Fish Aquat Sci 14(1), 43-54, 2011

KOSFAS FISHERIES AND AQUATIC SCIENCES
The Korean Society of Fisheries and Aquatic Science

Characterization of Copper/Zinc-Superoxide Dismutase (Cu/Zn-SOD) Gene from an Endangered Freshwater Fish Species *Hemibarbus mylodon* (Teleostei; Cypriniformes)

Sang Yoon Lee[1], Keun-Yong Kim[2], In-Chul Bang[2] and Yoon Kwon Nam[1]*

[1]*Department of Marine Bio-Materials and Aquaculture, Pukyong National University, Busan 608-737, Korea*

[2]*Department of Marine Biotechnology, Soonchunhyang University, Asan 336-745, Korea*

Gene structure of copper/zinc-superoxide dismutase (Cu/Zn-SOD; *sod1*) was characterized in *Hemibarbus mylodon* (Teleostei; Cypriniformes), an endangered freshwater fish species in Korean peninsula. Full-length cDNA of *H. mylodon SOD1* consisted of a 796 bp open reading frame sequence encoding 154 amino acids, and the deduced polypeptide sequence shared high sequence homology with other orthologs, particularly with regard to metal-coordinating ligands. Genomic structure of the *H. mylodon sod1* gene (*hmsod1*; 1,911 bp from the ATG start codon to the stop codon) was typical quinquepartite (*i.e.,* five exons interrupted by four introns); the lengths of the exons were similar among species belonging to various taxonomic positions. The molecular phylogeny inferred from *sod1* genes in the teleost lineage was in accordance with the conventional taxonomic assumptions. 5'-flanking upstream region of *hmsod1,* obtained using the genome walking method, contained typical *TATA* and *CAAT* boxes. It also showed various transcription factor binding motifs that may be potentially involved in stress/immune response (*e.g.,* sites for activating proteins or nuclear factor kappa B) or metabolism of xenobiotic compounds (*e.g.,* xenobiotic response element; *XRE*). The *hmsod1* transcripts were ubiquitously detected among tissues, with the liver and spleen showing the highest and lowest expression, respectively. An experimental challenge with *Edwardsiella tarda* revealed significant up-regulation of the *hmsod1* in kidney (4.3-fold) and spleen (3.1-fold), based on a real-time RT-PCR assay. Information on the molecular characteristics of this key antioxidant enzyme gene could be a useful basis for a biomarker-based assay to understand cellular stresses in this endangered fish species.

Key words: Gene and promoter structure, *Hemibarbus mylodon*, Immune challenge, *Sod1*

Introduction

Generation of reactive oxygen species (ROS) is an unavoidable process in most aerobic organisms. Although ROS are important in host defenses against invading microorganisms as well as in certain cellular signaling events, excessive ROS cause oxidative stress, accompanied by significant damage to cellular macromolecules including lipids, proteins, and DNA (Zelko et al., 2002). Of variety of enzymes involved in the antioxidant pathway in aerobic organisms, copper/zinc-superoxide dismutase (Cu/Zn-SOD; SOD1; EC 1.15.1.1) plays pivotal roles in the first line of defense against oxidative stress. This enzyme primarily scavenges cytosolic ROS via the dismutation of the ROS to hydrogen peroxide, which is then converted to water by the catalytic action of catalase (EC 1.11.1.6) (Zelko et al., 2002). Given that in aquatic animals, many stimulatory factors often enhance the formation of ROS in a variety of organs, SOD has been proposed as a versatile biomarker that can be used to detect environmental problems as well as health risks for animals (Almeida et al., 2002; Regoli et al., 2002; Hansen et al., 2006).

Genetic determinants of Cu/Zn-SODs have been isolated from a variety of teleost species; however, most previous studies focused on coding sequences at the mRNA level (cDNAs), and there is little information on this gene at the genomic level in teleosts. This lack of knowledge regarding the

*Corresponding author: yoonknam@pknu.ac.kr

regulatory region of the fish *sod1* gene may be an obstacle in the pursuit of getting a deeper insight into the molecular mechanism of *sod1* regulation in response to biotic and abiotic stressors.

Hemibarbus mylodon (Teleostei; Cypriniformes) is an endemic freshwater species found along the Korean peninsula. Over the last decade, the habitats of this species have been largely destroyed by anthropogenic and/or industrial activities, raising serious concern about the extinction of wild populations (Jang et al., 2003; Kim et al., 2008). Recently, the Korean Ministry of Environment launched a research program for the conservation of endangered fishes, including *H. mylodon*, and studies addressing the biological and ecological issues involved in successful restoration are underway (Ministry of Environment, Korea; http://nre.me.go.kr/eweb/main/index.jsp). The development of fine molecular biomarkers for a threatened species can provide an improved understanding of stress phyiology and adaptive capacity at both the cellular and organismal levels, which could be of importance in the development of conservation strategies (Snape et al., 2004; Miracle and Ankley, 2005; Cho et al., 2008).

Consistent with our long-term goal to provide comprehensive information on a variety of molecular biomarkers that could aid in *in situ* and *ex situ* restoration activities for *H. mylodon*, the objective of this study was to cast knowledge on the molecular structure and mRNA expression of *sod1,* a key antioxidant enzyme gene from *H. mylodon*. Thus, we isolated and characterized the cDNA and genomic sequences of *H. mylodon sod1*; determined the molecular features of the 5'-flanking regulatory regions by genome walking method; and examined basal and induced mRNA expression upon immune challenge.

Materials and Methods

Isolation of full-length cDNA and genomic gene

We conducted a survey of our *H. mylodon* expressed sequence tag (EST) database (Bang et al., 2007), which revealed a clone (HMM-152) obtained from the muscle cDNA library that matched preiously determined vertebrate *sod1* sequences. This EST clone appeared to contain potentially the full-length *sod1* cDNA sequence, including the 5'-untranslated region (UTR) and the ATG start codon. The insert DNA was further sequenced in both directions using the primer walking method. The *sod1* open reading frame (ORF) fragment was amplified from liver total RNA (500 ng) by RT-PCR amplification using the primers HM SOD1 FW and HM SOD1 RV.

Information regarding the primers and thermal cycling conditions used in this study is provided in Table 1. RT-PCR was carried out using the *AccuPower* RT-PCR Premix (Bioneer, Daejeon, Korea) according to the manufacturer's recommendation. Purified RT-PCR product was cloned into the pGEM®-T Easy Vector (Promega, Madison, WI, USA), and the insert DNA was sequenced from six randomly chosen recombinant clones.

Genomic gene sequence of *H. mylodon* Cu/Zn-SOD (*hmsod1*), containing a fragment spanning the start codon to the stop codon, was isolated by PCR amplification. Genomic DNA was prepared from fin tissue using a conventional SDS/proteinase K method (Cho et al., 2008), and 100 ng of purified genomic DNA was subjected to PCR amplification using the Fidelity^{Plus} PCR System (Roche Applied Science, Mannheim, Germany). The HM SOD1 FW and HM SOD1 RV primers were used, and PCR product was gel-purified using the *AccuPrep* Gel Elution Kit (Bioneer), TA-cloned, and sequenced ($n=6$) as described above.

