

## MULTIPLE SEQUENCE ALIGNMENT AND PHYLOGENETIC STUDIES OF COX-1 AND COX-2 FROM MUS MUSCULUS


A.Rama Mohan Reddy and Chukka Santhaiah

Department of C.S.E, S.V.U.College of Engg, S.V.University, Tirupati, A.P, INDIA

**ABSTRACT:** The biological datasets of two genes of *Mus musculus* which are Cyclooxygenase (COX)-1 and COX-2 officially known as prostaglandin-endoperoxide synthase (PTGS), were selected for Phylogenetics studies using algorithm. This analysis has focused on Cyclooxygenase gene families for which initial characterizations have been achieved for individual members. Phylogenetic trees of each family define the evolutionary relationships of the members to each other. These families contain numerous members, indicating diverse functions in vivo. Closely related isoforms and separate subfamilies exist within many of these gene families, indicating possible redundancies and specialized functions. To facilitate their further study, we have developed phylogenetic tree of the analyzed genes.

**Key Words:** Key words: Cyclooxygenase, Phylogenetics, Alignment, Algorithm, *Mus musculus*

\*Corresponding author: Chukka Santhaiah, Department of C.S.E, S.V.U.College of Engg, S.V.University, Tirupati, A.P, INDIA

Copyright: ©2016 Chukka Santhaiah. This is an open-access article distributed under the terms of the Creative Commons Attribution License , which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

### INTRODUCTION

Cyclooxygenase (COX) is the enzyme that catalyzes the oxidation and subsequent reduction of arachidonic acid to form Prostaglandin G<sub>2</sub> and Prostaglandin H<sub>2</sub> (PGH<sub>2</sub>) [1,2]. PGH<sub>2</sub> can then undergo additional reactions to produce the primary prostaglandins (PGs), which participate in a variety of physiological functions in the vertebrates, including inducing fever, maintaining pregnancy, and regulating ion transport [3-5]. These functions have been extensively studied in mammals [6], but comparatively little functional data exist for other animals despite the sequencing of COX genes in several other vertebrates, particularly the teleosts [7-9]. Based on functional studies of COX in the teleosts, it seems that some functions are conserved [10], whereas others may be altered in some species [11] or are novel [12]. A phylogenetic tree is described as, a branching diagram that shows, for each species, with which other species it shares its most recent common ancestor [13]. The evolutionary tree or cladograms were traditionally used to draw evolutionary relationship among the organism; a more modern version of the same is phylogenetic tree which uses protein sequences to draw the evolutionary relationship. These trees dictate the relationship among the organisms based on the similarity and dissimilarity among the nucleotide or nucleic acid sequences [14]. The tree construction can be done through variety of tree-building methods which include methods based on distances, likelihood and characters. After a phylogenetic tree is constructed, it is important to test its accuracy which refers to the degree to which a tree is close to the true tree. Phylogenetics is the study of evolutionary relationships among organisms or genes. The purpose of phylogenetic studies is to reconstruct evolutionary ties between organisms and to estimate the time of divergence between organisms since they last shared a common ancestor. There are several types of data that can be used to build phylogenetic trees like DNA sequences, Gene Sequences and protein sequences.

## METHODOLOGY

Reference proteins of well-established molecular function, representing each of the protein families investigated, were chosen as query sequences for searches in the *Mus musculus* genome databases. These reference proteins were BG086892; AW555640; BG075861; BG085367. Searches were made using the TBLASTN tool against GenBank database non-redundant (NR), with search specifications for *O. sativa*. The other databases used were SWISSPROT and Universal Protein resource Uniprot. The BLAST server used was that of the National Center for Biotechnology Information. As selection criteria of BLAST hits for genomic sequences, a cut off e-value of  $e^{-10}$  was previously set. The genomic sequences found were used to predict putative genes contained within them. Whenever possible, genes were predicted on the basis of sequences generated by the *Mus musculus* Genome database, since these sequences present a higher degree of accuracy. To that end, a mixed procedure was adopted combining ab initio gene prediction algorithms of genomic sequence alignments with similar sequences from expressed genes (ESTs and cDNAs). The prediction algorithms were GenScan, GenomeScan, FGENESH, GeneMark.hmm and GraileXP. Such expressed sequences were found by BLAST searches against EST and NR databases of GenBank, using the genomic sequence as query. The algorithm of choice for the multiple alignments of protein sequences was ClustalX1.8, available through the BCM Search Launcher server. The multiple alignments were edited with the help of GENEDOC. All the genes with greater than 30% identity, with at least one of the reference proteins used in the searches, were regarded as functionally similar (homologous) to the reference proteins, receiving the same name. Those sequences that did not conform to this criterion were discarded. Prediction of homology and signature sequences for the putative transporter proteins were carried out with PROSITE and Pfam databases. Sequences were included into families based on homology and presence of signature sequences. Protein alignments obtained with ClustalX 1.8 were used as starting points for phylogenetic analysis. Unrooted trees were prepared by the neighbor-joining method using either Clustal, PHYLIP, or and 1000 bootstrap replicates were performed. Bold lines on trees indicate protein sequences that were confirmed by cDNA sequencing. In this work, we aimed to reveal phylogenetic distances across the species using experimental values, rather than sequence information in the graphs. Hence we used the data of COX1 and COX2 experimental values. In the relation network, enzymes and genes are represented as nodes, while the substrate and product compounds are represented as edges. The related structural information from the graphs was used for computing phylogenetic distances.

## RESULTS AND DISCUSSION

### Collection of COX1 and COX-2 sequence from data base

Using NCBI and UNIPROT, we have collected COX-1 and COX-2 sequences from *Mus musculus* for phylogenetic analysis using the algorithm and treeview softwares. The developed algorithm was used to align the different cox-1 sequences, and the aligned sequences were converted into phylogenetic tree. With the availability of the data in GenBank, it was possible to construct an overview of COX-1 and COX-2 in *Mus musculus*. As a starting point, the protein families in COX-1 and COX-2 which have positive molecular implications on pain transport, intracellular targeting and storage in *Mus Musculus*, were chosen for analysis. Taking specific members of these families as query sequences, searches were carried out for orthologous sequences in GenBank, Mouse database and Uniprot current databases using TBLASTN. After searching the databanks with TBLASTN sequences, clones having genomic sequences to the related family were taken and converted to amino acid sequences. In each family, similar sequences were removed and the sequences were subjected to PROSITE and Pfam databases to see the presence of signature sequences for the corresponding families. After subjecting the sequences to PROSITE 19 new putative genes, 10 characterized genes, 9 COX family like genes were predicted in COX-1. The percent identity for all the sequences was calculated in each family with the corresponding query sequence using GENEDOC. Phylogenetic analysis of the sequences of transporters revealed that the COX proteins were divergent, showing branches in tree view. The phylogenetic analysis shows four branches indicating different transporting function to each family. Some of the orthologous sequences are available as full-length cDNA clones. The expressed sequence tags were mentioned as accession numbers for the sequences.

