Genetics and Molecular Biology, 29, 3, 576-579 (2006)

www.sbg.org.br

Short Communication

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.

Send correspondence to Liu-wang Nie. Life Science College, Anhui Normal University, 1 East Beijing Road, wuhu, 241000 Anhui, China. E-mail: lwnie@mail.ahnu.edu.cn.

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'-ACGAGGTCGATA(CT)TT(AG)TA(AG)T(CT)(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 Mg^{2+}, 120 μM dNTP, 0.3 μM of each primer, 1.25 units of Taq polymerase and H_2O. 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 method (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/clustalxguess.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* (DQ 067427), *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 RPMNAFMVW (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.

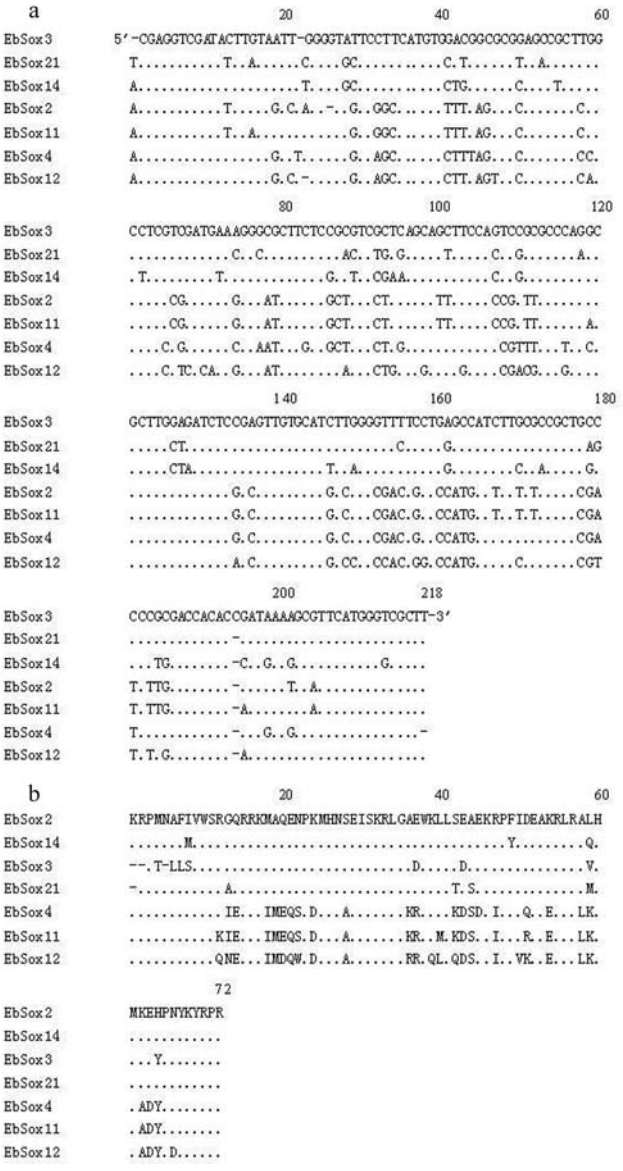

Figure 1 - *Sox* gene HMG-boxes of *Eremias breuchleyi*. (a) DNA sequence; dots indicate identities with *EbSox3*; (b) putative amino acid sequence; dots indicate identities with *EbSox2*.

