

Reconstruction of Cyclooxygenase Evolution in Animals Suggests Variable, Lineage-Specific Duplications, and Homologs with Low Sequence Identity

Justin C. Havird^{1,2} · Kevin M. Kocot^{1,3} · Pamela M. Brannock¹ ·
Johanna T. Cannon^{1,4} · Damien S. Waits¹ · David A. Weese^{1,5} ·
Scott R. Santos¹ · Kenneth M. Halanych¹

Received: 14 November 2014 / Accepted: 2 March 2015 / Published online: 11 March 2015
© Springer Science+Business Media New York 2015

Abstract Cyclooxygenase (COX) enzymatically converts arachidonic acid into prostaglandin G/H in animals and has importance during pregnancy, digestion, and other physiological functions in mammals. COX genes have mainly been described from vertebrates, where gene duplications are common, but few studies have examined COX in invertebrates. Given the increasing ease in generating genomic data, as well as recent, although incomplete descriptions of potential COX sequences in Mollusca, Crustacea, and Insecta, assessing COX evolution across Metazoa is now possible. Here, we recover 40 putative COX orthologs by searching publicly available genomic resources as well as ~250 novel invertebrate transcriptomic datasets. Results suggest the common ancestor of Cnidaria and Bilateria

possessed a COX homolog similar to those of vertebrates, although such homologs were not found in poriferan and ctenophore genomes. COX was found in most crustaceans and the majority of molluscs examined, but only specific taxa/lineages within Cnidaria and Annelida. For example, all octocorallians appear to have COX, while no COX homologs were found in hexacorallian datasets. Most species examined had a single homolog, although species-specific COX duplications were found in members of Annelida, Mollusca, and Cnidaria. Additionally, COX genes were not found in Hemichordata, Echinodermata, or Platyhelminthes, and the few previously described COX genes in Insecta lacked appreciable sequence homology (although structural analyses suggest these may still be functional COX enzymes). This analysis provides a benchmark for identifying COX homologs in future genomic and transcriptomic datasets, and identifies lineages for future studies of COX.

Data deposition GenBank Accession Numbers KM437900–KM437941.

Electronic supplementary material The online version of this article (doi:10.1007/s00239-015-9670-3) contains supplementary material, which is available to authorized users.

✉ Justin C. Havird
justin.havird@colostate.edu

¹ Department of Biological Sciences & Molette Biology
Laboratory for Environmental and Climate Change Studies,
Auburn University, Auburn, USA

² Present Address: Department of Biology, Colorado State
University, Fort Collins, CO, USA

³ Present Address: School of Biological Sciences, The
University of Queensland, Brisbane, Australia

⁴ Present Address: Department of Zoology, Naturhistoriska
riksmuseet, Stockholm, Sweden

⁵ Present Address: Department of Biological and
Environmental Sciences, Georgia College and State
University, Milledgeville, USA

Keywords Prostaglandin synthesis · Transcriptomics ·
Phylogenetics · Metazoa · Gene duplication and loss ·
Annelida

Introduction

Cyclooxygenase (COX; aka prostaglandin G/H synthase, PGHS; prostaglandin-endoperoxide synthase) is responsible for converting arachidonic acid to prostaglandin G and subsequently prostaglandin H in vertebrates (Simmons et al. 2004). COX is part of the larger myeloperoxidase superfamily, which includes myeloperoxidase, eosinophil peroxidase (EPO), peroxidase, peroxinectin, and other enzymes with a ~500 amino acid peroxidase domain (Daiyasu and Toh 2000). COX is further characterized by an N-terminal signal peptide, the EGF-like domain, the

membrane-binding domain (MBD), and the catalytic domain with distinct peroxidase and COX catalytic sites (Kulmacz et al. 2003; Simmons et al. 2004; Ishikawa et al. 2007; Havird et al. 2008), some of which may be missing or incomplete in organisms that are distantly related to vertebrates (Kanamoto et al. 2011). After conversion via COX, prostaglandins (PGs) go on to participate in autocrine and paracrine signaling in mammals, contributing to a wide range of physiological functions including stimulating the inflammation response, altering gastric acid secretion, and inducing labor during pregnancy (see Vane et al. 1998 for a review). Accordingly, drugs targeting COX (e.g., aspirin and other non-steroidal anti-inflammatory drugs) are among the most heavily prescribed worldwide (Singh and Triadafilopoulos 1999).