### SEQUENCE ALIGNMENT

These sequences were aligned using CLUSTALW program. The multiple alignments were used to develop phylogenetic tree.

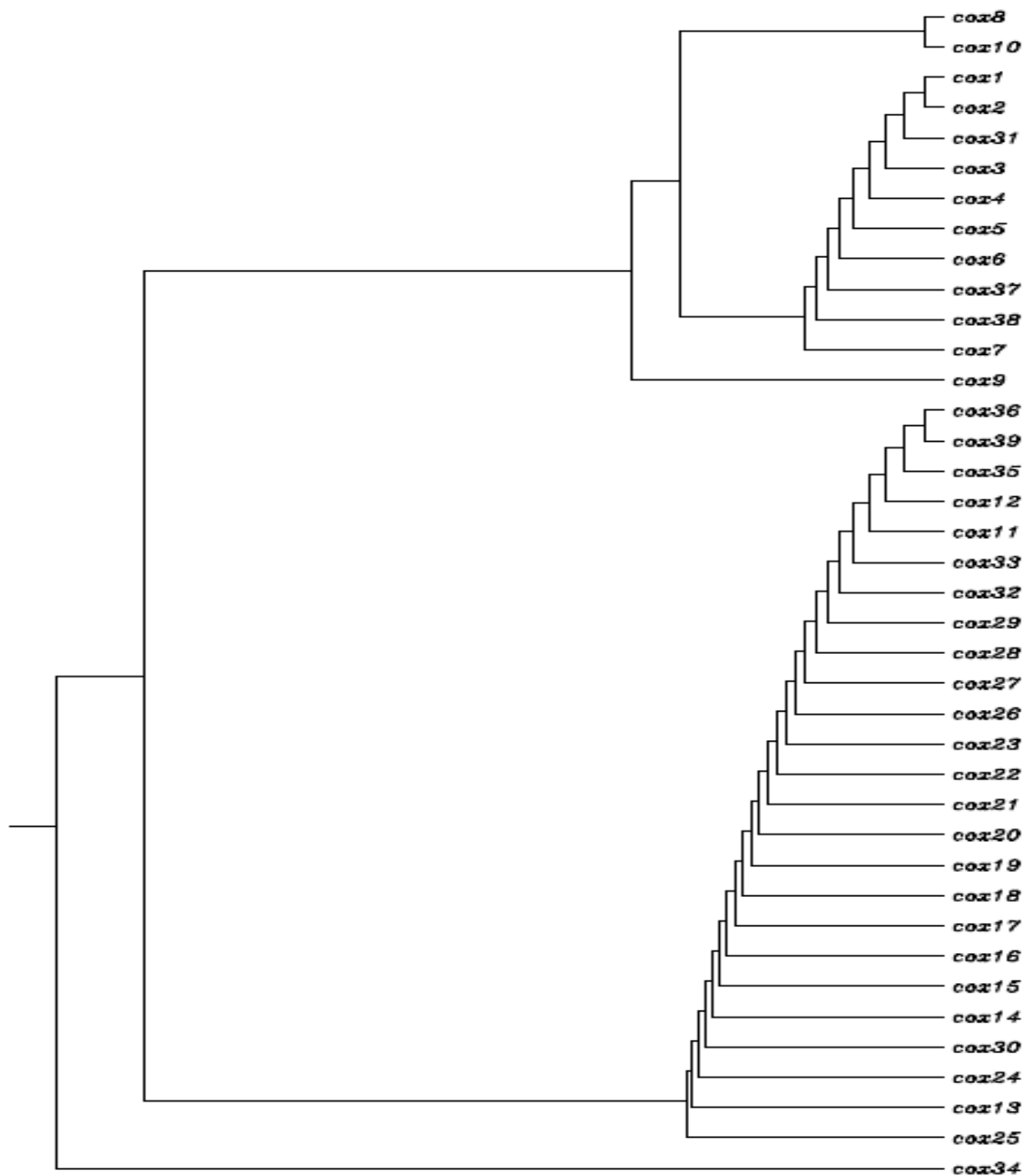


Fig 1: ClustalW result: Rooted Phylogenetic Tree with branch length (UPGMA) of Cox1