Genome walking to the 5'-flanking region

Genome walking was performed using the GenomeWalker™ Universal Kit (BD Biosciences Clontech, Mountain View, CA, USA) to obtain the sequence of the 5'-upstream regulatory region of *hmsod1*. Briefly, genomic DNA (2.5 μg) was digested with a restriction endonuclease (*Eco*RV, *Dra*I, *Pvu*II, or *Stu*I) and ligated with the adaptors according to the manufacturer's instructions. First and second rounds of PCR amplifications were performed using two primer pairs, AP1 (adaptor forward primer provided in the kit; BD Biosciences Clontech) with HM SOD1-GW1, and AP2 with HM SOD1-GW2. Amplified products were TA-cloned and sequenced to generate a contiguous sequence. This 5'-upstream sequence was confirmed by sequence analysis of the PCR clones ($n=5$) amplified directly from the genomic DNA using HM SOD1-pFW and HM SOD1-pRV primers. The sequence data have been submitted to the GenBank databases under accession numbers FJ975147 (cDNA) and FJ975148 (genomic sequence).

Sequence characterization and bioinformatic analysis

cDNA sequence of *hmsod1* was subjected to

Table 1. List of oligonucleotide primers used in this study

Name	Sequence (5′-3′)	Thermal cycling conditions	Purpose
HM SOD1 FW HM SOD1 RV	CTCGCTTATTAACGCTGTCGGT CACTTTGTAGTTGTGAGAAGGCAG	Reverse transcription at 42°C for 60 min followed by an inactivation step (94°C for 4 min) and then 30 cycles at 94°C for 30 s 58°C for 30 s and 72°C for 45 s (for cDNA) while 30 cycles at 94°C for 45 s 58°C for 1 min and 68°C for 2.5 min followed by a final elongation at 72°C for 5 min (for gDNA)	Isolation of *sod1* cDNA and genomic gene
HM SOD1-GW1 HM SOD1-GW2	CAATGTGCGTGGTTGAGGGCTAGCTCG AAAACAGTTCCGGTCACTTCCCCTGT	7 (the first PCR) or 5 (nest PCR) cycles at 94°C for 24 s and 68°C for 3 min, followed by 32 (the first PCR) or 20 (nest PCR) cycles at 94°C for 24 s and 72°C for 3 min, followed by a final elongation at 72°C for 7 min	Genome walking to 5'-flanking region of *sod1*
HM SOD1-pFW HM SOD1-pRV	AACAAGACCATATGAGCTTCTCCCATG AACAGTTCCGGTCACTTCCCCTGT	30 cycles at 94°C for 45 s, 60°C for 1 min and 72°C for 2.5 min	Isolation of continuous 5'-flanking region of *sod1* for sequence confirmation
Fi18S rRNA 1R	CAAGAATTTCACCTCTAGCGGC	Reverse transcription at 37°C for 60 min	Preparation of normalization control
qHM SOD1 1F qHM SOD1 1R	CGCACTTCAACCCTCACAAT CTCTACCCAGTGATGCCAAT	26 (for end-point PCR) or 45 (for real-time PCR) cycles at 94°C for 20 s, 58°C for 20 s and 72°C for 20 s	Semi-quantitative or real-time RT-PCR of *sod1*
q18S rRNA 1F q18S rRNA 1R	AGAAACGGCTACCACATCCA GGACACCCAGTTAAGGGCAT	20 (for end-point PCR) or 40 (for real-time PCR) cycles at 94°C for 20 s, 58°C for 15 s and 72°C for 20 s	Semi-quantitative or real-time RT-PCR of 18S rRNA

NCBI BLASTx search (http://www.ncbi.nlm.nih.gov/blast) to examine homology with previously known vertebrate *sod1* orthologs. The full-length cDNA sequence was also analyzed to identify putative ORFs using the ORF Finder in the NCBI web page (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Predicted molecular mass and theoretical pI values were calculated using the ProtParam tool (http://www.expasy.org/tools/protparam.html). Multiple sequence alignments at the nucleotide and amino acid levels were performed using ClustalW (http://align.genome.jp/). To compare the genomic structure and organization of *hmsod1* with representative orthologs, we deciphered animal *sod1* genes from the GenBank and Ensembl databases (http://www.ensembl.org/index.html) using text and BLAST searches.

Putative transcription factor binding sites in the 5'-flanking upstream region were identified using TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) and Transcription Element Search System (TESS; http://www.cbil.upenn.edu/cgi-bin/tess/tess).

Molecular phylogeny

Nucleotide sequences of the *sod1* genes for representative teleosts were retrieved from GenBank and Ensembl and compiled. An abalone *Haliotis discus discus* (Mollusca; Gastropoda) sequence was used as an outgroup. The sequence data were aligned using ClustalW (Thompson et al., 1994) with the default setting for gap penalties in BioEdit 7.0.5 (Hall, 1999). The aligned nucleotide sequence data matrix was subjected to maximum likelihood (ML) analysis in PAUP* 4.0b10 (Swofford, 2002). Model selection strategy using the Akaike Information Criterion (AIC) implemented in Modeltest 3.7 (Posada and Crandall, 1998) was used to determine the best-fit evolutionary model of nucleotide substitutions. ML tree was reconstructed with the GTR+I+Γ model using the likelihood settings, determined from Modeltest. ML analysis was performed using the heuristic search option with random addition of sequences (10 replicates) and tree-bisection-reconnection branch swapping. Robustness of tree topologies was evaluated by bootstrap analysis with

1,000 pseudoreplicates. Bayesian inference (BI) analysis was carried out using MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). A model selection strategy of AIC implemented in MrModeltest 2.3 (Nylander, 2004) was used to determine the best-fit evolutionary model, and the GTR+I+Γ model was selected. Two independent Markov chains were performed with four simultaneous chains (three heated and one cold) with random starting trees for 1,000,000 generations, sampling trees at intervals of 10 generations. A total of 1000 of 10,001 resulting trees were discarded as "burn-in." The remaining trees were used to construct a 50% majority-rule consensus tree and to estimate statistical supports for tree topologies determined on the basis of posterior probabilities.