Although COX has been extensively characterized for mammals, its roles in, and distributions among, non-mammalian animals are less well known. For example, COX in teleosts appears to be involved in osmoregulation (Choe et al. 2006), reproduction (Sorbera et al. 2001; Flippin et al. 2007), and circadian rhythm (Paredes et al. 2014). COX genes have also been characterized through targeted sequencing efforts in relatively early-branching lineages of chordates including lamprey, hagfish, and cephalochordates (Havird et al. 2008). In invertebrates, COX has been described and biochemically characterized in regards to PG synthesis in two amphipod crustaceans (*Gammarus sp.* and *Caprella sp.*) and two soft coral species (*Gersemia fruticosa* and *Plexaura homomalla*) (Koljak et al. 2001; Valmsen et al. 2001, 2004; Varvas et al. 2009). While COX sequences have also been found in genomic databases of the crustacean *Daphnia pulex* and in two molluscs (the oyster *Crassostrea gigas* and the mussel *Mytilus edulis*), searches for COX sequences in other invertebrates have been less successful. For example, although putative COX sequences (sharing 30 % sequence identity with human COX) have been identified in the genomes of the body louse (*Pediculus humanus*) and pea aphid (*Acyrtosiphon pisum*) (Kawamura et al. 2014), genomes of several other insects (e.g., the fly *Drosophila*, the mosquito *Aedes aegypti*, and the honeybee *Apis mellifera*) apparently lack COX homologs based on sequence homology (Varvas et al. 2009; Kawamura et al. 2014). Finally, a functional COX homolog has been described from one non-metazoan lineage, the red alga *Gracilaria vermiculophylla* (Kanamoto et al. 2011; Varvas et al. 2013). Interestingly, the red alga COX lacks sequence homology to other known COX genes (28 % sequence identity with human COX) and is missing several COX functional domains and residues, including one of three helices from the membrane-binding domain, the glycosylation sites, and the aspirin acetylation site present in all vertebrate COX proteins (Kanamoto et al. 2011). However, it maintains structural homology to other COXs and has been experimentally shown to convert arachidonic acid to PGs (Varvas et al. 2013).

While COX has apparently been lost or has become highly diverged at the sequence level in some animals (Kawamura et al. 2014), it has also undergone duplications and potential functional diversification. For instance, multiple copies of COX genes exist in nearly every chordate lineage examined to date (Havird et al. 2008). As first described in mammals, and later for other families, there are two COX paralogs in vertebrates: COX-1 and COX-2, which have functional specializations (Vane et al. 1998). Phylogenetic assessment of COX in chordates suggested the duplication event leading to COX-1 and COX-2 likely occurred within Craniata, as lamprey and hagfish COX sequences are most likely homologs of vertebrate COX-1/2 (Havird et al. 2008). However, independent COX duplications have occurred in both Tunicata (=Urochordata) and Cephalochordata, resulting in genes named COXa/b and COXc/d, respectively (Havird et al. 2008). Lastly, although amphipod COX was found in single copy in the two species examined (Varvas et al. 2009), multiple copies in the Arctic soft coral *G. fruticosa* (named COX-A/B; Järving et al. 2004) imply duplicate and diversified copies of COX may be common and represent the ancestral condition among animals (Kawamura et al. 2014).

To date, COX genes in many lineages are only known from a few exemplar sequences (e.g., Mollusca and Crustacea) and a number of major invertebrate lineages (e.g., Annelida) remain to be surveyed for these genes. Broader taxonomic sampling (if present) and analyses of COX amongst animals could address questions such as: Was a COX enzyme similar to the one found in vertebrates present in the last common metazoan ancestor? Which invertebrate lineages possess COX homologs? Are COX duplications common or do COX genes tend to be single copy? Fortunately, the increase in publically available genomic and transcriptomic resources representing a wide swath of animal diversity, as well as the ability to generate such data rapidly and economically, now makes it possible to shed light on these questions. Here, we screened publically available and novel genomic and transcriptomic data representing all major animal lineages for COX homologs in order to generate a phylogenetic hypothesis for its evolution across Metazoa. Based on this phylogeny, a discussion is presented highlighting lineage-specific duplications, the possibility of COX homologs with low-sequence homology in some lineages, and the likely origin of a vertebrate-like COX in metazoans.