Cox2 Protein sequence

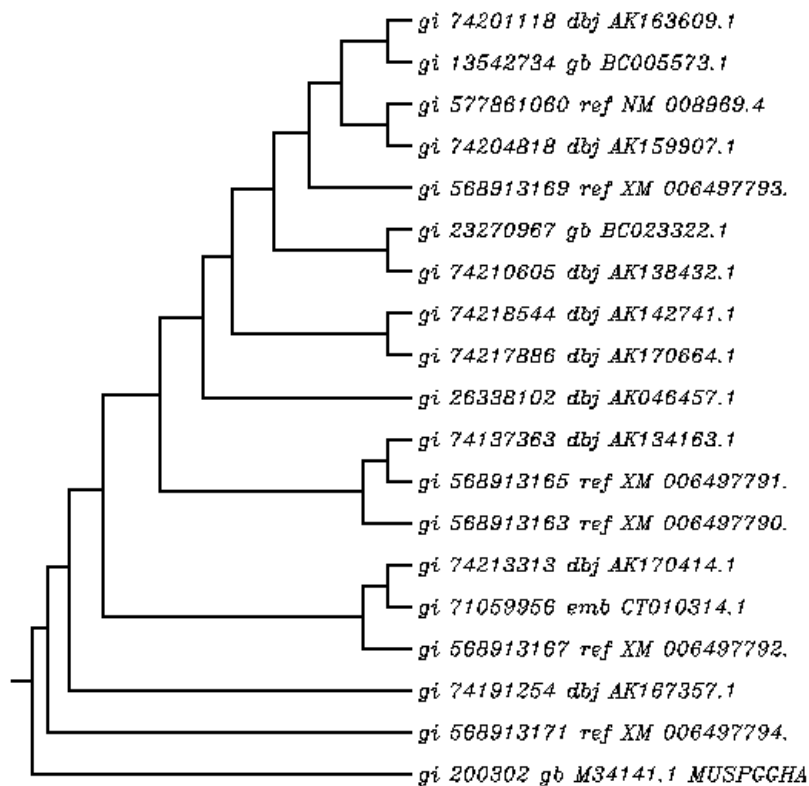
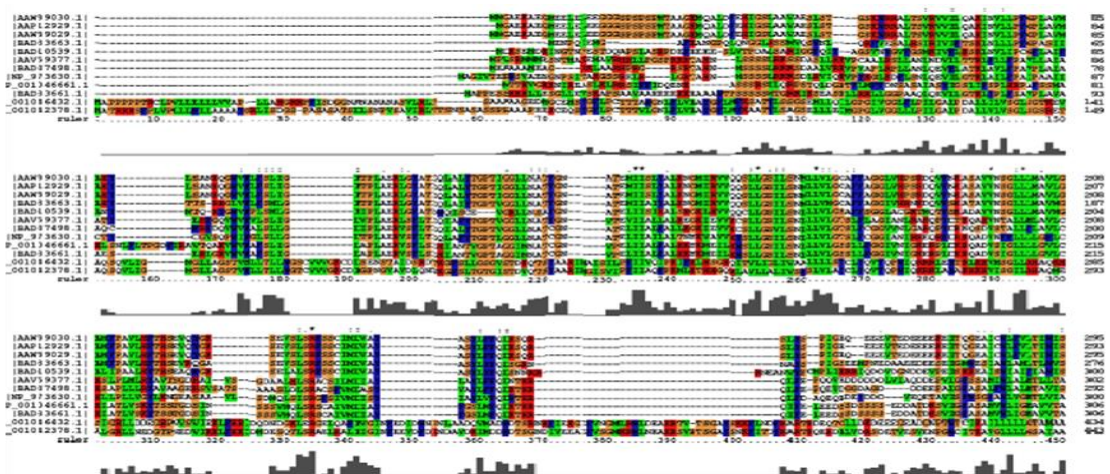


Fig 2: ClustalW result: Rooted Phylogenetic Tree with branch length (UPGMA) of Cox2

COX-1 sub family

COX-1 subfamily1 is an important family related to COX transporters. Nearly eleven genes related to COX-1 proteins were identified and named as COX1a. These are numbered based on their alignment. The genes, which are showing more than 30% identity with query, are said to be homologous sequences.



PDF Created with deskPDF PDF Writer - Trial :: http://www.docudesk.com

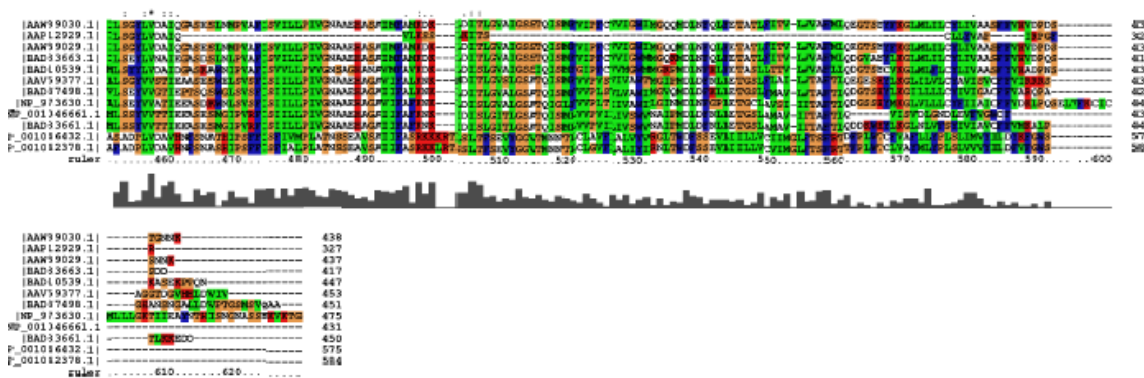
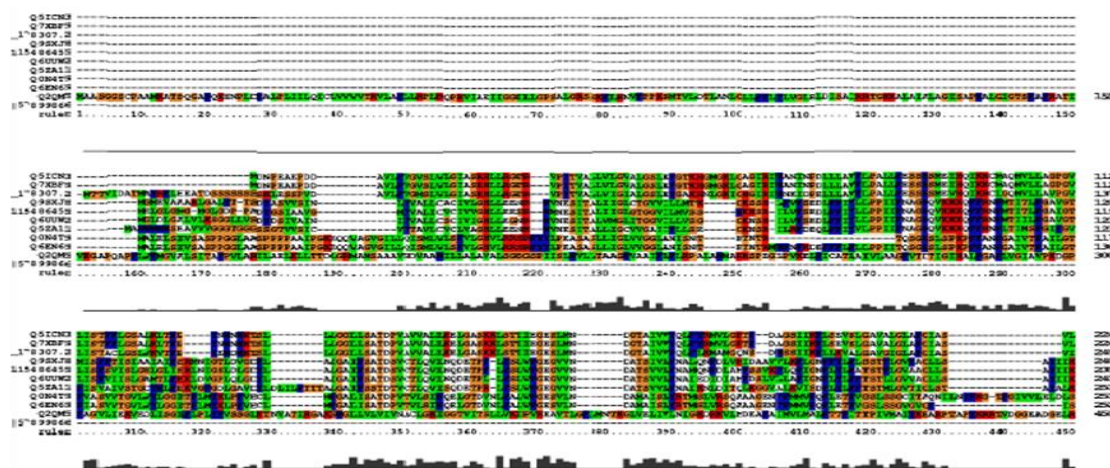


Fig. 3: Phylogenetic tree of *Mus musculus* COX-1. Programs used were ClustalX (Thompson et al., 1997) for alignments, and Treeview (Page, 1996) for graphical output. Values indicate the number of times (in %) that each branch topology was found during bootstrap analysis. Scale bar corresponds a distance of 10 changes per amino acid positions.

### COX-1 sub family2

COX-1 subfamily 2 is the monovalent, which is the largest gene family. Nearly twenty-nine genes were identified and named, as COX-1b. These are also numbered based on the alignment with reference protein of *Mus musculus*.