Tissue distribution assay of *sod1* transcripts

To examine the tissue distribution and basal expression of *hmsod1* transcripts, semi-quantitative RT-PCR assay was performed. Total RNA was extracted from eight somatic tissues including the brain, gill, heart, intestine, kidney, liver, muscle and spleen. Tissue samples were obtained from 16 healthy individuals (16.5±2.0 g body weight) and pooled within each tissue type prior to RNA extraction. Preparation of total RNA was carried out using the RNeasy Midi Kit (Qiagen, Hilden, Germany) including the DNase I treatment step according to the manufacturer's recommendations. The integrity and purity of extracted total RNA were checked using 28S : 18S rRNA ratio in ethidium bromide (EtBr)-stained gels and/or spectrophotometry. A 2 μg of total RNA in a reaction volume of 20 μL was reverse-transcribed at 37°C for 60 min using the Omniscript RTase (Qiagen) and the oligo $d(T)_{20}$ primer (1 μM final concentration) according to the manufacturer's protocol. To prepare a normalization control, a conserved reverse primer for fish 18S rRNA, Fi18S rRNA 1R primer (Cho et al., 2008), was also included in the reverse transcription (RT) reaction at 0.05 μM. When the RT reaction was finished, the cDNA was diluted two-fold (for *sod1*) or eight-fold (for 18S rRNA) with sterile water and then 1 μL of the diluted sample was subjected to PCR amplification using the *AccuPower* PCR Premix (Bioneer). PCR primers for *sod1* (qHM SOD1 1F and qHM SOD1 1R) and 18S rRNA (q18S rRNA 1F and q18S rRNA 1R) were designed to amplify the 280- and 338 bp amplicons, respectively. Triplicate amplifications were carried out in independently. The amplified product was separated on agarose gel (1.5%) and visualized by EtBr staining. EtBr-stained bands for *sod1* were normalized against 18S rRNA bands using the image analysis software, Quantity-One™ implemented in VersaDoc 4000 (Bio-Rad, Hercules, CA, USA).

Bacterial challenge and real-time RT-PCR assays

To examine the effect of bacterial challenge on the transcriptional modulation of *hmsod1*, an experimental injection of pathogenic bacteria, *Edwardsiella tarda* Gram (-) (FSW910410) was performed. Preparation of bacterial culture was carried out as described previously (Cho et al., 2009a). A freshly grown *E. tarda* (5×10^3 cells) was resuspended in 100 μL of 0.85% NaCl solution. *H. mylodon* juveniles (13.2±2.5 g body weight; $n=6$) were given an intraperitoneal injection of the bacterial suspension. Similarly, a control group ($n=6$) was also prepared by injecting 0.85% NaCl solution only. Fish were transferred to cages within a tank containing 100 L of 5 μm-filtered tap water. The water temperature was held at 24±1°C and dissolved oxygen was 6±1 ppm. Fish were not fed after injection. Two days post-injection, the kidney and spleen tissues were surgically removed from each individual and pooled within tissues prior to total RNA extraction. Reverse transcription was performed using the total RNA samples from the kidney and spleen as described above. A 1 μL aliquot of the diluted cDNA template (two-fold for *sod1* and eight-fold for 18S rRNA) was subjected to real-time PCR amplification using 2× SYBR Green Supermix (Bio-Rad) and iCycler Real-Time Detection Module (Bio-Rad) according to the manufacturer's protocol. The oligonucleotide primers were the same as those used for the end-point RT-PCR. Fluorescence was measured after each elongation step to estimate the average threshold cycle number (CT). The dissociation curve was examined after the final elongation to confirm the specificity of the PCR. Based on the standard curves, *sod1* transcript levels were normalized against the 18S rRNA levels, and the relative expression of Mn-SOD transcripts in the bacteria-challenged group was expressed as a fold change relative to the saline-injected control using the comparative CT method described by Schmittgen and Livak (2008). Assays were carried out in triplicate per cDNA sample in an independent fashion.

Statistical analyses

The relative levels of basal expression of *hmsod1* transcripts across different tissues were assessed by ANOVA followed by Tukey's *post-hoc* test at a confidence level of 95% to ascertain whether any differences between the tissue groups were

statistically significant. The difference in *hmsod1* expression between bacteria- and saline-injected groups was tested by Student's *t*-test using the SPSS software (ver. 10.1.3). Differences were considered significant when $P<0.05$.

Results and Discussion

Sequence characteristics of cDNA and deduced amino acids

H. mylodon sod1 cDNA (796 bp) consisted of a 68 bp 5'-UTR, a 462-bp ORF encoding 154 amino acids, a 266-bp 3'-UTR including the stop codon and 20 bp poly (A+) tail (see GenBank accession number FJ975147). A putative polyadenylation signal (AATAAA) was found 16 bp prior to the poly (A+) tail. The deduced amino acid sequence contained putative Cu ligands (His^{47}, His^{49}, His^{64} and His^{124}; positions in Fig. 1) and Zn ligands (His^{64}, His^{72}, His^{84} and Asp^{87}), as well as two Cys residues (Cys^{58} and Cys^{150}) that potentially form a disulfide bridge (Fig. 1). Calculated molecular mass was 15755.64 Da, and the theoretical pI was 6.28. *H. mylodon* SOD1 exhibited considerable similarity in amino acid sequence to previously characterized teleost SOD1s, with the highest similarity of 84.4% to zebrafish SOD1 (Fig. 1). Common features, particularly with respect to the metal-coordinating residues, were clearly conserved in all of the species examined. Due to the essential role of the SOD1, the structure and function are remarkably conserved across diverse organisms (Fukuhara et al., 2002; Zelko et al., 2002). The two Cys residues (Cys^{58} and Cys^{150}; Fig. 1), forming a disulfide bridge, have been reported to play an important role in stabilizing the electrostatic loop located at the C-terminal region of the SOD1 enzyme (Ciriolo et al., 2001; Zelck et al., 2005). The clustering of charged residues (i.e. Glu^{136}, Glu^{137}, Lys^{140} and Arg^{147} in Fig. 1) within the electrostatic loop region is important for the electrostatic attraction of substrates toward the active site (Ciriolo et al., 2001).

