African Journal of Biotechnology Vol. 8 (8), pp. 1441-1448, 20 April, 2009
Available online at http://www.academicjournals.org/AJB
ISSN 1684–5315 © 2009 Academic Journals

Full Length Research Paper

Cloning and analysis of the HMG domains of ten *Sox* genes from *Bombina maxima* (Amphibia: Anura)

Jingjing Wang, Ning Wang and Liu-wang Nie*

The Provincial Key Lab. of the Conservation and Exploitation Research of Biological Resources in Anhui, College of Life Sciences, Anhui Normal University, Wuhu, Anhui 241000, China.

Accepted 5 March, 2009

***Sox* is a large gene family which encodes *Sry*-related transcription factors and contains a HMG box that is responsible for the sequence-specific DNA binding. In this paper, we obtained ten clones representing HMG box-containing *Sox* genes (*BmSox1a*, *BmSox1b*, *BmSox3a*, *BmSox3b*, *BmSox3c*, *BmSox11a*, *BmSox11b*, *BmSox11c*, *BmSox14, BmSox33*) from male and female *Bombina maxima* respectively and no sexual differences were observed, using highly degenerate primers designed from the conservative motif (HMG-box) of the human *SRY* gene. We want to reveal the gene duplication process and gene trees for the *Sox* genes in this toad. The sequences analysis indicated that *Sox1*, *Sox3* and *Sox11* gene may be duplicated. The cloned *Sox* genes shared high sequence identity to the homologous human *SOX* genes. Based on the amino acid sequence similarities, the phylogenetic analysis was carried out and the results suggested that nine of ten HMG domain-encoding sequences are members of the *SoxB* and *SoxC* groups. *BmSox33* and its homologous gene *xSox33*(*Xenopus laevis*) were found only in Discoglossidae and Xenopodidae in amphibian, but *Sox33* has not been found in other vertebrates up to now, and the origin of *Sox33* gene is an interesting target in phylogenesis.**

Key words: *Bombina maxima, Sox* genes, HMG box, *BmSox33.*

INTRODUCTION

Sox genes form a large gene family by homology to the HMG-box region of *Sry* (sex-determining region of the Y chromosome) (Pevny et al., 1997; Wegner et al., 1999). The HMG- box domain of *Sox* proteins, which located on the C-terminal end by a basic tail, is a characteristic L-shaped domain of about 79 amino acid residues and an α-helical DNA-binding domain that binds within the minor groove to induce DNA bending (Ner ,1992; Phillips et al, 2006). Besides the HMG domain, most *Sox* proteins harbor several functional domains, such as the trans-activation domain, the transrepression domain and the dimerization domain (Lefebvre et al., 2007).

The members of *Sox* genes family have been identified in a broad range of animal taxa, through studies in mammals, birds, reptiles, amphibians and fish, insects and nematodes (Pevny et al., 1997; Wegner et al., 1999; Koopman et al., 2004, Sessa and Bianchi, 2007). More than 30 *Sox* genes had been characterized in mammals, leading to a recent classification of *Sox* proteins into 10 groups (from A to I), with two B subgroups, B1 and B2 (Bowles et al., 2000; Lefebvre et al., 2007). *Sox* proteins within the same group share a high degree of amino acid sequences identity (generally 70–95%) both within and outside the HMG box, whereas *Sox* proteins from different groups share partial identity (≥46%) in the HMG box domain and none outside this domain (Lefebvre et al., 2007; Sessa and Bianchi, 2007). Most of them are expressed in a variety of tissues and have diverse roles including sex differentiation, stemness, neurogenesis and gliogenesis, neural crest development, skeletogenesis and so on (Bowles et al., 2000; Lefebvre et al., 2007).

Bombina maxima (*B.maxima*) belongs to discoglossidae which is the most conservative family in Anura. It is very significant to understand the *Sox* gene duplication process and phylogenetic analysis of this species in toad evolution. We have cloned the HMG domain-encoding sequences of ten *Sox* genes from the genome of *B. maxima (BmSox)*. Because the HMG domain may be

*Corresponding author. E-mail: lwnie@mail.ahnu.edu.cn

considered as an independent evolutionary unit, we predict that HMG domain variation will be an accurate marker of the pattern of evolution of the family (Bowles et al., 2000).

MATERIALS AND METHODS

Animals and DNA extraction

Two males and two females of *B. maxima* (Anura: Discoglossidae) obtained from Yunnan province in China were used in this study. Genomic DNA was extracted from fresh muscle tissues according to standard phenol-chloroform procedures (Jifang et al., 2006).