PDF Created with DeskPDF PDF Writer - Trial :: <http://www.docdesk.com>

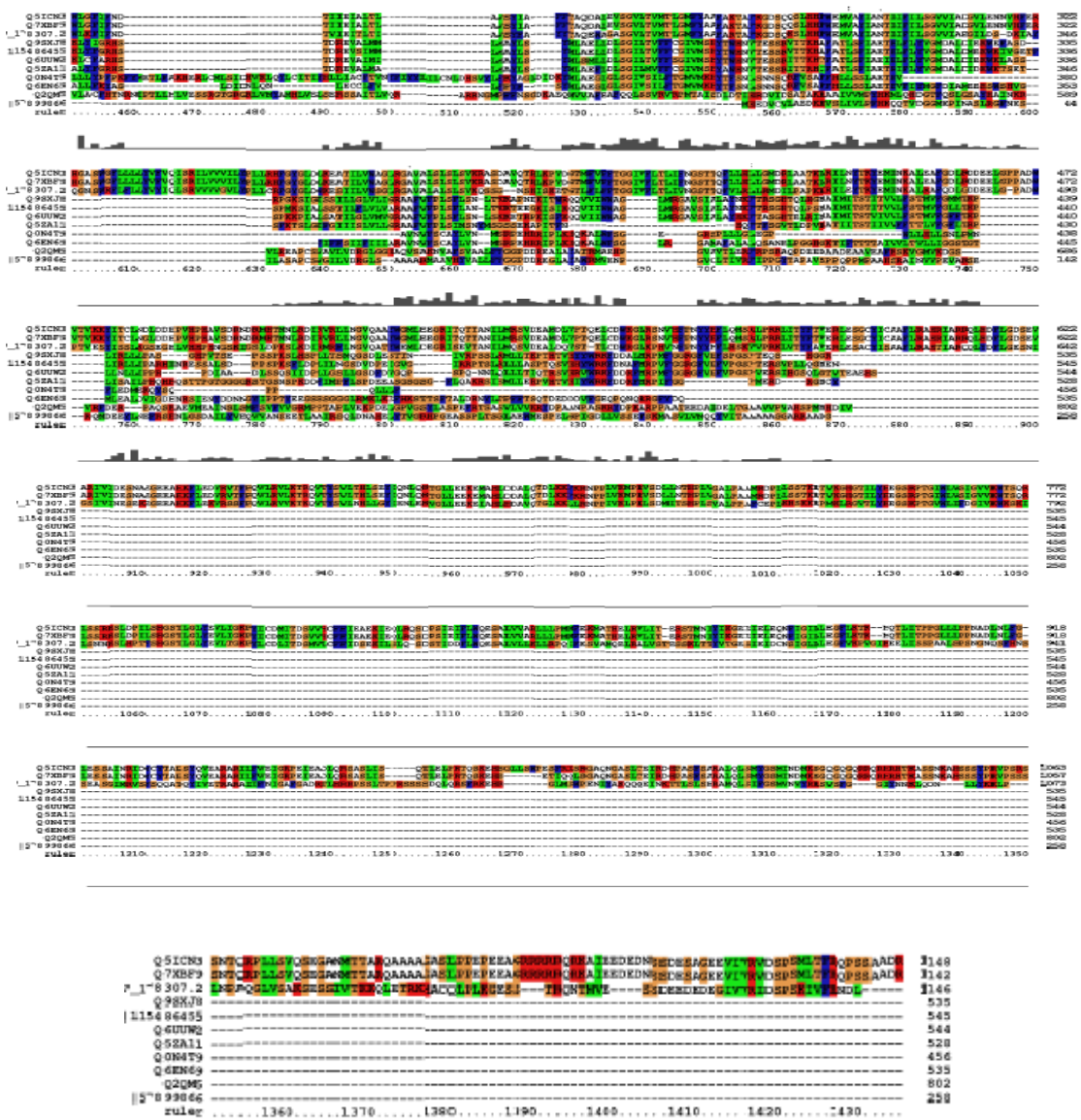
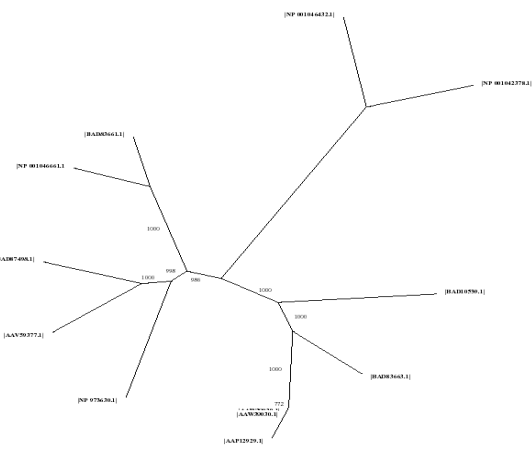


Fig. 4: Phylogenetic tree of *Mus musculus* COX-1b. Programs used were ClustalX (Thompson et al., 1997) for alignments, and Treeview (Page, 1996) for graphical output. Values indicate the number of times (in %) that each branch topology was found during bootstrap analysis. Scale bar corresponds a distance of 10 changes per amino acid positions.



### COX-1 subfamily 3

Smallest gene family in the COX-1 family. Only three genes were identified in this family. The genes, which have more than 30% identity with query, are regarded as homologous sequences.