Genomic structure

A single band for the *sod1* fragment from *H. mylodon* genome was successfully produced by PCR. The amplified fragment for *hmsod1* (2,171 bp) contained a 1,911 bp genomic sequence from the start codon to the stop codon, in which five exons (72, 97, 70, 118 and 108 bp for exons I to V, respectively) were interrupted by four introns (875, 303,177 and 91 bp for introns I to IV, respectively). The coding sequences of the genomic clearly matched the cDNA sequence. The consensus GT/AG rule was conserved at all exon/intron boundary regions. The quinque-partite organization (five exons interrupted by four introns) is a typical structure for most known animal *sod1* genes (Fig. 2). Despite the relative abundance of

```
                                  20                  40                  *60                  80
H. mylodon      MAKKAVCVLK GTGEVTGTVF FEQETDGSPV KLSGTISGLT AGKHGFHVHV FGDNTNGCIS AGPHFNPHNK NHGGPTDGD-
D. rerio        .VN..... .........Y .N..GEKK.. .VT.E.T... P.........A ..........  ........D. T......SV-
O. mykiss       ..M..... .......... ....GADG.. ..I.E....A P.E......A Y.......M.  .........Q T......AV-
S. salar        ..L..... .......... ....G..A.. ..T.E.A... P.E......A ........M.  .........H T......TV-
T. rubripes     ..M..... .A.DTS...Y ....NESA.. ..T.E.K... P.E......A ..........  ....Y..... T.A....A.-
T. nigroviridis .VI..... .A..TS...Y ...QDEKA.. ..T.E.K... ..E......A ..........  ....Y..... T.A..N.ENS
O. fasciatus    .VL..... .A..T....Y ....S.SA.. ..T.E.K... P.E......A ..........  .......... ..A..N.AE-
P. major        .VQ..... .A..T..V.H ....SESA.. T.K.E..... PDE......A ..........  .......... ..A....AE-
A. schlegeli    .VL..... .A..T..V.H ....SESA.. ..T.E.K... P.E......A ..........  ....L...G. .......EE-
E. coioides     .DL..... .A..TS...Y ....S.SA.. ..T.E.K... P.E......A ..........  .......... Q.A....A.-
E. malabaricus  .VL..... .A..TS...Y ......SA.. ..T.E.K... P.E...Q..A ..........  .......... H.A....AE-
O. mossambicus  .VL..... ...DTS...Y ....N.SA.. ..T.E.K... P.E......A ..........  .......Y.. .....K.AE-
G. aculeatus    .VV..... .A..T...IY ....S.KAA. ..T.E.K... P.E......A ..........  ........S. T.A..N.EI-
O. latipes      .VL..... ....TN.V.N ....S.SA.. .VT.E.K... P......I.. Y.......V.  .......Y.. .....E.AE-

                                100                 120                 140         *     157
H. mylodon      --RHVGDLGN VTAGESGVAK IDIVDKMLTL SGQHSIIGRT MVIHEKEDDL GKGGNEESLK TGNAGGRLAC GVIGITG
D. rerio        --........ ...DA..... .E.E.A.... .......... .......... .......... .......... .....Q (84.4%)
O. mykiss       --........ ....ADN... .N.Q...... T.PD..... ......A... .......... .......Q.. .....AQ (81.0%)
S. salar        --........ ...AADS... .N.Q.EI.S. A.P..... ......A... ...D....R. .....S.... .....AQ (79.8%)
T. rubripes     --..L..... ....ADNI.. ...K.S.... T.PY..... ......A... .......... .......... .....Q (79.2%)
T. nigroviridis LK........ ...EADQI.. ...T.SVIS. H.KF..... ......A... .......... .......... .....Q (74.5%)
O. fasciatus    --........ ....ADN... ...K.HII.. T.PD..... ......A... .......... .......... .....Q (80.5%)
P. major        --........ ....ADN... ...T..... N.PF..... ......A... .......... .......... .....CQ (80.5%)
A. schlegeli    --........ ....ADN... ...T..... T.PL..... ......T... .......... .......... .....Q (81.1%)
E. coioides     --........ ....GDN... ...T..... N.PY..... ......A... .R...D.... .......... .....AQ (80.5%)
E. malabaricus  --........ ....GDN... ...T..II.. N.PY..... ......A... .T...... .......... .....SQ (79.2%)
O. mossambicus  --........ ....ADN... .E.T..VI.. T.RD..... ......V... X..... .......... .....Q (79.2%)
G. aculeatus    --........ ....GDNI.. ...T..VI.. T..... ......A... .......... .....S.... .....AQ (77.9%)
O. latipes      --........ ....DNN... ...T..LIR. ..PD..V... V.V...V... .....D.... .....A.... .....AQ (75.3%)
```

Fig. 1. Alignment of *Hemibarbus mylodon* SOD1 polypeptide sequence along with its corresponding orthologs. The percent identity with *H. mylodon* SOD1 is noted at the end of each sequence. Metal-coordinating residues are noted by boldface letters and two Cys residues forming a disulfide bond are indicated by asterisks. GenBank accession number or Ensembl ID for each sequence can be referred to phylogenetic tree provided in Fig. 4.

cDNA information on fish *sod1*, the genomic sequence of the teleost *sod1* gene has not been fully determined. We also deciphered several preliminary *sod1* gene sequences from publically available fish genome databases. We found that all of the fish species examined, with the exception of *Gasterosteus aculeatus*, which possesses only four exons, showed highly conserved quinquepartite genomic organization of the *sod1* genes (Fig. 2). Exon sizes were as follows: 72-75 bp for exon I, 91-99 bp for exon II, 66-70 bp for exon III, 115-120 bp for exon IV, and 108-111 bp for exon V (also refer to GenBank accession numbers or Ensembl IDs provided in Fig. 2). Thus neither sizes of each exon in the *hmsod1* gene nor the total exon length differed significantly from those of other animal species. In contrast, intron lengths were quite variable depending upon species, with human *sod1* exhibiting the largest introns (8,344 bp in total). Consequently, the genomic length from the ATG start codon to the stop codon of *sod1* was related, but only roughly so, with the genome size (*C* value) of the animal species.

Molecular phylogeny

Phylogenetic tree was reconstructed based on *sod1* sequence data matrix (Fig. 4). In ML and BI trees, teleost fishes were divided into two major clades; *Danio rerio* and *H. mylodon* belonging to the order Cypriniformes in the Ostariophysi formed a monophyletic group and were phylogenetically separated from the other teleosts, which formed another monophyletic group. The overall structure of the phylogenetic relationship among fishes was similar to previous relationships established using *sod1* sequences (either nucleotide or amino acid sequences) in the vertebrate lineage (Kim and Nam, 2006). Within the latter clade, *Oncorhynchus mykiss* and *Salmo salar* belonging to the Salmoniformes in the Protacanthopterygii, and the other teleosts belonging to the Tetraodontiformes, Perciformes or Beloniformes in the Percomorpha were subdivided into two monophyletic groups. Within the percomorph clade, tetraodontiform *T. nigroviridis* came up first and was followed by a monophyly of perciform and beloniform fishes in the BI analysis, although this relationship was not supported statistically in the ML analysis. Overall, the phylogenetic assumption inferred from the fish *sod1* gene was in accordance with the conventional higher-level teleostean relation-ships estimated using the complete mitogenome sequences of Miya et al. (2001), which showed the sequential ramification of Ostariophysi-Protaca- nthopterygii-Percomorpha.

Characteristics of the 5'-flanking regulatory region

A 1,981 bp 5'-flanking upstream fragment was obtained by the genome walking and putative transcription factor binding sites were predicted (Fig. 3). The upstream region of *hmsod1* exhibited a *TATA* box (TATAAA) at a relatively distant position (742 bp before the start codon) and three *CAAT* boxes (CCAAT; two copies at the proximal region and one at the distal region). In addition to these universal