PCR amplification

In accordance with previously described (Zhang et al., 2008), a pair of degenerate oligonucleotide sequences primers was designed and synthesized (*SoxN*: 5'-ATGAAYGCNTTYATGGTNTGG-3' and *SoxR*: 5'-GGNCGRTAYTTRTARTCNGG-3'). To amplify the HMG-box motifs of the *SOX* genes, about 200 ng of genomic DNA was added to a 25 µL reaction mix containing 10 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, 200 µM dNTP, 0.5 µM oligonucleotide primers, and 1 unit *Taq* DNA polymerase. The cycling conditions were 4 min at 95°C, followed by 35 cycles with 40 s at 94°C, 40 s at 52°C, and 1 min at 72°C and finally with a 10 min elongation at 72°C.

Cloning and sequencing

The PCR products were detected by 1.5% agarose gels and the desired PCR products were purified and cloned into the vector pMD18-T (Takara, Dalian, China). Recombined vector was transformed into *Escherichia coli* strain DH5α (Sangon, Shanghai, China) and more than 300 white clones were transferred to a plate of clones from an initial culture plate of lysogeny broth (LB) media containing X-gal and IPTG. 200 positive clones with insert PCR fragment were confirmed using colony PCR. To avoid errors during the PCR amplification and positive clones screening, more positive clones were screened and PCR conditions were improved. The frequency of the occurrence of individual sequences was presented in Table 2 in 200 distinct positive clones.

The distinct positive clones were screened using SSCP (single-strand conformation polymorphism) analysis method (Nie et al., 1999) and sequenced using the universal sequencing primer on an ABI377 auto-sequencer. DNA sequences were analyzed by BLAST, CLUSTAL W and Bioedit programs.

Construction of the molecular phylogenetic tree

41 published *Sox* genes sequences accessible at the NCBI BLAST server (Table 1) were chosen from *Homo spcies*, *Mus musculus*, *Gallus gallus*, *Xenopus laevis*, *Rana tientaiensis, Bufo bufogargarizans Cantor, Danio rerio*, *Takifugu rubripes* to research their evolutionary relationships among the ten *Sox* proteins of *B. maxima* and other vertebrate. These amnio acid sequence data were put into the computer software MEGA3.1 for estimation of molecular evolutionary distances. Using the values of the molecular evolutionary distances, the two molecular phylogenetic trees were constructed with the neighbor-joining (NJ) and minimum-evolution (ME) methods. To assess the robustness of branching, 100 bootstrap replicates were carried out and % support values are marked in.

RESULTS

PCR amplification and clone sequenced

Using genomic DNA as the template, we amplified a 215 bp PCR fragment. 200 distinct positive clones from male and female were screened using SSCP analysis and sequenced. And the ten different HMG domain-encoding sequences had been obtained by sequencing from both male and female toads, with no sexual differences between them. Each of these genes was represented in at least two independent clones (Table 2), making it very unlikely that any of the sequences presented in this paper contain PCR artefacts. GenBank accession numbers of them are: *BmSox1a* EU921549, *BmSox1b* EU921550, *BmSox3a* EU921551, *BmSox3b* EU921552, *BmSox3c* EU921553, *BmSox11a* EU921554, *BmSox11b* EU921555, *BmSox11c* EU921556, *BmSox14* EU921557, *BmSox33* EU921558.

Amino acid sequence analysis of the ten *Sox* genes

The amino acid sequences comparison of 44 *Sox* genes (10 different HMG domain-encoding sequences of *B. maxima* and 34 published HMG domain-encoding *Sox* gene sequences accessible at NCBI BLAST server) was shown (Figure 3). Following database searches and amino acid sequences homology analysis, nine of the ten clones were named as representing: *BmSox1a*, *BmSox1b*, *BmSox3a*, *BmSox3b*, *BmSox3c*, *BmSox11a*, *BmSox11b*, *BmSox11c*, *BmSox14*, their amino acid sequences exhibit 98, 98, 95, 91, 91, 95, 95, 92 and 96% identical to the homologous human HMG domains.

The amino acid sequence of the 10^{th} clone had 79% similarity to the human *SOX4*, 78% similarity to *SOX11*, but it had 94% homology to the *Xenopus laevis Sox-K1* protein which was isolated and sequenced from African clawed frog (*Xenopus laevis*) and named *xSox33*, a novel *SRY*-related gene (Accession number of GenBank: NP_001079045) (Hagiuda et al., 2003). According to the homologous of the amino acid sequence, we named this sequence as *BmSox33*.

Construction of the molecular phylogenetic trees

The relationship between the proteins encoded by the *B. maxima* and the corresponding proteins from other vertebrates were analyzed using neighbour-joining (NJ) and minimum-evolution (ME) methods. The bootstrapped phylogenetic trees are shown in Figure 1a, b, and an unrooted phylogenetic tree for HMG domain of the *SoxC* group analysed by NJ method is shown in Figure 2. From Figure 1, no difference is observed between the two phylogenetic trees using NJ and ME methods, which could conclude that the analysis about these *Sox* genes

Table 1. The *SOX/Sox* genes and accession number of GenBank used in this study from other animal.