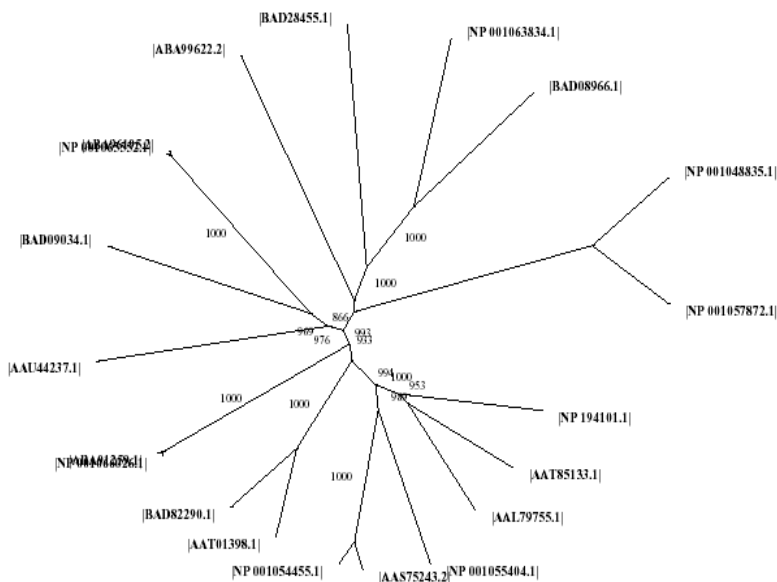
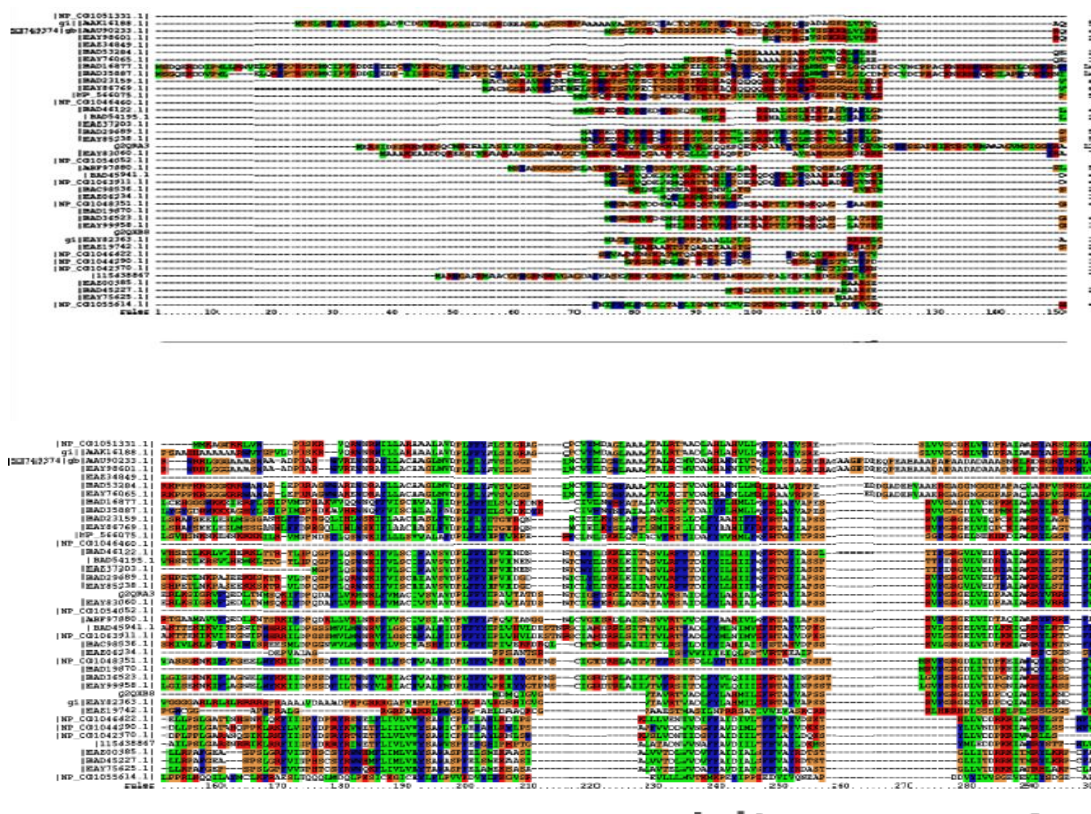


Fig. 5: Phylogenetic tree of *Mus musculus* COX-1c. Programs used were ClustalX (Thompson et al., 1997) for alignments, and Treeview (Page, 1996) for graphical output. Values indicate the number of times (in %) that each branch topology was found during bootstrap analysis. Scale bar corresponds a distance of 10 changes per amino acid positions.

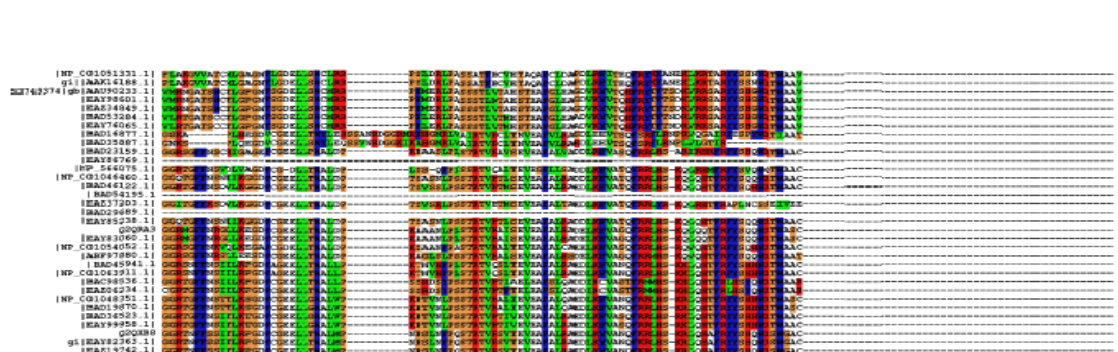
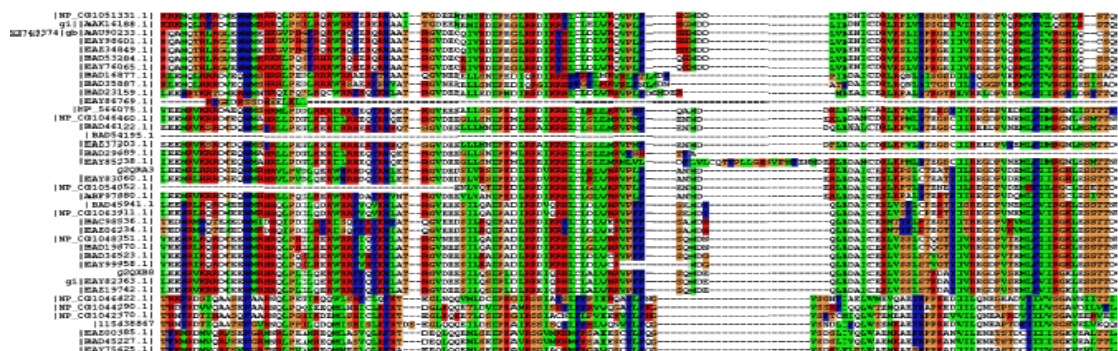
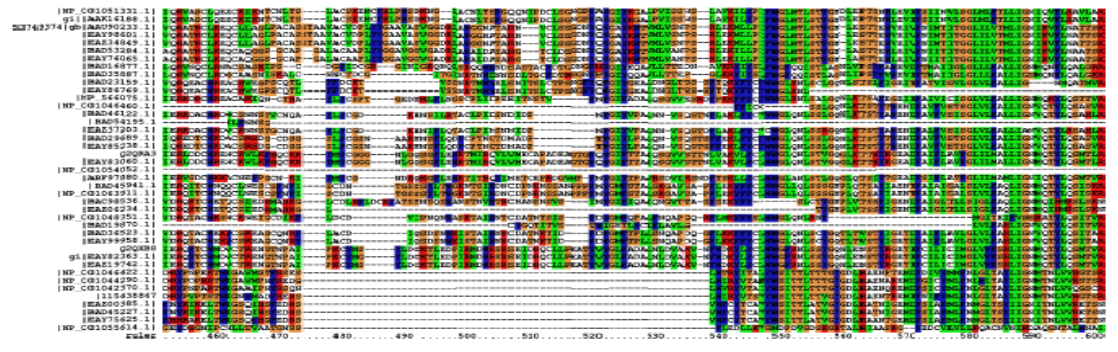
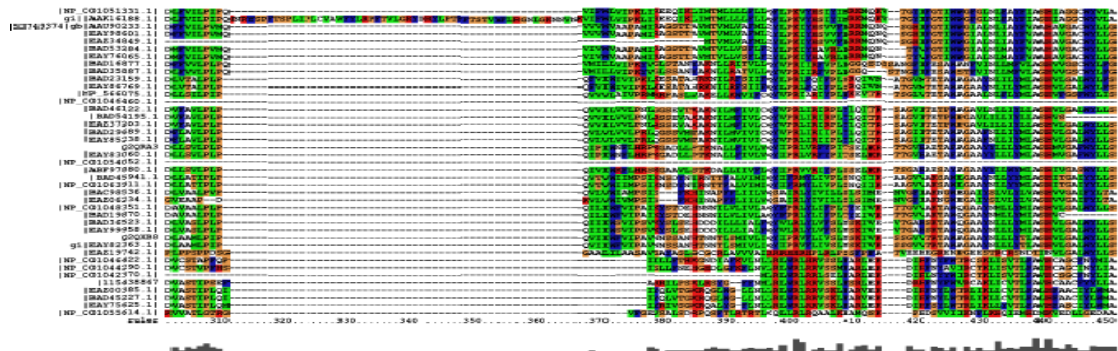
**COX-2 subfamily 1**

