Pakistan J. Zool., vol. 52(4), pp 1377-1382, 2020 DOI: https://dx.doi.org/10.17582/journal.pjz/20190122060142

In Silico Analysis of Cu-Zn Superoxide Dismutase and Mn Superoxide Dismutase Genes in Fugu (*Takifugu rubripes*)

Mehtap Bayir*
Department of Agricultural Biotechnology, Faculty of Agriculture, Atatürk University, 25240, Erzurum, Turkey

ABSTRACT

Superoxide dismutases are the best-known enzymatic antioxidants because of their central role in the antioxidant defense system. In this study, the phylogeny, gene structure, and conserved gene synteny of *sod1* and *sod2* genes in fugu—a model organism—were determined. Maximum amino acid similarity identity was found between putative fugu Sod1 and Sod2 proteins and their orthologs from teleost fish and tetrapods. Phylogenetic clustering was seen between *sod* genes in fugu and their orthologs. Finally, highly conserved gene synteny was determined between fugu *sod* genes and their orthologs from teleost fish and human.

Article Information
Received 22 January 2019
Revised 22 May 2019
Accepted 27 August 2019
Available online 02 April 2020

Key words
Bioinformatics, Fugu, Model organism, *sod1*, *sod2*

INTRODUCTION

Fugu (Japanese pufferfish, *Takifugu rubripes*) is an ideal model organism for vertebrate genome research and developmental biology studies (Uji *et al*., 2011). Japanese pufferfish has more advantages for genomic studies compared with other vertebrates. Its 400-Mb genome size is quite small compared with the 3,000-Mb human genome size (Watabe and Ikeda, 2006). It is also known that the pufferfish genome has more genes than those of coelacanths and air-breathing fish due to teleost-specific whole genome duplication (tsWGD) event (Van de Peer, 2004). The small size of the genome, which is a factor that facilitates the detection and analysis of genes, has regular sequences like those of the other vertebrates, requiring less work to obtain comparable data. Green spotted pufferfish (*Tetraodon nigroviridis*) can also be used as model organisms for vertebrate genomes (Close *et al*., 2016). Many human genes have been uncovered by comparing these two pufferfish genomes with the human genomes (Brenner *et al*., 1993).

Aerobic biological systems generate reactive oxygen species (ROS), including hydrogen peroxide (H_2O_2) and superoxide anion ($O_2^{\cdot-}$). ROS attacking essential cellular components may result in irreversible damage in their structure. However, organisms can initiate ROS removal and restrict generation of ROS via their well-developed antioxidant defense system (ADS).

The superoxide dismutases (SODs) have an important role in ADS due to their ability to prevent the generation of hydroxyl radical ($OH^{\cdot}$) and removing $O_2^{\cdot-}$ and catalyzing the dismutation of $O_2^{\cdot-}$ into H_2O_2 and molecular oxygen. SODs using metal ions in their activities have two forms of eukaryotic systems, namely, Cu^{+2}/Zn^{+2} and Mn^{+2} SOD. While the first isoform (Cu^{+2}/Zn^{+2} SOD) is found in cytosols and its activity is not affected by oxidative stress, the second isoform (Mn^{+2} SOD) is located in mitochondria, and its activity is increased in proportion to oxidative stress (Babior, 1997; Davies, 2000). Several studies have been carried out on SOD enzymes in aerobic organisms (Lopes *et al*., 2001; Fink *et al*., 2002; van der Oost *et al*., 2003; Ken *et al*., 2003; Farombi *et al*., 2007; Cho *et al*., 2009). Limited work, however, is available on the genomic organization and gene structure of teleost *sods* (Cho *et al*., 2009) and regulatory region of the *sod* genes in teleost fish (Maehara *et al*., 1999; Mao *et al*., 2006). It is known that transcription of the Mn^{+2} *SOD* gene can be regulated by environmental factors (Valavanidis *et al*., 2006; Kim *et al*., 2007; Cho *et al*., 2009), and the function and structure of *SOD2* are well conserved in variant organisms (Fink *et al*., 2002). Therefore, the goals of the current study are determining *sod* genes in the fugu genome using bioinformatic tools, leading to future molecular works on antioxidant enzyme genes in teleost fishes.