Sequence	Accession number of GenBank
A	
homoSRY	AAT37462
B	
homoSOX1	NP-005977
homoSOX2	CAA83435
homoSOX3	CAA50465
homoSOX14	AA106731
homoSOX21	NP-009015
MusSOX1	BAC75667
MusSOX2	NP-035573
MusSOX3	AAH52024
MusSOX14	XP-284529
GalSOX1	BAA25092
GalSOX14	NP-990092
XenSOX1	BAE72677
C	
homoSOX4	NP-003098
homoSOX11	BAA88122
musSOX4	NP-033264
musSOX11	NP-033260
galSOX11	NP-990518
danioSOX4	BC065354
danioSOX11	CAB87379
takSOX4	AAQ18501
takSOX11	AAQ18502
ranaSOX11	AAQ23863
bufoSOX11	ABW90114
xenSOX4	AAG17026
xenSOX11	Q91731
D	
homoSOX5	CAG32994
homoSOX6	AAK26243
musSOX5	BAA32567
musSOX6	CAA09270
E	
homoSOX9	CAA86598
galSOX9	BAA25296
F	
homoSOX7	NP-113627
musSOX7	NP-035576
G	
homoSOX15	NP-008873
musSOX15	NP-0333261
H	
homoSOX30	NP-848511
musSOX30	AAF99391
I	
xenSOX31	BAA32249
J	
ceSOXJ	U51998
Outgroup	
homoTCF-1	CAB56795

Species abbreviations used as gene prefixes: homo, *Homo sapiens*; mus, *Mus musculus*; gal, *Gallus gallus*; danio, *Danio rerio*; tak, *Takifugu rubripes*; xen, *Xenopus laevis*; bufo, *Bufo bufogargarizans Cantor;* rana, *Rana tientaiensis;* ce, *Caenorhabditis elegans*.

Table 2. The frequency of the occurrence of individual sequence.

Gene	Number of positive clones
BmSox1a	2 (♀,1:♂, 1)
BmSox1b	3 (♀,1:♂, 2)
BmSox3a	3 (♀,1:♂, 2)
BmSox3b	4 (♀,2:♂, 2)
BmSox3c	5 (♀,2:♂, 3)
BmSox11a	3 (♀,2:♂, 1)
BmSox11b	2 (♀,1:♂, 1)
BmSox11c	2 (♀,1:♂, 1)
BmSox14	3 (♀,1:♂, 2)
BmSox33	2 (♀,1:♂, 1)

Other fragments were not selected because of low identity or unknown sequence based on amino acid sequences homology analysis, only one clone represented and so on.

is credible. Phylogenetic analysis showed 46 *Sox* genes were clustered into ten subfamilies (A-J), and the nine clones from *B. maxima* (*BmSox1a*, *BmSox1b*, *BmSox3a*, *BmSox3b*, *BmSox3c*, *BmSox11a*, *BmSox11b*, *BmSox11c*, *BmSox14*) belonged to groups B and C. Three members of *SoxB* and *SoxC* groups were obtained from *B. maxima*, which contained three isoforms of *Sox3* (*Sox3a*, *Sox3b* and *Sox3c*) and *Sox11* (*Sox11a*, *Sox11b*, *Sox11c*), two copies of *Sox1* (*Sox1a* and *Sox1b*). Besides, *BmSox33* and *xSox33* HMG box sequence clusters close to group C sequence in phylogenetic analysis (Figures 1, 2). And according to Figure 3, *Sox33* also has C signature amino acid sequences, so *BmSox33* and *xSox33* should belong to group C.

DISCUSSION

Phylogenesis analysis of *Sox* genes

From Figure 1, 46 *Sox* genes were subdivided into ten groups (A-J) based on their amino acid sequences. *B. maxima* contains ten orthologues of two *Sox* groups: B (*Sox1*, *-3*, and-*14*), C (*Sox11*, *-33*). In terms of full-length sequences and functional roles, *Sox* genes of B group (*Sox1*, -2, -3, -14 and -21) can be more correctly separated into two distinct subgroups, B1 and B2 (Koopman et al., 2004; Uchikawa et al., 1999). Our research indicates that this species has duplicates of *Sox1*, *Sox 3*, *Sox11*. But most of these groups are represented by a single gene in the invertebrate model organisms *Drosophila melanogaster* and *Caenorhabditis elegans*, which suggests that expansion of this single gene into multiple related genes occurred during vertebrate evolution and genome expansion (Koopman et al., 2004; Soullier et al., 1999). In addition, *SoxA*, *SoxG* and *SoxH* groups are all represented by single gene (*Sry*, *Sox15* and *Sox30*, respectively) in mammals and have no