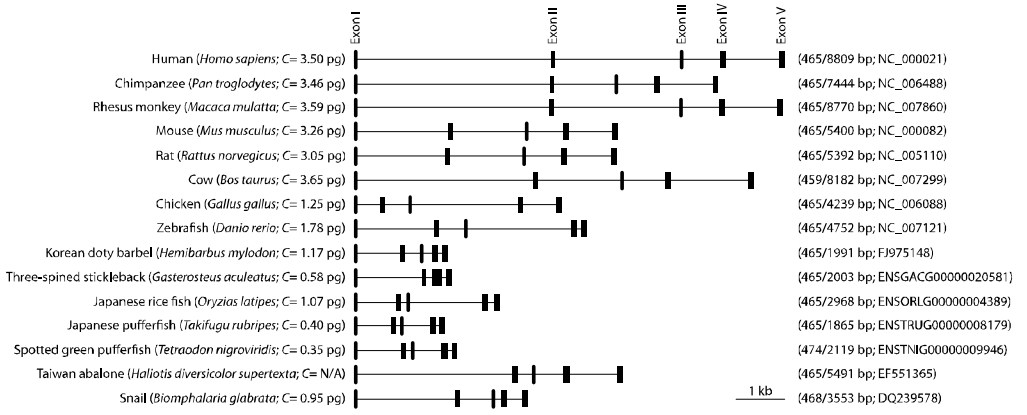

Fig. 2. Schematic drawing to compare the genomic structure of *Hemibarbus mylodon sod1* with representative orthologs from other animal species. Exons are indicated by vertical boxes, while non-translated sequences by horizontal lines. Numbers in each parenthesis at the right are the total length of exons/genomic gene from the ATG start codon to the stop codon and GenBank accession number (or Ensembl ID). Haploid genome size (*C* value=1.17 pg) of the *H. mylodon* is referred to Bang et al. (2008), while those of other animals are referred to Animal Genome Size Database (http://www.genomesize.com/). N/A=data not available.

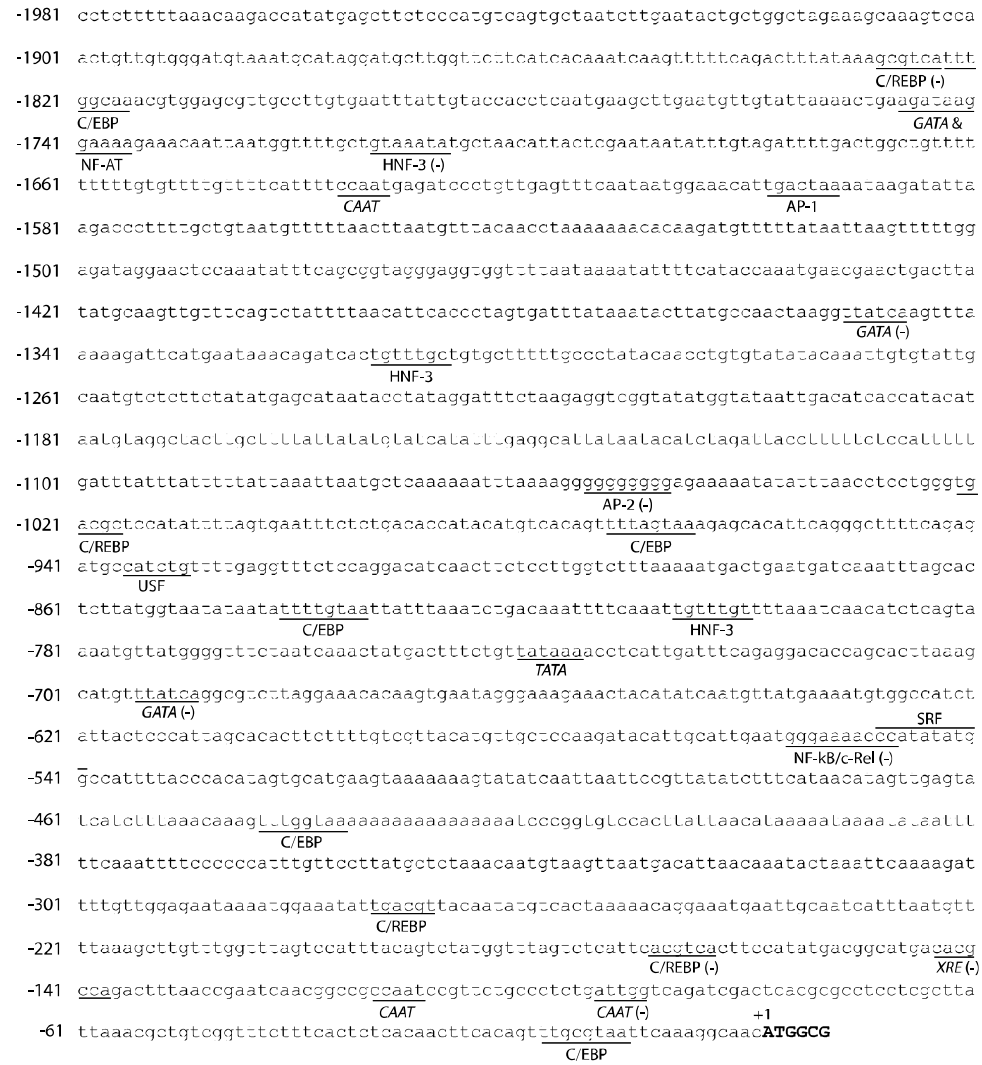

Fig. 3. Nucleotide sequence of 5'-flanking upstream region of *Hemibarbus mylodon sod1*. Putative transcription factor binding motifs are underlined based on the prediction using TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) and TESS (http://www.cbil.upenn.edu/cgi-bin/tess/tess) analysis. The reverse orientation of the transcription factor binding motif is noted by (-).

transcription factor binding elements, the *hmsod1* regulatory region represented a variety of binding motifs for transcription factors involved in stress- and immune-responses. These included cAMP-response element binding protein (C/REBP; consensus binding sequence=TGACGY), CAAT-enhancer binding protein (C/EBP; TTDNGNAA), nuclear factor for activated T-cells (NF-AT; WGGAAA), activating protein-1 (AP-1; TGASTMA), AP-2 (CCCMNSSS), upstream stimulatory factor (USF; CANNTG), nuclear factor kappa B (NF-κB; GGGRNNYYCC) and serum response factor [SRF; also known as a *CArG* box binding protein; $CC(A/T)_6GG$]. The *hmsod1* promoter also contained a copy of xenobiotic response element (*XRE*; TNGCGTG) in the proximal region (145 bp from the start codon in a reverse orientation). Although the specific roles of these factors in the orchestration of the *sod1* gene expression remain to be explored in teleosts, many of them are known to be involved in the transcriptional regulation of mammalian *sod1*s under stimulatory or inflammatory conditions (Rojo et al., 2004; Afonso et al., 2006). Thus, extensive examinations of *hmsod1* expression in response to designed stress treatments or immune challenges are needed to better hypothesize the specific or coordinated roles of these transcription factors. Conversely, the *XRE* is the target of a protein complex including a ligand-