MATERIALS AND METHODS

Fugu *sod1* and *sod2* gene sequences were obtained by performing BLAST (http://useast.ensembl.org/Multi/blastview) searches with an identical orthologous

* Corresponding author: mehtap.bayir@atauni.edu.tr
0030-9923/2020/0004-1377 $ 9.00/0

zebrafish (*Danio rerio*) Sod protein sequence as an inquiry to the Ensembl genome sequence database (http://useast.ensembl.org/index.html). Zebrafish *sod* cDNA sequences provided from the Ensembl genome database were used to identify ESTs coded by specific fugu *sod1* and *sod2* genes in a BLAST search of the NCBI database (http://blast.ncbi.nlm.nih.gov).

Teleost fish exhibits strict evolutionary conservation for the gene structure in the same gene family. Therefore, the exon-intron junctions of the *sod1* and *sod2* genes were determined using zebrafish *sod1* and *sod2* gene structures as a reference. To confirm whether *sod1* and *sod2* genes were transcriptionally active or served as pseudogenes or non-functional genes, tBLASTn searches in the NCBI database were conducted to identify their ESTs using their Ensembl-derived amino acid sequences as queries (Table I). Superoxide dismutase genes in fugu with their corresponding zebrafish query sequence IDs, NCBI cDNA IDs, chromosomal locations and length of Sod polypeptides were also determined using Ensembl and the NCBI genome databases (Table II).

Table I. Ensembl gene ID and expressed sequence tags (EST) coded by fugu and zebrafish *sod1* and *sod2* genes.

Gene	Ensembl gene ID	EST
Fugu *sod1*	ENSTRUG00000021322	Not detected
Fugu *sod2*	ENSTRUG00000005242	Not detected
Zebrafish *sod1*	ENSDARG00000043848	EH544614.1
Zebrafish *sod2*	ENSDART00000062556.4	EH537787.1

Phylogenetic analysis

CLUSTALW (Thompson *et al*., 1994) at BioEdit software (http://www.mbio.ncsu.edu/bioedit/page2.html) was used for sequence alignment of the *sod1* and *sod2* genes. The protein sequence of fugu Sod1 and Sod2 was aligned with Sod/SOD protein sequences from fugu, tetraodon, medaka, zebrafish, human, and mice. The pairwise alignment of the BLOSUM62 matrix (Gromiha, 2010) was used for sequence identity and similarity. A maximum-likelihood tree with the Poisson correction distance model based on amino acid substitution per site was built using MEGA6 (Tamura *et al*., 2013) to determine the phylogenetic relationships of the fugu, tetraodon, medaka, zebrafish, human, and mouse Sod sequences. A bootstrapped neighbor-joining tree was also constructed before construction of the maximum-likelihood tree to confirm the phylogeny of *sod* genes (data not shown). As in a previous study, the protein sequence of human lymphocyte cytosolic protein LCP2) was used as an external group (Kell *et al*., 2018).

Conserved gene synteny

The conserved gene synteny of fugu sod genes with the sod/SOD of zebrafish, medaka and human was arranged manually using the region conceptus selection of Ensembl database to recognise co-localized gene (Thirumaran and Wright, 2014).