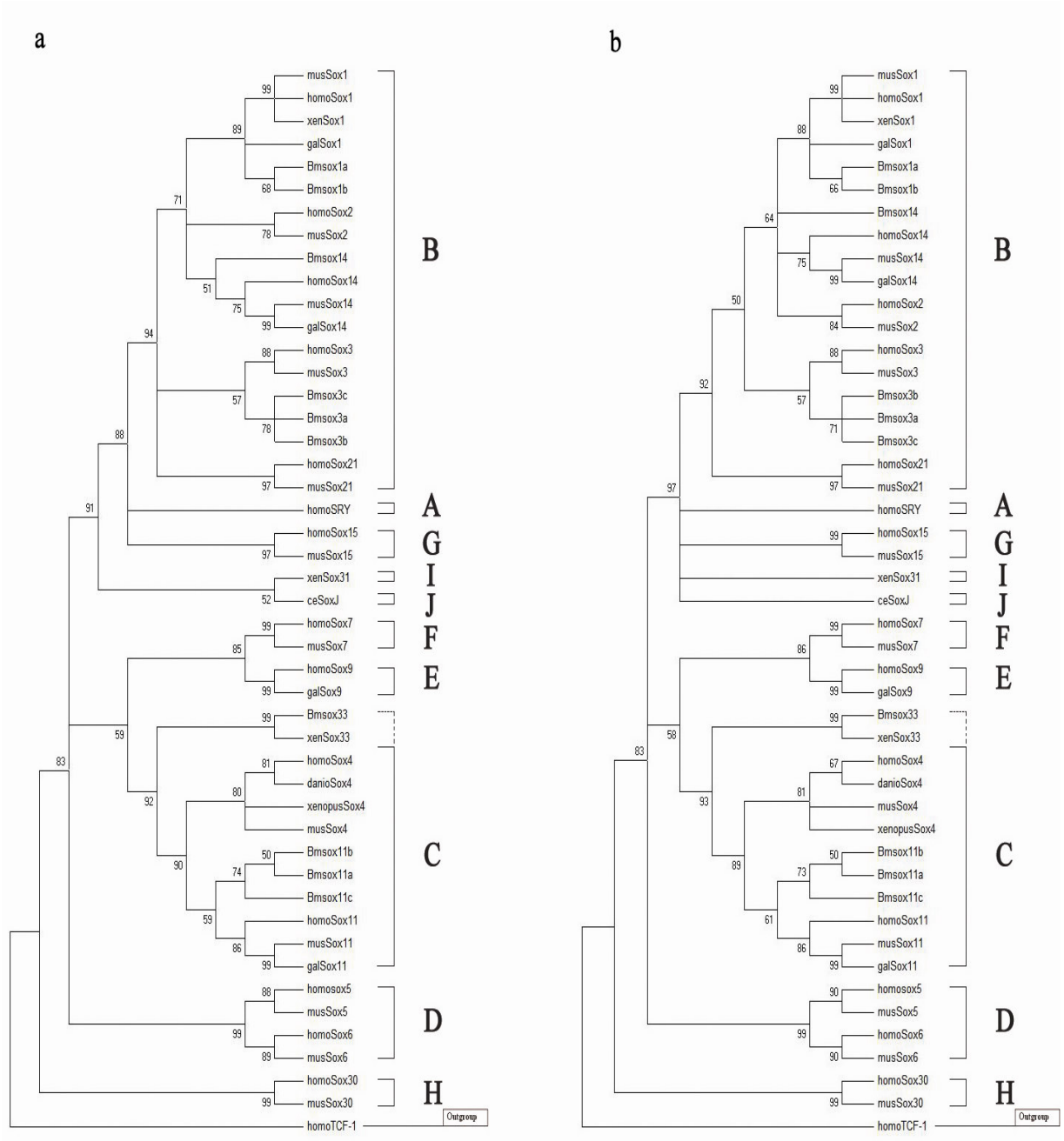

Figure 1. Phylogenetic neighbor-joining (a) and minimum-evolution (b) tree based on the SOX/*Sox* HMG-box amino acid sequences showing the different groups of the *Sox* gene family(A-H). *BmSox33* and *xSox33* clusters close to group C. Human TCF-1 gene was included as outgroup. Numbers show bootstrap values based on 100 replicates. Abbreviations and accession numbers of GenBank are in Table 1.

homologues in *C. elegans* or *Drosophila* (Bowles et al., 2000). In the past, the study addressed specifically the question of the evolution of the mammalian Y-chromosomal *Sox* gene, *Sry*, from the X-linked gene *Sox3* (Kato and Miyata, 1999; Bowles et al., 2000). *Sox15* HMG box sequence is close to group B sequences in phylogenetic analyses (Figure 1), but it lacks group B signature amino acid sequences (Figure 3). *Sox30* sequence is highly diverged from that of any other mammalian *Sox* genes.