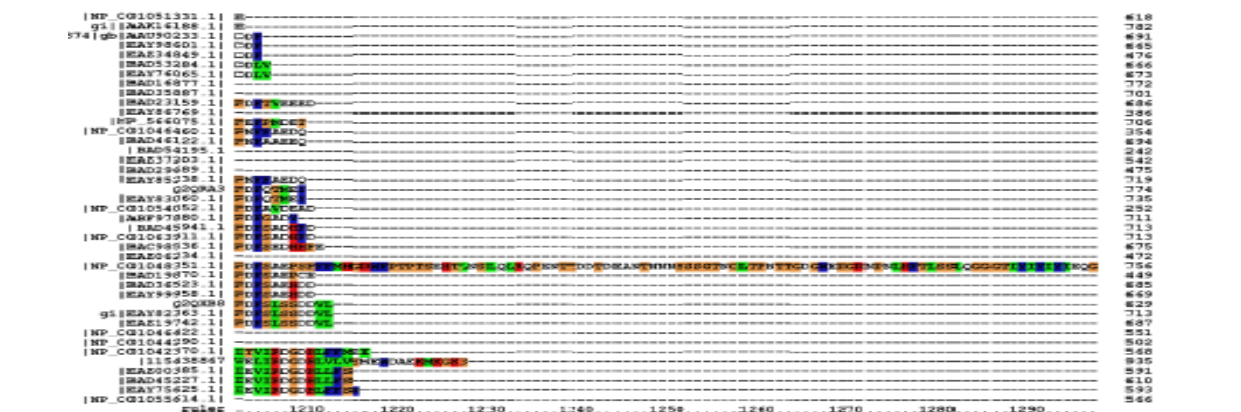
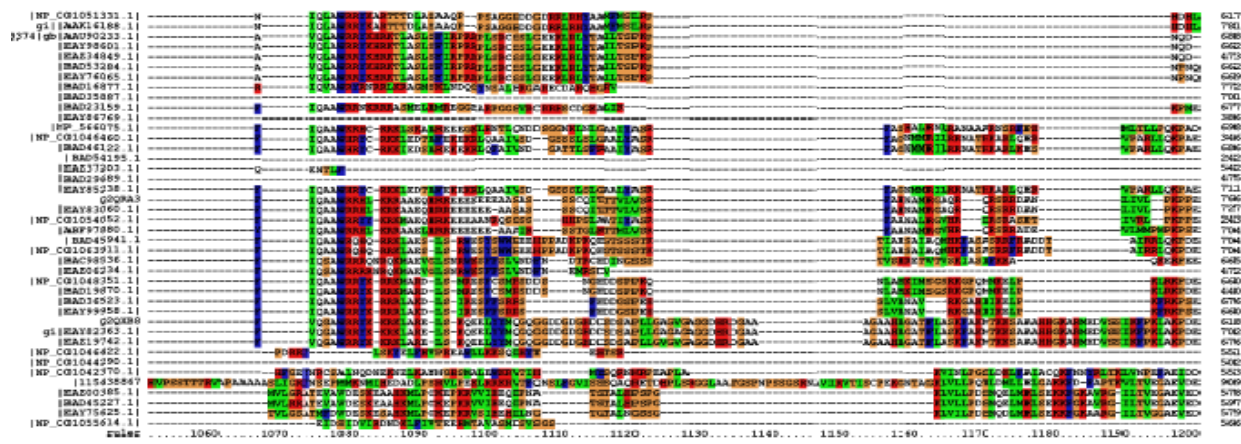
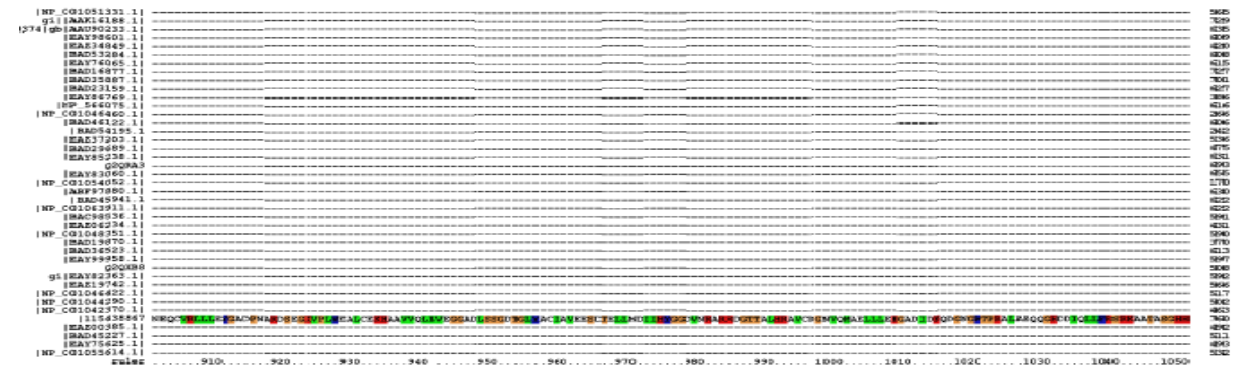
COX-2 subfamily 1 contains a total of forty-two genes and these are named as COX-2a. These are numbered based on the alignment of the gene with query. The genes, which are showing more than 30% identity, are homologous sequences.