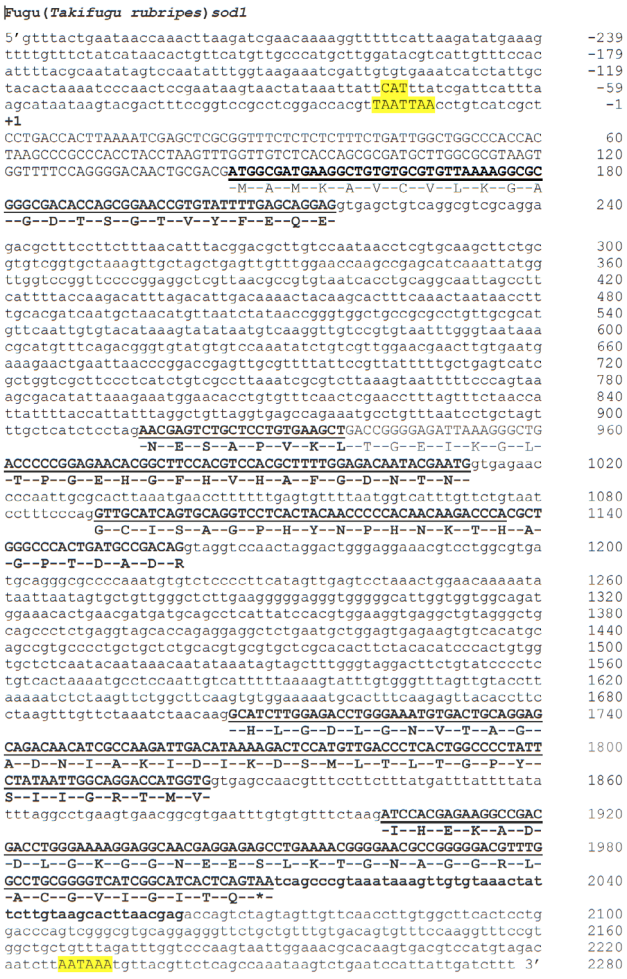

Fig. 1. The nucleotide sequence of the fugu *sod1* gene. Exons of the *sod1* gene are shown in capital letters and nucleotide position is indicated the numbers end the rows. The start site of transcription is indicated by +1. The 5' upstream sequence, 3'downstream sequence and introns are shown in lower-case letters. A putative CAAT box, TATA box and polyadenylation signal (AATAAAA) are shown capital letters highlighted in yellow. The stop codon (TAA) is indicated by an asterisk.

Table II. Superoxide dismutase genes of fugu with their corresponding zebrafish query sequence ID, NCBI cDNA ID, locations of chromosome and length of *sod* polypeptides.

Gene	Zebrafish query sequence ID	NCBI cDNA ID	Location	Number of amino acids
sod1	NP_571369.1	XP_003971372.1	Chromosome 10: 25 694 395-25 699 454	154
sod2	NP_956270.1	XP_003971923.1	Chromosome 16: 9 495 906-9 498 317	227

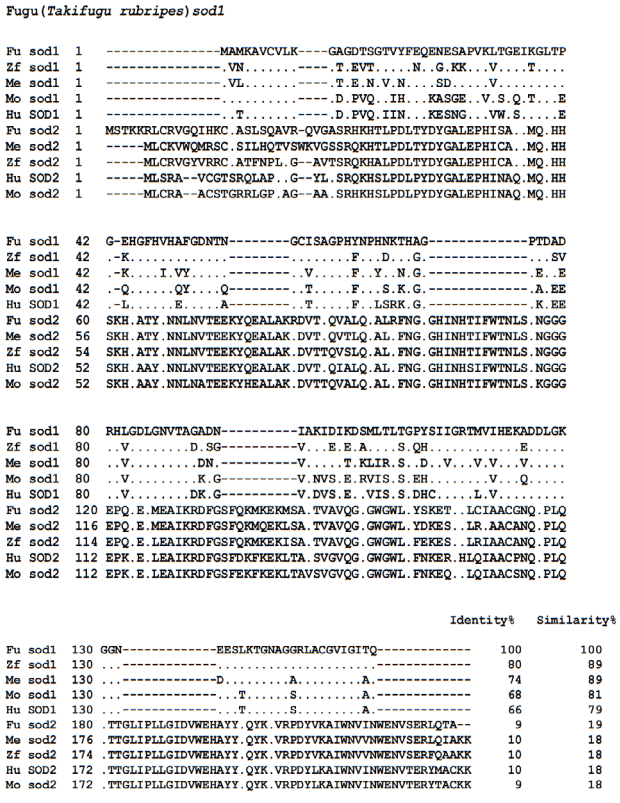

Fig. 2. The percentage of sequence identity and similarity among fugu *sod1* gene and human, mouse, zebrafish, medaka *sod1* and *sod2* gene. The dots and dashes in the alignment represent sequence identity and missing amino acids respectively.