Sox33 gene

Among vertebrates, orthologous genes in different species are highly similar to each other in terms of HMG-box

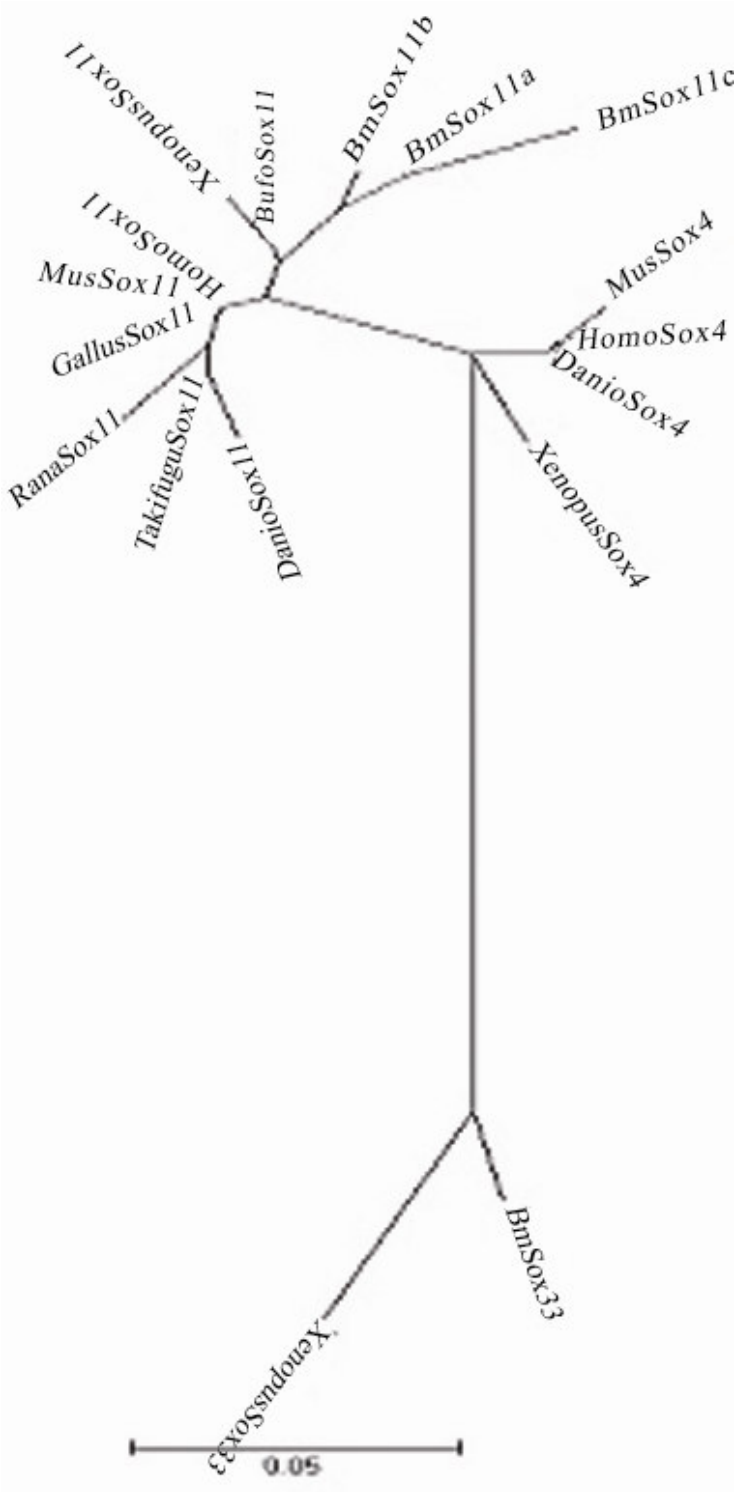

Figure 2. Unrooted phylogeny for the HMG domains of *SoxC* group. The molecular phylogenetic tree was reconstructed with the neighbor-joining (NJ) method. Species abbreviations are as for Table 1.

homology (Bowles et al., 2001). In previous studies, the subgroups of *SOX/Sox* were defined by taking 80% homology as a cut-off value (Hiraoka et al., 1997; Wright et al., 1993). According to amino acid sequences comparison (Figure 3), this analysis is also obvious. Besides, some signature amino acid sequences are detected in