PDF Created with deskPDF PDF Writer - Trial :: <http://www.docudesk.com>







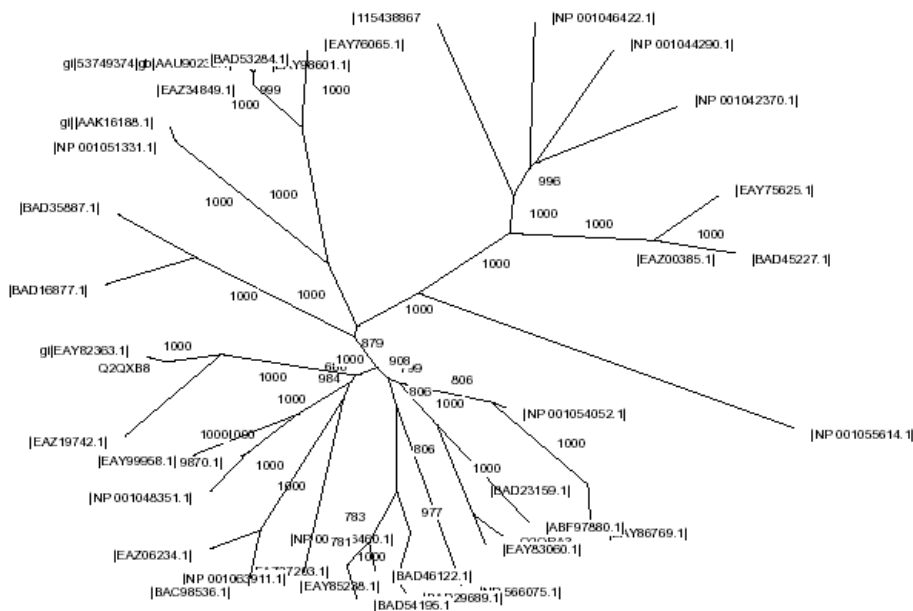
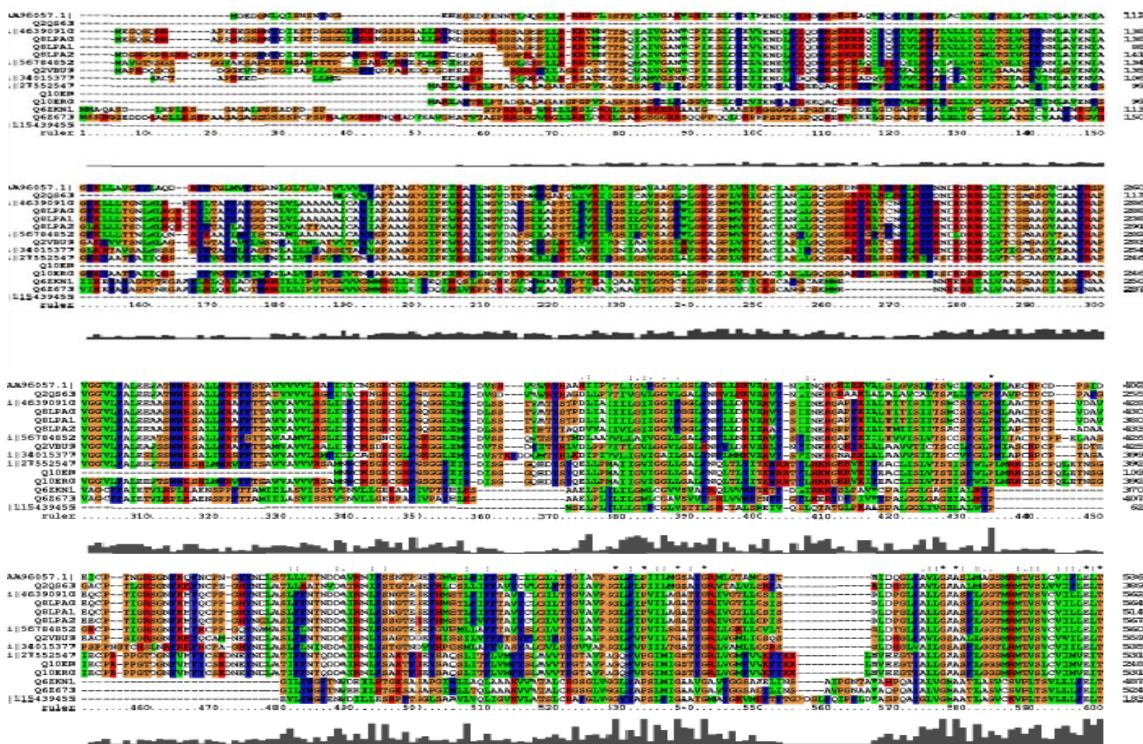
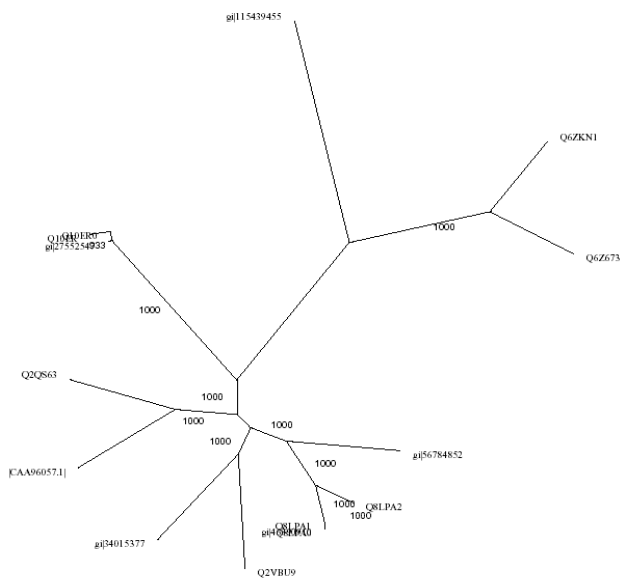
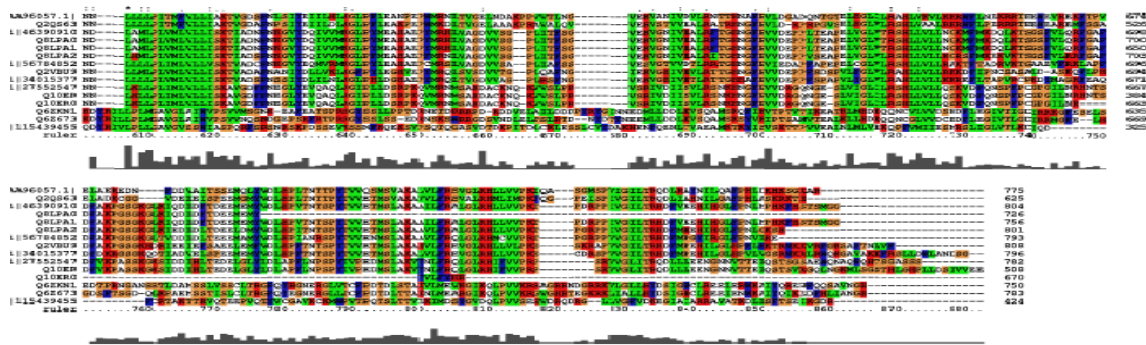