RESULTS AND DISCUSSION

Bioinformatics and computational analysis of fugu sod genes

As a known procedure in bioinformatics studies, first, statistical knowledge is collected using biological data; then, a model is generated for solving any computational modeling problem. Finally, testing and evaluation of a computational algorithm for solving a bioinformatics problem are conducted (Can, 2014). In this study, I retrieved some statistics using the Ensembl genomic database, NCBI database, BioEdit software, pairwise alignment of the BLOSUM62 matrix program and MEGA6 program.

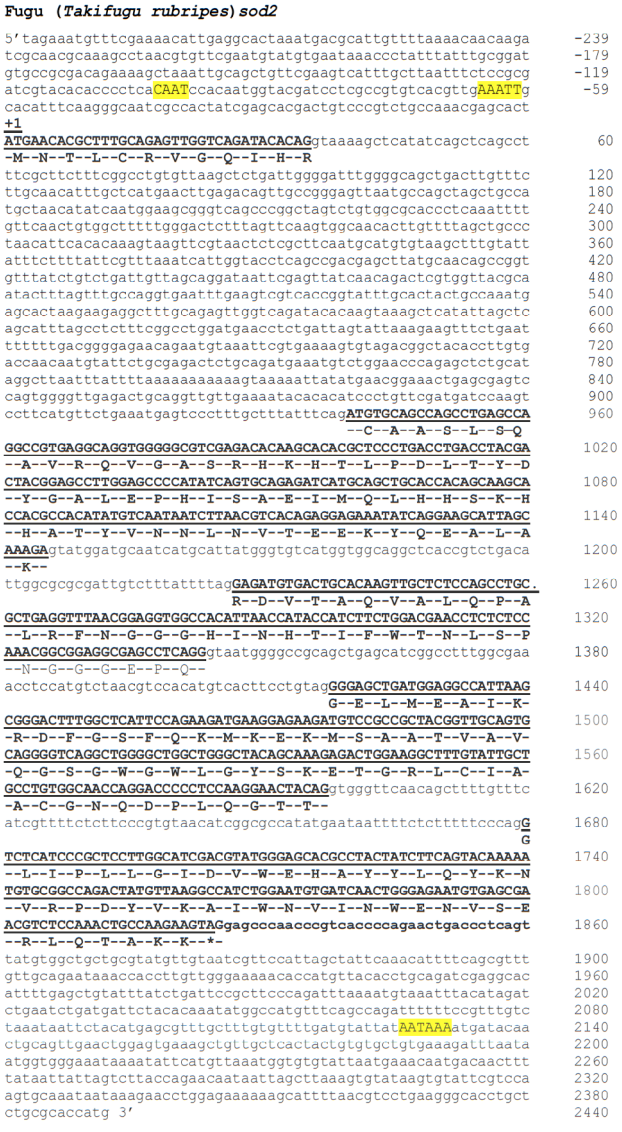

Fig. 3. The nucleotide sequence of the fugu *sod2* gene. Exons of the *sod2* gene are shown in capital letters and nucleotide position is indicated the numbers end the rows. The start site of transcription is indicated by +1. The 5' upstream sequence, 3'downstream sequence and introns are shown in lower-case letters. A putative CAAT box, TATA box and polyadenylation signal (AATAAAA) are shown capital letters highlighted in yellow. The stop codon (TAG) is indicated by an asterisk.