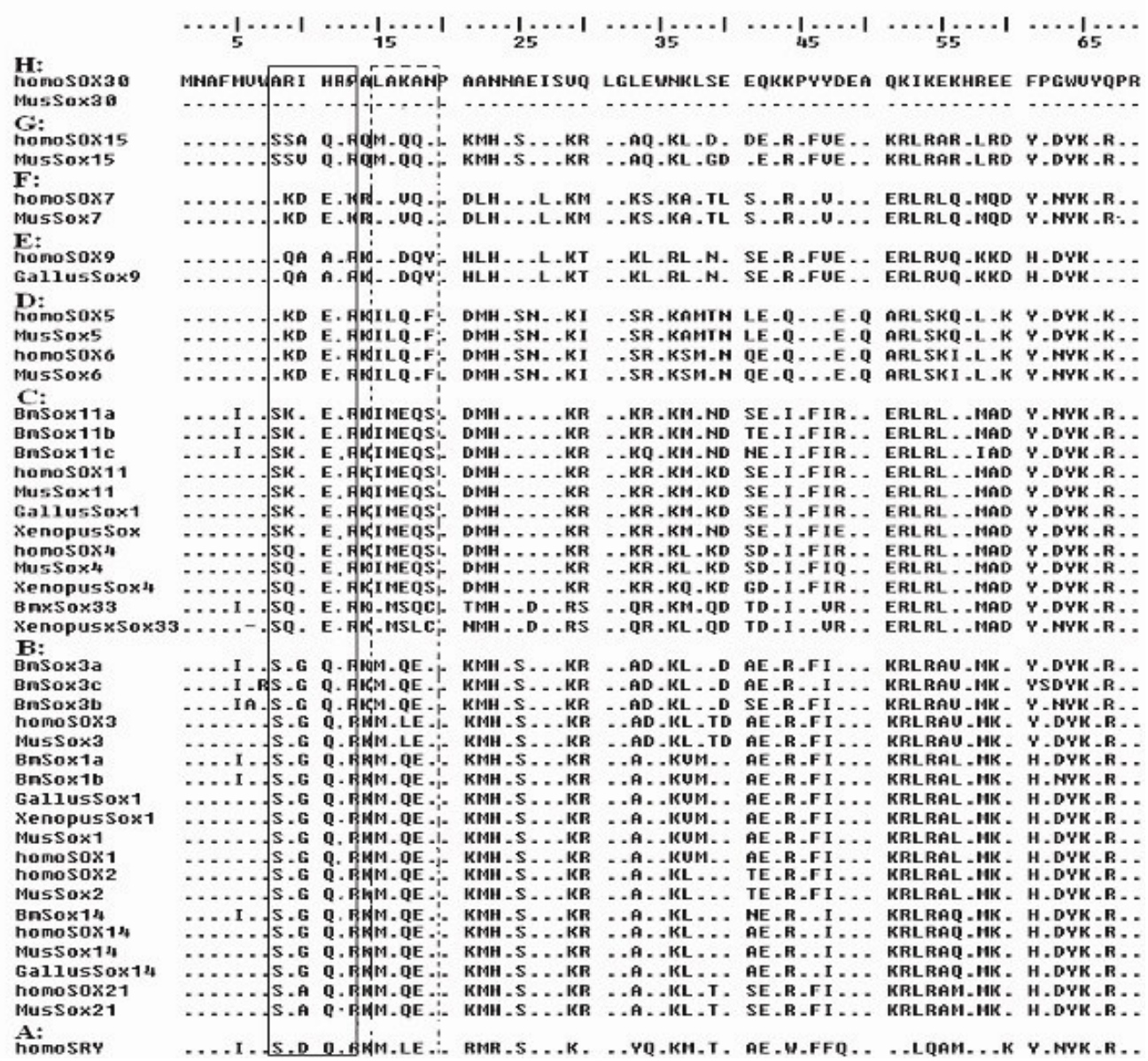

Figure 3. The HMG box amino acid sequences from the *Sox* genes found in this study and known *Sox* proteins identified previously were aligned. Sequences of group I and group J because of few Sox gene represented were not selected. Black dot means amino acid identity between the ten clones and *Sox* genes from vertebrate. Two boxes drawn by straight line and dotted line show two conserve positions in every subgroup. Gene orthology are boxed drawing by straight line and dotted line.

every subgroup. The sequence “SRG/AQRR” for group B, “SK/QIERR” for group C, “AKDERR” for group D were at position 8-13 (Bowles et al., 2001). In addition, on position 15-19 the sequence “MAQE/DN” was in B group, “IMEQS” in C group, “IAQAF” in D group. The protein sequence of *BmSox33* at position 8-13 was “SQIERR” which is similar with the sequence of C subgroup, but the sequence of *BmSox33* and *xSox33* “LMSQ/LC” at position 15-19 was different from group C (Zhang et al., 2008). According to Bowles’s viewpoint, *BmSox33* should belong to C subgroup, which is consistent with the conclusion researched by Hagiuda who studied the full length of cDNA sequence (Hagiuda. et al., 2003).