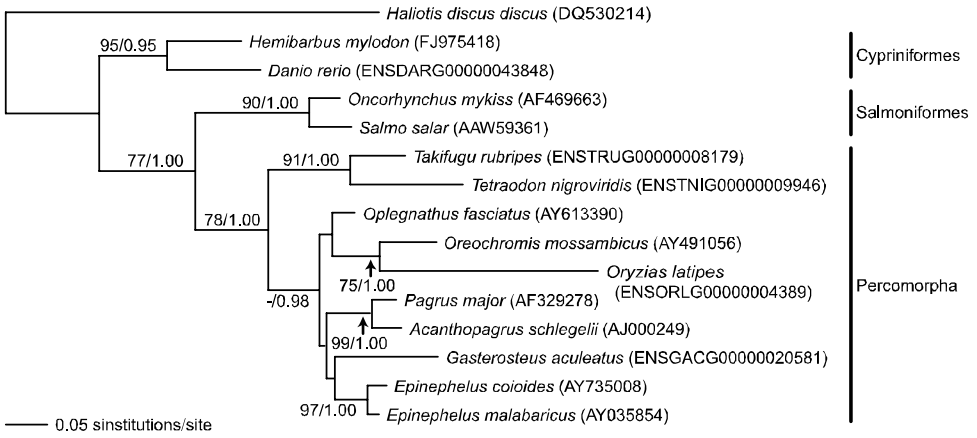

Fig. 4. Maximum likelihood (ML) tree inferred from nucleotide sequences of fish *sod1* genes. Molluscan *Haliotis discus discus* was used as an outgroup. Numbers at each branch node indicate bootstrap value of ML analysis above 50% and posterior probability of Bayesian analysis above 0.90. Taxonomic placements according to the conventional classification scheme of the Teleostei were indicated for each taxonomic group.

activated aryl hydrocarbon receptor (AhR), widely known to be responsible for the recognition of exogenous environmental pollutants in eukaryotic cells (Ryu et al., 1996). *XRE* is usually found in the promoter region of xenobiotics-metabolizing battery genes, such as cytochrome P450 (Ryu et al., 1996), glutathione S-transferase (Pimental et al., 1993), and UDP-glucuronosyltransferase 1A1 (Yueh et al., 2003), as well as mammalian *sod1* (Yoo et al., 1999a; Park and Rho, 2002). The identification of *XRE* in this study is also in agreement with previous observations on the differential modulation of *sod1* in other fish species exposed to xenobiotic compounds (Pedrajas et al., 1998; Shi et al., 2005; Vega-López et al., 2007). However unlike murine *sod1* (Yoo et al., 1999b), the *hmsod1* did not possess a metal-responsive element (*MRE*) in its promoter region, although heavy metal exposure could stimulate *sod1* transcription in this species as well (see Cho et al., 2009b). This suggests that the induced expression of *hmsod1* during heavy metal exposure might be related to the general defense mechanism against prooxidative forces arising from excess heavy metal ions rather than a direct MRE-mediated pathway (however see, Yoo et al., 1999b).

Tissue distribution and basal mRNA expression

RT-PCR assay detected ubiquitous expression of *sod1* transcripts in all the tissues examined, with various basal expression levels among tissues (Fig. 5). Of the eight tissues examined, the liver displayed the highest expression of *sod1* transcripts, and followed by the heart and muscle ($P<0.05$). The brain and intestine exhibited moderate levels of *sod1* expression, while gill and kidney tissues expressed only weak level expression levels and spleen exhibited the

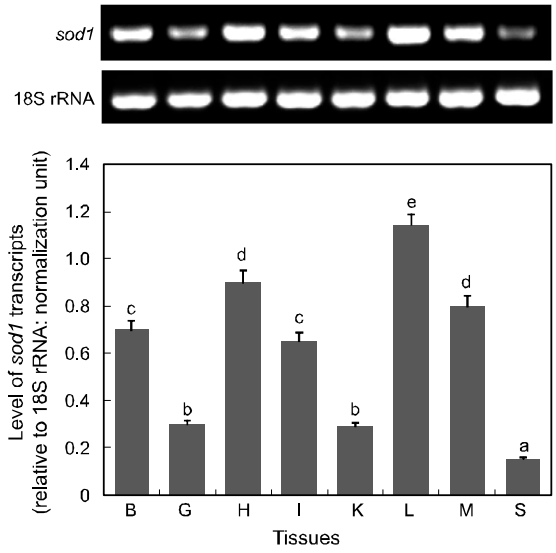

Fig. 5. Distribution of *Hemibarbus mylodon sod1* transcripts in different tissues as assessed by semi-quantitative RT-PCR analysis. Total RNAs from the brain (B), gill (G), heart (H), intestine (I), kidney (K), liver (L), muscle (M) and spleen (S) were subjected to RT-PCR analysis. Representative gels are provided to show the amplified RT-PCR products for both *sod1* and 18S rRNA. Densitometric analysis of *sod1* bands based on the normalization against 18S rRNA was performed using Quantity-One image analysis software implemented in VersaDoc 4,000 (Bio-Rad). Means with the same letters (a-e) are not significantly different based on ANOVA followed by Tukey's range test at a confidence level of 95%. Standard deviations are indicated by T-bars.

lowest level of *sod1* transcripts ($P<0.05$). The detection of *sod1* transcripts in a wide array of tissue types is not surprising, given that superoxide dismutase is one of the key housekeepers involved in general host protection against biological oxidation (Zelko et al., 2002). The results of this study are consistent with many previous reports on the variable basal levels of *sod1* transcripts across different tissue types (Cho et al., 2006; Nam et al., 2006). Furthermore, our finding of robust expression of *sod1* transcripts in the liver is consistent with the presence of multiple binding motifs for liver-enriched transcription factors such as C/EBP and hepatocyte nuclear factors (HNFs), in the 5'-upstream regulatory region of *hmsod1* [see Cho et al. (2009a)]. The important role of C/EBP in the liver-predominant expression of rat *sod1* has also been reported previously (Kim et al., 1997).

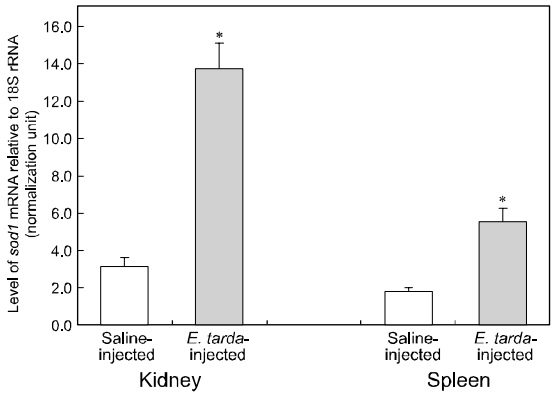

Fig. 6. Real-time RT-PCR assay of *Hemibarbus mylodon sod1* transcripts in kidney and spleen after an experimental challenge with *Edwardsiella tarda*. Freshly grown *E. tarda* (5×10^3 cells suspended in 100 μL of 0.85% NaCl) were delivered to each of juvenile *H. mylodon* ($n=6$), and kidney and spleen total RNA samples were assessed at 48 h post injection. Relative expression of *sod1* in the *E. tarda*-injected group to sham control group (injected with 0.85% NaCl only) was estimated by triplicate real-time RT-PCR assays using comparative CT method. T bars indicate the standard deviations. Asterisks indicate significant difference between bacteria-challenged and sham-challenged groups based on the Student's *t*-test ($P<0.05$).