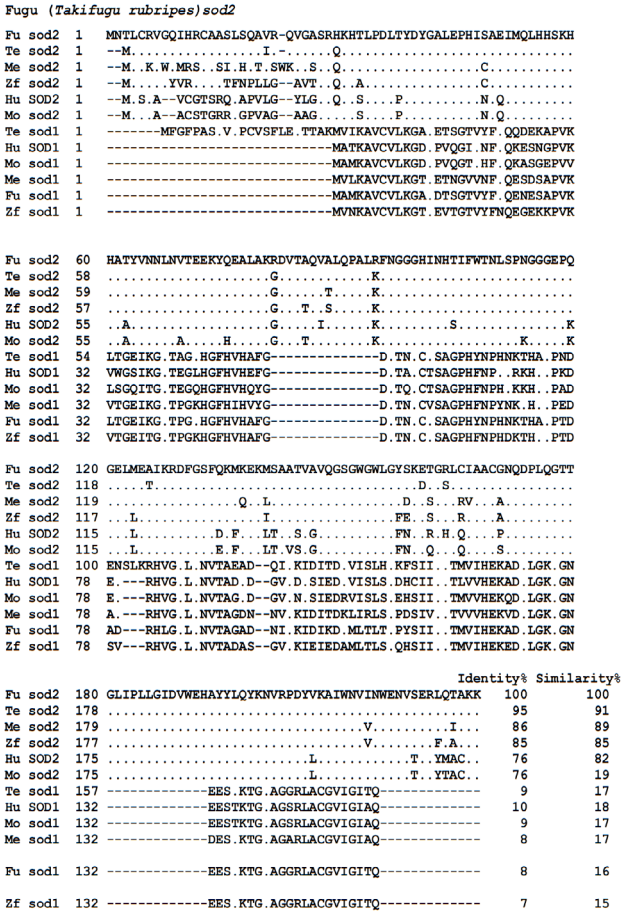

Fig. 4. The percentage of sequence identity and similarity among fugu *sod2* gene and human, mouse, zebrafish, medaka *sod1* and *sod2* gene. The dots and dashes in the alignment represent sequence identity and missing amino acids respectively.

The cDNA sequence was obtained from the Ensembl genomic database to determine the exon–intron structure of the fugu *sod* genes, and it was found that these genes had five exons separated by four introns. It was determined that introns of both genes followed the gt-ag rule. Moreover, they have putative TATA and CAAT boxes and polyadenylation signals. The results clearly show that *sod* genes in fugu exhibit a highly conserved gene structure (Figs. 1 and 3).

I searched the ESTs of *sod1* and *sod2* by BLAST searches in the NCBI genomic database using cDNA nucleotide sequences obtained from the Ensembl genomic database (Table I). However, as shown in Table I, there was no EST for these genes. The sequence identity and similarity among the fugu and human, mouse, zebrafish, and medaka *sod* genes are given in Figures 2 and 4. The highest identity and similarity rates for fugu *sod1* and *sod2* genes were determined with their orthologs (tetraodon, medaka, and zebrafish). After that, a maximum-likelihood phylogenetic tree was built based on identity and similarity results (Fig. 5). At the end of the phylogeny analysis, high phylogenetic clustering was seen between the *sod* genes in fugu and their orthologs. Conserved gene synteny evidence was determined for *sod* genes of fugu and *Sod/sod* genes of other teleost fishes and human (Fig. 6). The syntenic genes of fugu *sod1* and *sod2*, located on chromosomes 15 and 16, exhibit conserved gene synteny with human *SOD1* and *SOD2* located on chromosomes 21 and 6, medaka *sod1* and *sod2* located on chromosomes 14 and 24, and zebrafish *sod1* and *sod2* located on chromosomes 10 and 20. The results of conserved gene synteny between *sod* genes in fugu and their orthologs in medaka, zebrafish, and human show that teleost fish lost duplicated copies of the *sod1* and *sod2* genes after tsWGD because of mutations. This would be an interesting result for future studies on whether all teleost species have single copies of *sod* genes and their transcriptional controls.