Sox genes contain a conserved DNA sequence which is responsible for sequence specific DNA binding, and the sequence similarity of the conserved sequence to that of *Sry* is at least 50% (Bowles et al., 2001). The HMG domain sequence MNAF (position1-4) appears to be conserved for all *Sox* proteins. Besides, other conserved positions can be found, like P at position 20, S at position 28, L and G at position 31 and 32.

Until now, *Sox33* gene is not represented in other vertebrates, except Discoglossidae (this paper) and Xenopodidae (Hagiuda et al., 2003) which also belonged

to Salientia of Amphibia. Discoglossidae was the most conservative family in Anura, which could mean that *Sox33* is an ancestral *SoxC* gene in the evolution of *Sox* gene. From Figure 2, it is possible that *Sox33* sequence diverged slightly from the other *SoxC* sequences such as *Sox4* and *Sox11,* but sequences are still similar enough to be classed as a subgroup of group C (79% similarity to the human *SOX4* gene, 78% similarity to human *SOX11* gene). The *Sox33* gene has not been found in other vertebrates up to now, so the origin of *Sox33* gene will be an interesting target in phylogenesis research.

Functions of *SOX* genes

In this paper, three members of these two subgroups were obtained, containing three isoforms of *Sox3*, two isoforms of *Sox1* and *Sox14*. *Sox3*, the closest relative of *Sry*, is located on the X chromosome (Graves, 1998; Katoh and Miyata, 1999). In humans, *SOX3* mutations cause significant gonadal defect, including small testes, in addition to X-linked hypopituitarism, recessive hypoparathyroidism, and mental retardation (Rousseau et al., 1991; Laumonnier et al., 2002). However, in *Xenopus*, *Sox3* acts as a negative regulator of Xnr5, which encodes a nodal-related TGFh-family protein (Weiss et al., 2003). And maternal B1-type *Sox* (*Sox1*, -2, -3) functions together with the VegT/h-catenin system to regulate nodal expression and to establish the normal pattern of germ layer formation (Nitta et al., 2006; Zhang et al., 2004), but their specific functions are not clear in *B. maxim*, which needs us to research in the future.

Genome duplication

According to Prince and Pickett, gene duplication is a mechanism by which new gene functions may be acquired (Prince et al., 2002). A large number of duplicate copies of single copy mammalian genes have been identified in fish such as zebrafish and Fugu, for example in contrast to four clusters of Hox genes in mammals, zebrafish and Fugu contain seven and six Hox clusters, respectively (Amores et al., 1998; Aparicio et al., 2002). Comparative genomics and phylogenetic analysis have indicated that duplication in fish is the result of a large scale segmental duplication, lending to a “fish-specific whole-genome duplication” theory, such as the duplication of *Sox1*, *Sox4*, *Sox9* and *Sox14* in zerbrafish and six of the ancestral vertebrate *Sox* genes (*Sox1*, *Sox6*, *Sox8*, *Sox9*, *Sox10* and *Sox14*) duplicated in *Fugu rubripes* (Burgons et al., 2004; Koopman et al., 2004). Hence *Sox* gene duplication may have occurred relatively early during teleost evolution. This paper indicates that the species may also have duplications of *Sox3*, *Sox11* and *Sox1*, but genomic composition of *B. maxima* is not known. So it is not known whether this is due to segmental duplications or whole genome duplication, which needs further investigation.

ACKNOWLEDGEMENTS

We are grateful to Professor Hao jiasheng for providing the specimen of *Bombina maxima* and the reviewers for numerous valuable suggestions on the manuscript. This research was supported by National Natural Science Foundation of China (No.30770296), the Natural and Science Key Project of Anhui Educational Department (KJ2007A022), and the Key Lab. of Biotic Environment and Ecology Safety in Anhui Province (2006).

REFERENCES

Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang YL, Westerfield M, Ekker M, Postlethwait JH (1998). Zebrafish hox clusters and vertebrate genome evolution. Sci. 282: 1711-1714.

Aparicio S (2002). Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. Sci. 297: 1301-1310.

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

Burgons G.M, Llewellyn L, Mylonas CC, Canario AVM, Zanuy S, Sweeney GE (2004). Analysis of the *Sox* gene family in the European sea bass (*Dicentrarchus labrax*). Comp. Biochem. Physiol. Part B 137: 279-284.