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

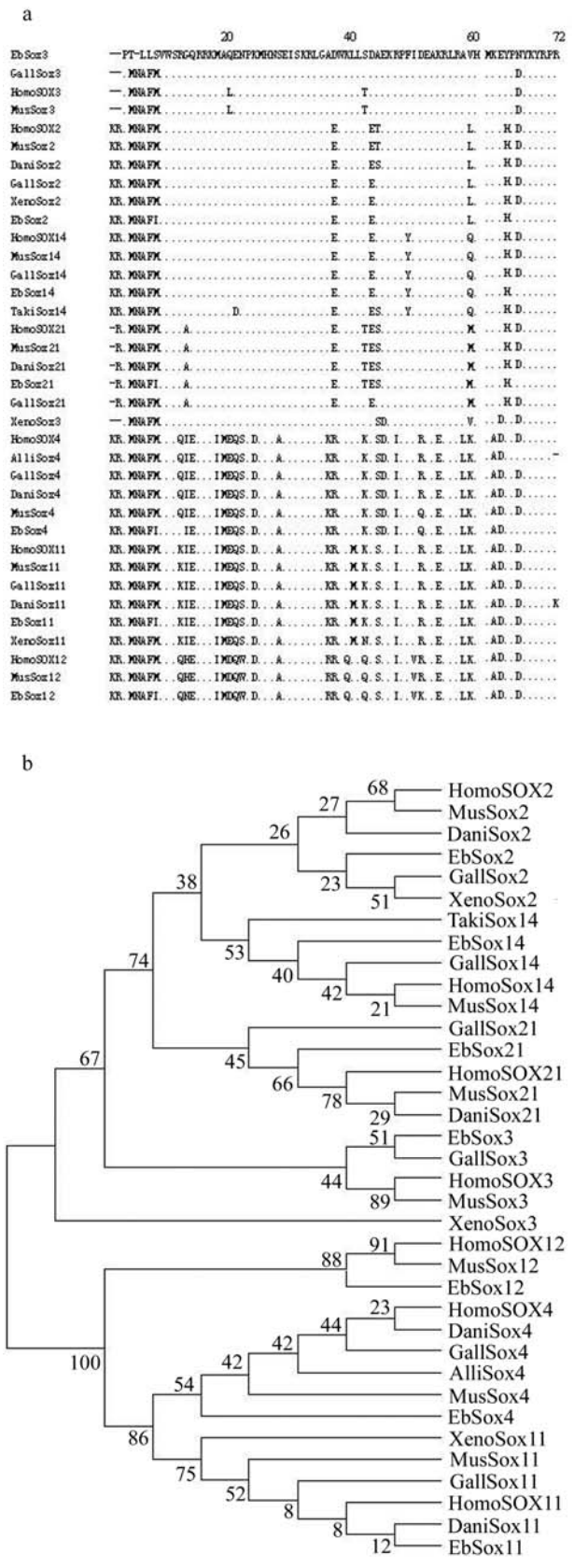

Figure 2 - Comparative analysis of *Sox* gene HMG-boxes in different species. (a) multiple alignment results; dots indicate identities with *EbSox3*; (b) phylogenetic analysis of the *Sox* gene family.

Table 1 - GenBank sequences searched in this study.

Sequence	Accession number	Sequence	Accession number
Homo sapiens		*Danio rerio*	
HomoSOX2	NP 003097.1	*DaniSox2*	NP 998283.1
HomoSOX3	NP 005625.2	*DaniSox4*	NP 998287.1
HomoSOX4	NP 003098.1	*DaniSox11*	NP 571411.1
HomoSOX11	NP 003099.1	*DaniSox21*	NP 001009888.1
HomoSOX12	NP 008474.2		
HomoSOX14	NP 004180.1	*Xenopus laevis*	
HomoSOX21	NP 009015.1	*XenoSox2*	AAB62821.1
		XenoSox11	AAH70707.1
Mus musculus			
MusSox2	NP 035573.2	*Alligator sinensis*	
MusSox3	NP 033263.1	*AlliSox4*	AA017690.1
MusSox4	NP 033264.2		
MusSox11	NP 033260.4	*Takifugu rubripes*	
MusSox12	NP 035568.1	*TakiSox14*	AAQ18498.1
MusSox14	XP 284529.3		
MusSox21	NP 808421.1		
Gallus gallus			
GallSox2	NP 990519.1		
GallSox3	NP 989526.1		
GallSox4	NP 989815.1		
GallSox11	NP 990518.1		
GallSox14	NP 990092.1		
GallSox21	BAA77266.1		

entry (Laudet, *et al*., 1993). 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 Provice, China. We thank two anonymous reviewers and the technical editor for critically reviewing of the manuscript.