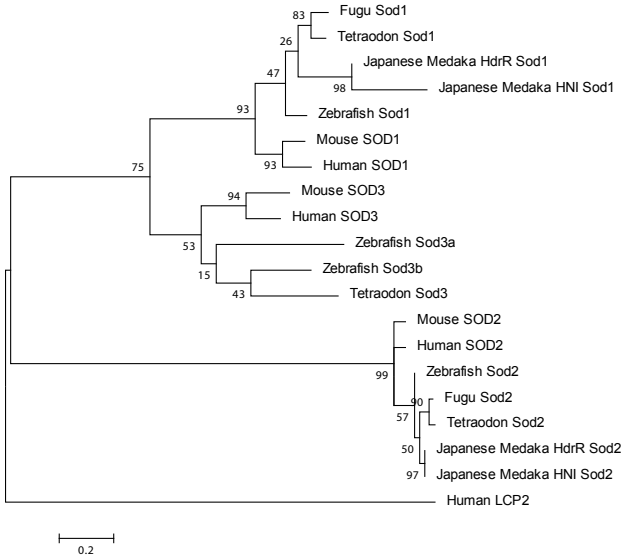

Fig. 5. Phylogenetic relationships between *sod1* and *sod2* sequence from fugu and *sod* sequences from other vertebrates. Accession numbers of the sequences used for phylogenetic tree are the following: Fugu Sod1/Sod2: HE602549, HE602550, Tetraodon Sod1/Sod2/Sod3: H3CX53, Q4S6V6, H3D497; Japanese medaka HDrR Sod1/Sod2: XP_004076261, GCA_002234675, Zebrafish Sod1/Sod2/Sod3: NP_571369, NP_956270, XP_001332758, Mouse Sod1/Sod2/Sod3: NP_035564, NP_038699.2, NP_035565, Human Sod1/Sod2/Sod3: NP_000445, NG_008729, NP_003093.

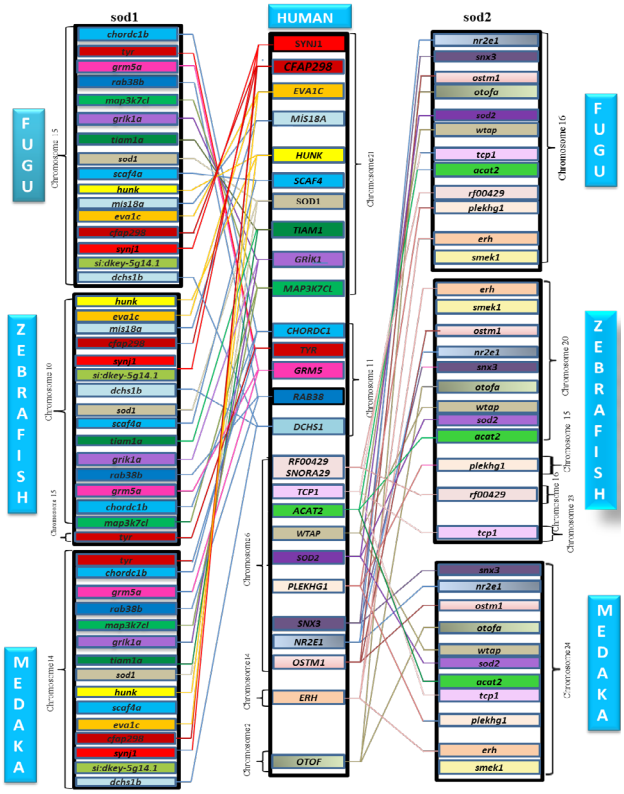

Fig. 6. Conserved gene synteny of fugu *sod* genes.

CONCLUSION

Antioxidant enzymes are gaining more importance, and they have been studied more in recent years. When antioxidant levels and production of ROS are in balance in normal conditions, the harmful effect of free radicals can be observed with oxidative stress-induced diseases (Scandalios, 1993). Stress responses in fishes can manifest multifaceted levels involving the actions of different sets of genes and their products. Understanding the genetic characteristics associated with model organisms that exhibit a stress response is important for molecular studies. Identification and characterization of the stress genes that are differentially expressed between stress-tolerant and intolerant fish will provide important genetic markers that may be used in aquaculture selection programs to help improve stress tolerance, as well as serving as a model for other vertebrates, including human (Iwama *et al*., 1999). However, it is known that oxidative stress has a big role in more than 100 diseases in human as their reason or effect (Halliwell *et al*., 1992; Gutteridge, 1993; Poljsak *et al*., 2013). For this purpose, I identified *sod* genes in fugu—serving as a model organism—using bioinformatic tools; the results can illuminate the path for future studies on molecular stress responses in fish.