Stimulated expression of *hmsod1* in response to bacterial challenge

No mortality was detected during the period (48 h) of the bacterial challenge using *E. tarda*. However injection of *E. tarda* suspension induced significant up-regulation of *hmsod1* transcripts in kidney and spleen. Rreal-time RT-PCR assay revealed that the amounts of *hmsod1* transcripts in kidney and spleen from bacteria-challenged fish were more than four-fold and three-fold, respectively higher than those observed in saline-injected control fish ($P< 0.05$). To date, the transcriptional regulation of fish *sod1* in response to bacterial challenge has rarely been exploited in other species; however, previous observations on non-teleostean species have indicated that exposure to endotoxic agents can induce oxidative damage (Ghosh et al., 1996; Choi et al., 2006; Li et al., 2008; Ji et al., 2009). The present expression data for *hmsod1* are also consistent with the presence of binding sites for immune-relevant transcription factors in its 5'-upstream regulatory region. Several studies on aquatic invertebrate species (*e.g.,* mollusks or crustaceans) have also indicated that *sod* transcription could be significantly modulated by infection with bacterial pathogen(s) (Cheng et al., 2006; Wang et al., 2008; Li et al., 2010). Additionally, Reyes-Becerril et al. (2008) reported that dietary supplementation of live yeast resulted in the elevation of SOD activity in the gilthead seabream (*Sparus aurata*) liver, which may support, at least in part, our finding of the stimulated transcription of *hmsod1*. Further, results of this study are similar to our recent observations of the up-regulated expression of the mitochondrial type of manganese SOD (Mn-SOD; *sod2*) from this species upon bacterial challenge (Cho et al., 2009c). Increased *sod* expression upon bacterial challenge has been generally recognized as an antioxidant defense action to scavenge or neutralize the ROS released during the proinflammatory cytokine response to bacterial toxins (Marikovsky et al., 2003). Further studies examining fish-to-fish variability in the transcriptional response to immune challenges would be of value. Specifically, real-time PCR analysis should be performed with individual samples rather than pooled RNA samples.

In summary, the molecular structure of the *sod1* gene was characterized in *H. mylodon*, an endangered freshwater fish species. The results of this study indicate that *hmsod1* exhibits the typical structure of vertebrate *sod1* genes at both the mRNA and genomic levels. A bioinformatic analysis of the 5'-flanking region suggested that various transcription factors may confer stress-related and/or tissue-specific modulation of *hmsod1*. From this basis, further studies should examine the unique or coordinated roles of predicted transcription factors in the regulation of *hmsod1* in response to different biotic and abiotic factors.

Acknowledgments

This study was supported by funds from the Korea-Ukraine International Cooperative Research Project (#2010-00091) from the National Research Foundation, Republic of Korea.

References

Afonso V, Santos G, Collin P, Khatib AM, Mitrovic DR, Lomri N, Leitman DC and Lomri A. 2006. Tumor necrosis factor-α down regulates human Cu/Zn superoxide dismutase 1 promoter via JNK/AP-1 signaling pathway. Free Rad Biol Med 41, 709-721.

Almeida JA, Diniz YS, Marques SFG, Faine LA, Ribas BO, Burneiko RC and Novelli ELB. 2002. The use of the oxidative stress responses as biomarkers in Nile tilapia (*Oreochromis niloticus*) exposed to *in vivo* cadmium contamination. Environ Intl 27, 673-679.

Bang IC, Lim YH, Cho YS, Lee SY and Nam YK. 2007. Survey of expressed sequence tags from tissue-specific cDNA libraries in *Hemibarbus mylodon*, an endangered fish species. J Aquacult 20, 248-254.

Bang IC, Lee YA and Lee WO. 2008. Cytogenetic analysis of three *Hemibarbus* species (Cypriniformes) from Korea. J Auqacult 21, 259-264.

Cheng W, Tung YH, Liu CH and Chen JC. 2006. Molecular cloning and characterization of copper/zinc superoxide dismutase (Cu,Zn-SOD) from the giant freshwater prawn *Macrobarachium rosenbergii*. Fish Shellfish Immunol 21, 102-112.

Cho YS, Choi BN, Kim KH, Kim SK, Kim DS, Bang IC and Nam YK. 2006. Differential expression of Cu/Zn superoxide dismutase mRNA during exposures to heavy metals in rockbream (*Oplegnathus fasciatus*). Aquaculture 253, 667-679.

Cho YS, Lee SY, Kim K-Y, Bang IC, Kim DS and Nam YK. 2008. Gene structure and expression of metallothionein during metal exposures in *Hemibarbus mylodon*. Ecotoxicol Environ Saf 71, 125-137.

Cho YS, Lee SY, Kim KH, Kim SK, Kim DS and Nam YK. 2009a. Gene structure and differential modulation of multiple rockbream (*Oplegnathus fasciatus*) hepcidin isoforms resulting from different biological stimulations. Dev Comp Immunol 33, 46-58.

Cho YS, Bang IC, Lee IR and Nam YK. 2009b. Hepatic expression of Cu/Zn-superoxide dismutase transcripts in response to acute metal exposure and heat stress in *Hemibarbus mylodon* (Teleostei: Cypriniformes). Fish Aquat Sci 12, 179-184.

Cho YS, Lee SY, Bang IC, Kim DS and Nam YK. 2009c. Genomic organization and mRNA expression of manganese superoxide dismutase (Mn-SOD) from *Hemibarbus mylodon* (Teleostei, Cypriniformes). Fish Shellfish Immunol 27, 571-576.

Choi YS, Lee KS, Yoon HJ, Kim I, Sohn HD and Jin BR. 2006. *Bombus ignitus* Cu,Zn superoxide dismutase (SOD1): cDNA cloning, gene structure, and up-regulation in response to paraquat, temperature stress, or lipopolysaccharide stimulation. Comp Biochem Physiol B Biochem Mol Biol 144, 365-371.

Ciriolo MR, Battistoni A, Falconi M, Filomeni G and Rotilio G. 2001. Role of the electrostatic loop of Cu, Zn superoxide dismutase in the copper uptake process. Eur J Biochem 268, 737-742.

Fukuhara R, Tezuka T and Kageyama T. 2002. Structure, molecular evolution, and expression of primate superoxide dismutases. Gene 296, 99-109.

Ghosh B, Hanevold CD, Dobashi K, Orak JK and Singh I. 1996. Tissue differences in antioxidant enzyme gene expression in response to endotoxin. Free Rad Biol Med 21, 533-540.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41, 95-98.

Hansen BH, Romma S, Garmo OA, Olsvik PA and Andersen RA. 2006. Antioxidative stress proteins and their gene expression in brown trout (*Salmo trutta*) from three rivers with different heavy metal levels. Comp Biochem Physiol C Toxicol Pharmacol 143, 263-274.

Jang MH, Lucas MC and Joo G-J. 2003. The fish fauna of mountain streams in South Korean national parks and its significance to conservation of regional freshwater fish biodiversity. Biol Conserv 114, 115-126.

