development and differentiation of reptiles has yet to be explored.

The lacertid lizard, *Eremias breuchleyi*, lacks identifiable sex chromosomes but it appears that the sex determination in this species might be genetic because incubation temperature does not influence sex development. As a prelude to understanding the involvement of *SRY*-like genes in the development and differentiation of reptiles, we attempted to clone the *Sox* genes family of *E. breuchleyi* using the polymerase chain reaction (PCR). In the present paper we report the cloning and nucleotide sequence of seven *E. breuchleyi* Sox genes which show extensive homology with the *Sox* genes of various other vertebrate taxa. The phylogenetic evolution of *Sox* genes is also discussed.

Two male and two female *E. breuchleyi* were captured from Qianshan, Suzhou, Anhui province, China and the genomic DNA isolated from muscle tissues using routine protocols (Sambrook et al., 1989). A pair of PCR primers was designed using a multiple alignment of a HMG-box sequence representative of *SRY/Sox* gene family, primer 1 being: 5’-AGCGACCCATGAA(CT)GC(AGCT)TT(CT)AT(AGCT)G-3’ and primer 2 being: 5’-ACGAGGTCCATA(C)TT(AG)TA(AG)(T)(AGCT)GG-3’. The amplification fragment length was 216 bp using these primer pairs.

Amplifications were carried out in a total volume of 25 μL containing 100 ng of sample genomic DNA, 1.5 mM Mg2+, 120 μM dNTP, 0.3 μM of each primer, 1.25 units of Taq polymerase and H2O. The PCR cycling condition were 5 min at 97 °C, followed by 35 cycles of 40 s at 94 °C, 40 s at 55.5 °C and 50 s at 72 °C with a final 10 min elongation at 72 °C.

The PCR products were detected on 1.7% (w/v) agarose gels and cloned using the pMD 18-T vector (purchased from TaKaRa). 100 white clones were transferred to a plate of clones from initial culture plate and 81 positive clones
with inserted Sox DNA were confirmed using colony PCR. The distinct positive clones were screened using single-strand conformation polymorphism (SSCP) analysis (Nie, Shan and Guo, 1999) and sequenced using the universal sequencing primers on an ABI377 auto-sequencer. DNA sequences were analyzed using the BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and CLUSTALX programs (http://www.igh.cnrs.fr/bin/clustalwguess.cgi) and a phylogenetic tree was constructed using the Molecular Evolutionary Genetic Analysis (MEGA) software.

We succeeded in cloning a 220 bp fragment using *E. breuchleyi* genomic DNA as template. Seven distinct Sox-positive clones, representing distinct Sox genes, were selected from male and female *E. breuchleyi* but there was no sexual difference between them (Figure 1a). We named the genes *Eremias breuchleyi* Sox (*EbSox*) based on BLAST analysis, the isolated genes being *EbSox2* (DQ067423), *EbSox3* (DQ067425), *EbSox4* (DQ067426), *EbSox11* (DQ067427), *EbSox12* (DQ067428), *EbSox14* (DQ067430), *EbSox21* (DQ067433) (sequence accession numbers of GenBank in parentheses). The putative amino acid sequences of these Sox genes are shown in Figure 1b. A database search for the cloned sequences revealed the following percentage identity with the homologous human SOX genes: *EbSox2* = 96%, *EbSox3* = 88%, *EbSox4* = 94%, *EbSox11* = 99%, *EbSox12* = 96%, *EbSox14* = 98%, *EbSox21* = 97%.

The HMG domain sequence RPMNAFMCW (positions 2–10) appears to be conserved for all SOX sequences (Bowles, Schepers and Koopman, 2000). Figure 2a shows the sequence comparison of 72 conserved HMG-box amino acid residues from the 29 Sox genes sequences (Table 1) and the seven sequences cloned by us. Comparison of these 36 HMG domains showed that they were clustered within distinct phylogenetic sub-groups (Figure 2b). The previous studies had showed that all characteristic SOX/Sox genes can be divided into ten groups (A-J) (Bowles, Schepers and Koopman, 2000). In our study we found that the seven *E. breuchleyi* Sox genes did not cluster together but were distributed between the B and C Sox groups. It is known that Sox1, Sox2 and Sox3 are members of the family of HMG DNA-binding domain containing transcription factors related to the testis-determination Sry gene and, along with the recently discovered Sox14 and Sox21 genes, comprise the group B subfamily of Sox genes (Collignon et al., 1996), while the members of group C (Sox4, Sox11, Sox12, Sox22 and Sox24) encode highly conserved N- and C- terminal amino acid sequences (Collignon et al., 1996; Pevny and Lovell-Badge, 1997; Arsic et al., 1998; Rex et al., 1997; Kamachi et al., 1995; Hargrave et al., 2000; Uchikawa et al., 1999). It should be remembered, however, that the PCR primer set used by us may have had a bias leading to preferential amplification of group B and C *E. breuchleyi* Sox genes.

Figure 1 - Sox gene HMG-boxes of *Eremias breuchleyi*. (a) DNA sequence; dots indicate identities with *EbSox3*; (b) putative amino acid sequence; dots indicate identities with *EbSox2*.

The members of the Sox genes family have been highly conserved though evolution and have been found in a wide variety of species. In the Sox3 gene product (Figure 2a), there is a D (Asp) amino acid at position 66 in bird, mouse and human but it is N (Asn) in *E. breuchleyi*. At position 20 and 42 the amino acid is identical between human and mouse SOX3 protein but shows a conservative change in other species, while in the rest of the protein the Sox3 gene amino acid sequences are highly conserved. It is probable that gene duplication has caused the diversity seen in the HMG box superfamily in which the Sox family of genes shows the highest mutation rate (Laudet, et al., 1993). In the phylogeny of the Sox family the Sox4 gene is considered to be an early offshoot and the SRY gene a recent
The occurrence of the sequence conservation among the Sox4 homologues in amniotes is interesting because the SRY gene shows rapid gene evolution in mammals which is possibly caused by Y-linked inheritance (Tucker and Lundrigan, 1993), although an ancient conserved function might also have restricted the divergence of Sox4 gene homologues in amniotes. In fact, the amino acid sequences in the HMG box regions are highly conserved among different species including Eremias breuchleyi, but their functional conservation in sex determination and differentiation needs to be further studied.

Acknowledgement

This work is supported by the science foundation of the Key Laboratory of Biotic Environment and Ecological Safety in Anhui Province, China. We thank two anonymous reviewers and the technical editor for critically reviewing of the manuscript.

References


Associate Editor: André Luiz Paranhos Perondini