Graves JA (1998). Interactions between SRY and SOX genes in mammalian sex determination. Bioessays. 20: 264-269.

Hagiuda J, Hiraoka Y, Hasegawa M, Ogawa M, Aiso S (2003). A novel Xenopus laevis SRY-related gene, xSox33. Biochem. Biophys. Acta. 1628(2): 140-145.

Hiraoka Y, Ogawa M, Sakai Y, Taniguchi K, Fujii T, Umezawa A, Hata J, Aiso S (1997). Isolation and expression of a human SRY-related cDNA hSOX20, Biochem. Biophys. Acta 1396: 132-137.

Jifang Z, Muyuan Z (2006). Isolation and Sequence Analysis of the *Sox-1,-2,-3* Homologs in *Trionyx sinensis* and *Alligator sinensis* Having Temperature-Dependent Sex Determination. Biochem. Genet. 44: 98-109.

Kato K, Miyata T (1999). A heuristic approach of maximum likelihood method for inferring phylogenetic tree and an application to the mammalian Sox-3 origin of the testis-determining gene SRY. FEBS Lett. 463: 129-132

Koopman P, Schepers G, Brenner S, Venkatesh B (2004). Origin and diversity of the *Sox* transcription factor gene family: genome-wide analysis in *Fugu rubripes*.Genet. 328: 177-186.

Laumonnier F, Ronce N, Hamel BC, Thomas P, Lespinasse J, Raynaud M, Paringaux C, Van Bokhoven H, Kalscheuer V, Fryns JP, Chelly J, Moraine C, Briault S (2002). Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. Am. J. Human Genet. 71: 1450-1455.

Lefebvre V, Dumitriu B, Penzo-Mendez A, Han Y, Pallavi B (2007). Control of cell fate and differentiation by *Sry*-related high-mobility-group box (Sox) transcription factors. Int. J. Biochem. Cell Biol. 39: 2195-2214.

Ner SS (1992). HMGs everywhere. Curr. Biol. 2: 208-210.

Nie LW, Shan XN, Guo CW (1999). The PCR amplification and SSCP analysis of Sox gene in turtles (*Platysternon megacephalum* and *Cistoclemmys flavomarginatus*). China J. Appl. Environ. Biol. 5(4): 378-381.

Nitta KR, Takahashi S, Haramoto Y, Fukuda M, Onuma Y, Asashima M (2006). Expression of *Sox1* during *Xenopus* early embryogenesis. Biochem. Biophys. Res. Commun. 351: 287-293.

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

Phillips NB, Jancso-Radek A, Singh VIR, Chan G, Haas E, Weiss MA (2006). SRY and Human Sex Determination: The Basic Tail of the HMG Box Functions as a Kinetic Clamp to Augment DNA Bending. J. Mol. Biol. 358: 172-192.

Prince VE, Pickett FB (2002). Splitting pairs: the diverging fates of

duplicated genes. Nat. Genet. 3: 827-837.
Rousseau F, Vincent A, Rivella S, Heitz D, Triboli C, Maestrini E, Warren ST, Suthers GK, Goodfellow P, Mandel JL, Toniolo D, Oberle I (1991). Four chromosomal breakpoints and four new probes mark out a 10-cM region encompassing the fragile-X locus (FRAXA). Am. J. Human Genet. 48: 108-116.
Sessa L, Bianchi ME (2007). The evolution of High Mobility Group Box (HMGB) chromatin proteins in multicellular animals. Gene. 387: 133-140.
Soullier S, Jay P, Poulat F, Vanacker J, Berta P, Laudet V (1999). Diversification pattern of the HMG and SOX family members during evolution. J. Mol. Evol. 48: 517-527.
Uchikawa M, Kamachi Y, 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. 84:103-120.
Wegner M (1999). From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res. 27: 1409-1420.
Weiss J, Meeks JJ, Hurley L, Raverot G, Frassetto A, Jameson JL (2003). Sox3 is required for gonadal function, but not sex determination, in males and females. Mol. Cell. Biol. 23: 8084-8091.
Wright EM, Snopek B, Koopman P (1993). Seven new members of the Sox gene family expressed during mouse development, Nucleic Acids Res. 21: 744.
Zhang C, Basta T, Hernandez-Lagunas L, Simpson P, Stemple DL, Artinger KB, Klymkowsky MW (2004). Repression of nodal expression by maternal B1-type SOXs regulates germ layer formation in Xenopus and Zebrafish. Dev. 273: 23-37.
Zhang Y, Nie LW, Yan L, Zheng PP, Chen W (2008). Clone and Sequence Analysis of Sox Genes in *Rana tientaiensis.* Asiat. Herpetol. Res. 11: 151-158.