Fig. 6: Phylogenetic tree of *Mus musculus* COX-2. Amino acids in putative transmembrane fragments are shadowed, and amino acids, which are conserved in most sequences, are highlighted. Alignments were made using the ClustalX program.

### COX-2 subfamily 2

COX-2 subfamily contains nearly fourteen genes were identified. These are named as COX-2b and numbered based on alignment with query.





**Fig. 7: Phylogenetic tree of Mus musculus COX-2b. Programs used were Clustalw (Thompson et al., 1997) for alignments, and Treeview (Page, 1996) for graphical output. Values indicate the number of times (in %) that each branch topology was found during bootstrap analysis. Scale bar corresponds a distance of 10 changes per amino acid positions.**

**CONCLUSION**

In this work, we have analyzed the phylogenetic relationships of COX-1 and COX-2 proteins. This analysis has focused on COX gene families for which initial characterizations have been achieved for individual members, including COX-1sub family1, sub family2 and subfamily3, in COX-2 subfamily1 and subfamily2. Phylogenetic trees of each family define the evolutionary relationships of the members to each other. These families contain numerous members, indicating diverse functions in vivo. Closely related isoforms and separate subfamilies exist within many of these gene families, indicating possible redundancies and specialized functions. To facilitate their further study, which includes alignment of the analyzed genes.

**REFERENCES**

[1] Smith WL, DeWitt DL, Garavito RM. 2000. Cyclooxygenases: structural, cellular, and molecular biology. *Annu Rev Biochem.* 69:145–182.  
 [2] Kam PC, See AU. 2000. Cyclo-oxygenase isoenzymes: physiological and pharmacological role. *Anaesthesia.*55:442–449.  
 [3] Katori M, Majima M. 2000. Cyclooxygenase-2: its rich diversity of roles and possible application of its selective inhibitors. *Inflamm Res.* 49:367–392.

- [4] Langenbach R, Loftin CD, Lee C, Tiano H. 1999. Cyclooxygenase-deficient mice. A summary of their characteristics and susceptibilities to inflammation and carcinogenesis. *Ann N Y Acad Sci.* 889:52–61.
- [5] Bosetti F, Langenbach R, Weerasinghe GR. 2004. Prostaglandin E2 and microsomal prostaglandin E synthase-2 expression are decreased in the cyclooxygenase-2-deficient mouse brain despite compensatory induction of cyclooxygenase-1 and Ca<sup>2+</sup>-dependent phospholipase A2. *J Neurochem.*91:1389–1397.
- [6] Choi SH, Langenbach R, Bosetti F. 2006. Cyclooxygenase-1 and -2 enzymes differentially regulate the brain upstream NF-kappaB pathway and downstream enzymes involved in prostaglandin biosynthesis. *J Neurochem.* 98:801–811.
- [7] Rao JS, Langenbach R, Bosetti F. 2005. Down-regulation of brain nuclear factor-kappa B pathway in the cyclooxygenase-2 knockout mouse. *Brain Res Mol Brain Res.* 139:217–224.
- [8] Feng Z, Li D, Fung PC, Pei Z, Ramsden DB, Ho SL. 2003. COX-2-deficient mice are less prone to MPTP-neurotoxicity than wild-type mice. *NeuroReport.* 14:1927–1929.
- [9] Iadecola C, Niwa K, Nogawa S, Zhao X, Nagayama M, Araki E, Morham S, Ross ME. 2001. Reduced susceptibility to ischemic brain injury and N-methyl-D-aspartate-mediated neurotoxicity in cyclooxygenase-2-deficient mice. *Proc Natl Acad Sci USA.* 98:1294–1299.
- [10] Vane JR, Bakhle YS, Botting RM. 1998. Cyclooxygenases 1 and 2. *Annu Rev Pharmacol Toxicol.* 38:97–120.
- [11] Zou J, Neumann NF, Holland JW, Belosevic M, Cunningham C, Secombes CJ, Rowley AF. (1999). Fish macrophages express a cyclo-oxygenase-2 homologue after activation. *Biochem J.* 340:153–159.
- [12] Roberts SB, Langenau DM, Goetz FW. 2000. Cloning and characterization of prostaglandin endoperoxide synthase-1 and -2 from the brook trout ovary. *Mol Cell Endocrinol.* 160:89–97.
- [13] Choe KP, Havird J, Rose R, Hyndman K, Piermarini P, Evans DH. 2006. COX2 in a euryhaline teleost, *Fundulus heteroclitus*: primary sequence, distribution, localization, and potential function in gills during salinity acclimation. *J Exp Biol.* 209:1696–1708.
- [14] DeWitt DL, Smith WL. 1988. Primary structure of prostaglandin G/H synthase from sheep vesicular gland determined from the complementary DNA sequence. *Proc Natl Acad Sci USA.* 85:1412–1416.
- [15] Merlie JP, Fagan D, Mudd J, Needleman P. 1988. Isolation and characterization of the complementary DNA for sheep seminal vesicle prostaglandin endoperoxide synthase (cyclooxygenase). *J Biol Chem.* 263:3550–3553.

# International Journal of Plant, Animal and Environmental Sciences