References

Arsic N, Rajic T, Stanojcic S, Goodfellow PN and Stevanovic M (1998) Characterisation and mapping of the human SOX14 gene. Cytogenet Cell Genet 83:139-146.

Bowles J, Schepers G and Koopman P (2000) Phylogeny of the *Sox* family of developmental transcription factors based on sequence and structural indicators. Dev Boil 227:239-255.

Collignon J, Sockanathan S, Hacker A, Cohen-Tannoudji M, Norris D, Rastan S, Stevanovic M, Goodfellow PN and Lovell-Badge R (1996) A comparison of the properties of *Sox3*, with *Sry* and two related genes, *Sox1* and *Sox2*. Development 122:509-520.

Foster JW, Dominguez-Steglich MA, Guioli S, Kwok C, Weller PA, Weissenbach J, Mansour S, Young ID, Goodfellow PN, Brook JD and Schafer AJ (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an *SRY*-related gene. Nature 372:525-530.

Hagiuda J, Hiraoka Y, Hasegawa M, Ogawa M and Aiso S (2003) A novel *Xenopus laevis* SRY-related gene, *xSox33*. Biochim Biophys Acta 1628:140-145.

Hargrave M, Karunaratne A, Cox L, Wood S, Koopmam P and Yamada T (2000) The HMG box transcription factor gene *Sox14* marks a novel subset of ventral interneurons and is regulated by Sonic hedgehog. Dev Boil 219:142-153.

Hawkins JR (1994) Sex determination. Hum Mol Genet 3:1463-1470.

Kamachi Y, Sockanathan S, Liu Q, Breitman M, Lovell-Badge R and Kondoh H (1995) Involvement of SOX proteins in lens-specific activation of crystallin genes. EMBO J 14:3510-3519.

Laudet V, Stehelin D and Clevers H (1993) Ancestry and diversity of the HMG box superfamily. Nucleic Acids Res 21:2493-2501.

Li M, Pevny L, Lovell-Badge R and Smith A (1998) Generation of purified neural precursors from embryonic stem cells by lineage selection. Curr Biol 8:71-974.

Nagai K (2001) Molecular evolution of Sry and *Sox* gene. Gene 270:161-169.

Nie LW, Shan XN and Guo CW (1999) The PCR amplification and SSCP analysis of *Sox* gene in turtles (*Platysternon megacephalum and Cistoclemmys flavomarginatus*). Chin J Appl Environ Boil 5:378-381.

Pevny LH and Lovell-Badge R (1997) *Sox* genes find their feet. Curr Opin Genet Dov 7:338-344.

Pevny LH, Sockanathan S, Placzek M and Lovell-Badge R (1998) A role for *SOX1* in neural determination. Development 125:1967-78.

Rex M, Uwanogho DA, Orme A, Scotting PJ and Sharpe PT (1997) *cSox21* exhibits a complex and dynamic pattern of transcription during embryonic development of the chick central nervous system. Mech Dev 66:39-53.

Sambrook J, Friston EF and Maniatis T (1989) Molecular Cloning, a Laboratory Mannual. Cold Spring Harbor Laboratory Press, New York.

Schilham MW, Oosterwegel MA, Moerer P, , de Boer PA, , Verbeek S, Lamers WH, Kruisbeek AM, Cumano A and Clevers H (1996) Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking *Sox-4*. Nature 380:711-714.

Tucker PK and Lundrigan BL (1993) Rapid evolution of the sex determining locus in old world mice and rats. Nature 364:715-717.

Uchikawa M, Kamachi Y and Kondoh H (1999) Two distinct subgroups of group B *Sox* genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. Mech Dev 84:103-120.

Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Bricarelli FD, Keutel J, Hustert E, Wolf U, Tommerup N, Schempp W and Scherer G (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene *SOX9*. Cell 79:1111-20.

Associate Editor: André Luiz Paranhos Perondini