Statement of conflict of interest

The authors declare there is no conflict of interest.

REFERENCES

Babior, B.M., 1997. Superoxide: a two edged sword. *Brazilian J. med. Boil. Res.*, **30**: 141-155. https://doi.org/10.1590/S0100-879X1997000200001

Brenner, S., Elgar, G, Sandford, R., Macrae, A., Venkatesh, B. and Aparicio, S., 1993. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature*, **366**: 265-268. https://doi.org/10.1038/366265a0

Can, T., 2014. Introduction to Bioinformatics. In: *miRNomics: microRNA biology and computational analysis. Methods in molecular biology (Methods and Protocols)* (eds. M. Yousef and J. Allmer), Vol. 1107. Humana Press, Totowa, NJ.

Close, R.A., Johanson, Z., Tyler, J.C., Harrington, R.C. and Friedman, M., 2016. Mosaicism in a new Eocene pufferfish highlights rapid morphological innovation near the origin of crown tetraodontiforms. *Palaeontology*, **59**: 499-514. https://doi.org/10.1111/pala.12245

Cho, Y.S., Lee, S.Y., Bang, I.C., Kim, D.S. and Nam, Y.K., 2009. Genomic organization and mRNA expression of manganese superoxide dismutase (Mn-SOD) from *Hemibarbus mylodon* (Teleostei, Cypriniformes). *Fish Shellf. Immunol.*, **27**: 571-576. https://doi.org/10.1016/j.fsi.2009.07.003

Davies, K.J.A., 2000. An overview of oxidative stress. *IUBMB Life*, **50**: 241-244. https://doi.org/10.1080/15216540051080895

Farombi, E.O., Adelowo, O.A. and Ajimoko, Y.R., 2007. Biomarkers of oxidative stress and heavy metal levels as indicators of environmental pollution in African catfish (*Clarias gariepinus*) from Nigeria Ogun River. *Int. J. environ. Res. Publ. Hlth.*, **4**: 158-165. https://doi.org/10.3390/ijerph2007040011

Fink, R.C. and Scandalios, J.G., 2002. Molecular evolution and structure-function relationships of the superoxide dismutase gene families in angiosperms and their relationship to other eukaryotic and prokaryotic superoxide dismutases. *Arch. Biochem. Biophys.*, **399**: 19–36. https://doi.org/10.1006/abbi.2001.2739

Gromiha, M.M., 2010. Protein sequence analysis. In: *Protein bioinformatics: From sequence to function*. Elsevier Inc., New Delhi, India, pp. 29-62. https://

doi.org/10.1016/B978-8-1312-2297-3.50002-3

Gutteridge, J.M.C., 1993. Free radicals in disease processes: a compilation of cause and consequence. *Free Rad. Res. Comm.*, **19**: 141-158. https://doi.org/10.3109/10715769309111598

Halliwell, B., Gutteridge, J.M.C. and Cross, C.E., 1992. Free radicals, antioxidants, and human disease: where are we now? *J. Lab. Clin. Med.*, **119**: 598-620.

Iwama, G.K., Vijayan, M.M., Forsyth, R.B. and Ackerman, P.A., 1999. Heat shock proteins and physiological stress in fish. *Am. Zool.*, **39**: 901-909. https://doi.org/10.1093/icb/39.6.901

Kell, M.J., Riccio, R.E., Baumgartner, E.A., Compton, Z.J., Pecorin, P.J., Mitchell, T.A., Topczewski, J. and LeClair E.E., 2018. Targeted deletion of the zebrafish actinbundling protein L-plastin (lcp1). *PLoS One*, **13**: e0190353. https://doi.org/10.1371/journal.pone.0190353

Ken, C.F., Lin, C.T., Shaw, J.F. and Wu, J.L., 2003. Characterization of fish Cu/Zn–superoxide dismutase and its protection from oxidative stress. *Mar. Biotech.*, **5**: 167-173. https://doi.org/10.1007/s10126-002-0058-1