Ji PF, Yao CL and Wang ZY. 2009. Immune response and gene expression in shrimp (*Litopenaeus vannamei*) hemocytes and hepatopancreas against some pathogen-associated molecular patterns. Fish Shellfish Immunol 27, 563-570.

Kim K-Y and Nam YK. 2006. Molecular phylogenetic analyses of *Scyliorhinus torazame* (Carcharhiniformes) inferred from Cu,Zn superoxide dismutase. Kor J Ichthyol 18, 293-299.

Kim K-Y, Lee SY, Cho YS, Bang IC, Kim DS and Nam YK. 2008. Characterization and phylogeny of two β-cytoskeletal actins from *Hemibarbus mylodon* (Cyprinidae, Cypriniformes), a threatened fish species in Korea. DNA Seq 19, 87-97.

Kim YH, Yoo HY, Chang MS, Jung G and Rho HM. 1997. C/EBPα is a major activator for the transcription of rat Cu/Zn superoxide dismutase gene in liver cell. FEBS Lett 401, 267-270.

Li C, Sun H, Chen A, Ning X, Wu H, Qin S, Xue Q and Zhao J. 2010. Identification and characterization of an

intracellular Cu, Zn-superoxide dismutase (icCu/Zn-SOD) gene from clam *Venerupis philippinarum*. Fish Shellfish Immunol 28, 499-503.

Li Q, Kumar A, Gui JF and Yu FSX. 2008. *Staphylococcus aureus* lipoproteins trigger human corneal epithelial innate response through toll-like receptor-2. Microb Pathogen 44, 426-434.

Marikovsky M, Ziv V, Nevo N, Harris-Cerruti C and Mahler O. 2003. Cu/Zn superoxide dismutase plays important role in immune response. J Immunol 170, 2993-3001.

Miracle AL and Ankley GT. 2005. Ecotoxicogenomics: linkage between exposure and effects in assessing risks of aquatic contaminants to fish. Reprod Toxicol 19, 321-326.

Miya M, Kawaguchi A and Nishida M. 2001. Mitogenomic exploration of higher teleostean phylogenies: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. Mol Biol Evol 18, 1993-2009.

Nam YK, Cho YS, Kim K-Y, Bang IC, Kim KH, Kim SK and Kim DS. 2006. Characterization of copper, zinc superoxide dismutase from a cartilaginous shark species, *Scyliorhinus torazame* (Carcharhiniformes). Fish Physiol Biochem 32, 305-315.

Nylander JAA. 2004. MrModeltest v2. Evolutionary Biology Centre, Uppsala University.

Park EY and Rho HM. 2002. The transcriptional activation of the human copper/zinc superoxide dismutase gene by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin through two different regulator sites, the antioxidant responsive element and xenobiotic responsive element. Mol Cell Biochem 24, 47-55.

Pedrajas JR, Bavilanes F, López-Barea J and Peinado J. 1998. Incubation of superoxide dismutase with malondialdehyde and 4-hydroxy-2-nonenal forms new active isoforms and adducts: an evaluation of xenobiotics in fish. Chemico-Biol Interact 116, 1-17.

Pimental RA, Liang B, Yee GK, Wilhelmsson A, Poellinger L and Paulson KE. 1993. Dioxin receptor and C/EBP regulate the function of the glutathione *S*-transferase Ya gene xenobiotic response element. Mol Cell Biol 13, 4365-4373.

Posada D and Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Regoli F, Gorbi S, Frenzilli G, Nigro M, Corsi I, Focardi S and Winston GW. 2002. Oxidative stress in ecotoxicology: from the analysis of individual antioxidants to a more integrated approach. Mar Environ Res 54, 419-423.

Reyes-Becerril M, Salinas I, Cuesta A, Meseguer J, Tovar-Ramirez D, Ascencio-Valle F and Esteban MA. 2008. Oral delivery of live yeast *Debaryomyces hansenii* modulates the main innate immune parameters and the expression of immune-relevant genes in the gilthead seabream (*Sparus aurata* L). Fish Shellfish Immunol 25, 731-739.

Rojo AI, Salinas M, Martin D, Perona R and Cuadrado A. 2004. Regulation of Cu/Zn-superoxide dismutase expression via the phophatidylinositol 3 kinase/Akt pathway and nuclear factor-κB. J Neurosci 24, 7324-7334.

Ronquist F and Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

Ryu DY, Levi PE, Fernandez-Salguero P, Gonzalez FJ and Hodgson E. 1996. Piperonyl butoxide and acenaphthylene induce cytochrome P450 1A2 and 1B1 mRNA in aromatic hydrocarbon-responsive receptor knock-out mouse liver. Mol Pharmacol 50, 443-446.

Schmittgen TD and Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3, 1101-1108.

Shi H, Sui Y, Wang X, Luo Y and Ji L. 2005. Hydroxyl radical production and oxidative damage induced by cadmium and naphthalene in liver of *Carassius auratus*. Comp Biochem Physiol C Toxicol Pharmacol 140, 115-121.

Snape JR, Maund SJ, Pickford DB and Hutchinson TH. 2004. Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. Aquatic Toxicol 6, 143-154.

Swofford DL. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver 4, Sinauer Associates, Sunderland, MA.

Thompson JD, Higgins DG and Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673-4680.

Vega-López A, Galar-Martínez M, Jiménez-Orozco FA, García-Latorre E and Dominguez-López ML. 2007. Gender related differences in the oxidative stress response to PCB exposure in an endangered goodeid fish (*Girardinichthys viviparus*). Comp Biochem Physiol A Mol Integr Physiol 146, 672-678.

Wang KJ, Ren HL, Xu DD, Cai L and Yang M. 2008. Identification of the up-regulated expression genes in hemocytes of variously colored abalone (*Haliotis diversicolor* Reeve, 1846) challenged with bacteria. Dev Comp Immnol 32, 1326-1347.

Yoo HY, Chang MS and Rho HM. 1999a. Xenobiotic-responsive element for the transcriptional activation of the rat Cu/Zn superoxide dismutase gene. Biochem

Biophys Res Commun 256, 133-137.

Yoo HY, Chang MS and Rho HM. 1999b. Heavy metal-mediated activation of the rat Cu/Zn superoxide dismutase gene via a metal-responsive element. Mol Gen Genet 262, 310-313.

Yueh MF, Huang YH, Hiller A, Chen S, Nguyen N and Tukey RH. 2003. Involvement of the xenobiotic response element (XRE) in Ah receptor-mediated induction of human UDP-glucuronosyltransferase 1A1. J Biol Che``m 278, 15001-15006.

Zelck UE, Janje B and Schneider O. 2005. Superoxide dismutase expression and H_2O_2 production by hemocytes of the trematode intermediate host *Lymnaea stagnalis* (Gastropoda). Dev Comp Immunol 29, 305-314.

Zelko IN, Mariani TJ and Folz RJ. 2002. Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. Free Radic Biol Med 33, 337-349.

(Received 15 February 2011; Revised 3 March 2011; Accepted 14 March 2011)