Kim, K.Y., Lee, S.Y., Cho, Y.S., Bang, I.C., Kim, K.H. and Kim, D.S., 2007. Molecular characterization and mRNA expression during metal exposure and thermal stress of copper/zinc- and manganese-superoxide dismutases in disk abalone, *Haliotis discus discus*. *Fish Shellf. Immunol.*, **23**: 1043-1059. https://doi.org/10.1016/j.fsi.2007.04.010

Lopes, P.A., Pinheiro, T., Santos, M.C., da Luz Mathias, M., Collares-Pereira, M.J. and Viegas-Crespo, A.M., 2001. Response of antioxidant enzymes in freshwater fish populations (*Leuciscus alburnoides* complex) to inorganic pollutants exposure. *Sci. Total Environ.*, **280**: 153-163. https://doi.org/10.1016/S0048-9697(01)00822-1

Maehara, K., Hasegawa, T., Xiao, H., Takeuchi, A., Abe, R. and Isobe, K., 1999. Cooperative interaction of NF-κB and C/EBP binding sites is necessary for manganese superoxide dismutase gene transcription mediated by lipopolysaccharide and interferon-γ. *FEBS Lett.*, **449**: 115-119. https://doi.org/10.1016/S0014-5793(99)00408-1

Mao, X., Moerman-Herzog, A.M., Wang, W. and Barger, S.W., 2006. Differential transcriptional control of the superoxide dismutase-2 κB element in neurons and astrocytes. *J. biol. Chem.*, **281**: 35863-35872. https://doi.org/10.1074/jbc.M604166200

Poljsak, B., Šuput, D. and Milisav, I., 2013. Achieving the balance between ROS and antioxidants: when to use the synthetic antioxidants. *Corpor. Oxid. Med. Cell. Long.*, Article ID 956792: 11. https://doi.org/10.1155/2013/956792

Scandalios, J.G., 1993. Oxygen stress and superoxide dismutases. *Plt. Physiol.*, **101**: 7-12. https://doi.org/10.1104/pp.101.1.7

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis, Version 6.0. *Mol. Biol. Evol.*, **30**: 2725–2729. https://doi.org/10.1093/molbev/mst197

Thirumaran, A. and Wright, J.M., 2014. Fatty acid-binding protein (*fabp*) genes of spotted green pufferfish (*Tetraodon nigroviridis*): comparative genomics and spatial transcriptional regulation. *Genome*, **57**: 289-301. https://doi.org/10.1139/gen-2014-0059

Thompson, J.D., Higgins, D.G. and Gibson, T.J., 1994. CLUSTALW: 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.

Uji, S., Kurokawa, T., Hashimoto, H., Kasuya, T. and Suzuki, T., 2011. Embryogenic staging of fugu, *Takifugu rubripes*, and expression profiles of *aldh1a2*, *aldh1a3* and *cyp26a1*. *Develop. Growth Differen.*, **53**: 715-725. https://doi.org/10.1111/j.1440-169X.2011.01281.x

Valavanidis, A., Vlahogianni, T., Dassenakis, M. and Scoullos, M., 2006. Molecular biomarkers of oxidative stress in aquatic organisms in relation to toxic environmental pollutants. *Ecotoxicol. Environ. Safe.*, **64**: 178-189. https://doi.org/10.1016/j.ecoenv.2005.03.013

Van Der Oost, R., Beyer, J. and Vermeulen N.P.E., 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: A review. *Environ. Toxicol. Pharmacol.*, **13**: 57-149. https://doi.org/10.1016/S1382-6689(02)00126-6

Van de Peer, Y., 2004. Tetraodon genome confirms Takifugu findings: Most fish are ancient polyploids. *Genome Biol.*, **5**: 250. https://doi.org/10.1186/gb-2004-5-12-250

Watabe, S. and Ikeda, D., 2006. Diversity of the pufferfish *Takifugu rubripes* fast skeletal myosin heavy chain genes. *Comp. Biochem. Physiol., Part D.*, **1**: 28-34. https://doi.org/10.1016/j.cbd.2005.12.001