Materials and Methods

Data Acquisition—Publically Available Data

The basic local alignment search tool (BLAST, Altschul et al. 1990) was used to screen publically available genomic and transcriptomic datasets (Online Resource 1) for

COX homologs. Remote BLASTs were performed for genomic datasets either via the Ensembl Genome Browser (Flicek et al. 2014), NCBI's Genome BLAST, or to the Joint Genome Institute (JGI). Specifically, BLASTp searches were performed by utilizing previously described COX sequences as queries: COX-1b from mummichog (*Fundulus heteroclitus*, Chordata: Vertebrata, GenBank accession ACH73265.1), COX-a from sea squirt (*Ciona intestinalis*, Chordata: Urochordata, Ensembl accession ENSCINP00000013352), COXa1 from oyster (*Crassostrea gigas*, Mollusca, GenBank accession ACP28169.2), COX from the water flea (*Daphnia pulex*, Arthropoda, GenBank accession EFX85708.1), COX-A from Arctic soft coral (*Gersemia fruticosa*, Cnidaria, GenBank accession AAF93168.1), and COX (15R specific) from sea whip (*Plexaura homomalla*, Cnidaria, GenBank accession AAF93169.1). This panel, which encompasses the known phylogenetic diversity of animal lineages possessing COX, was assembled to maximize the probability that COX homologs spanning varying levels of divergence would be recovered. BLAST searches were performed in a taxon-directed manner across all major metazoan lineages and complemented cases where new genetic resources were generated (see below). Publically available transcriptomic data were either downloaded as raw reads and assembled de novo (as for soft corals, see below) or as assembled sets of contigs (i.e., hard corals) and then searched via BLAST as above. The approximately top five (5) BLAST “hits” with e-values of 10^{-20} or less were retained from each search and used in preliminary phylogenetic analyses with previously characterized COX sequences to determine if candidates were in fact members of COX or closely related gene families (see [Sequence alignment and phylogenetic analyses](#) section).

In addition to the above searching of genomic and transcriptomic datasets, previously described COX sequences (Havird et al. 2008 and Table 1 of Kawamura et al. 2014) were also included in subsequent phylogenetic analyses. GenBank's protein (i.e., nr) database was also searched for COX sequences, with ones from chordate lineages being discarded (except for lineages of special interest such as coelacanth and Epaulette shark) and non-chordate lineages retained.

Data Acquisition—Novel Data

Novel transcriptomic data, generated from a range of invertebrate lineages and covering ~250 taxa (Online Resource 2), were also searched for COX homologs using either tBLASTn (which is based on translated protein similarity; Altschul et al. 1990) as described above or by searching for several specific COX-related terms (e.g., cyclooxygenase, PG, COX, PGHS, and arachidonic acid)

among annotation records of assembled contigs. While these transcriptomes will be described in their entirety elsewhere, methods for RNA extraction, cDNA library generation and sequencing as well as assembly and contig annotation generally followed those for previously described molluscan, crustacean, and hemichordate transcriptomes (Kocot et al. 2011; Genomic Resources Development Consortium et al. 2014; Cannon et al. 2014). When searching novel transcriptomes, top BLAST hits with e-values of 10^{-20} or less were retained from each transcriptome for preliminary phylogenetic analyses. A listing of transcriptomes from which COX candidates was queried is given in Online Resource 2.

Sequence Alignment and Phylogenetic Analyses

Amino acid sequences from COX candidates identified from BLAST searches were aligned with previously characterized COX sequences taken from Havird et al. (2008) and Kawamura et al. (2014) using MUSCLE (Edgar 2004) as implemented in SeaView version 4 with default values (Gouy et al. 2010). Poorly aligned and/or divergent regions of resulting alignments were trimmed using Gblocks version 0.91b (Castresana 2000) with the following parameters: the minimum number of 50 % of sequences was selected to identify conserved and flanking positions, maximum number of contiguous non-conserved sequences in a block set to 10, and gap positions allowed in all blocks. This approach resulted in final alignments of ~450 amino acids (roughly 25 % of amino acids were discarded). Preliminary phylogenetic trees were then generated from trimmed alignments using the parallelized form of FastTreeMP version 2.1.7 (Price et al. 2010). FastTreeMP was used in preliminary analyses instead of PhyML or RAxML because these latter programs were too computationally intensive for the numerous preliminary analyses that were performed (Price et al. 2010). Thirteen sequences from major animal lineages that were not COX homologs, but instead members of the related gene families myeloperoxidase (MPO) and peroxidase (PXDN), were retained as outgroups, consistent with a previous phylogenetic analysis indicating MPO and PXDN sequences are appropriate outgroups for COX (Daiyasu and Toh 2000). If candidate sequences clustered with previously characterized COX genes, they were retained as possible homologs, while sequences were discarded if they clustered with MPO/PXDN sequences to the exclusion of the COX clade. This approach was iterated several times (as opposed to one large, computationally intensive preliminary analysis) for subsets of candidate sequences (inferred via BLAST) and known COX homologs (totaling ~200 sequences per subset), with candidate sequences either being discarded or retained as potentially “true” homologs for each subset.

Table 1 Novel cyclooxygenase (COX) and peroxidasin (PXDN) sequences analyzed in the current study

Sequence name	Organism	GenBank Accession #	Size (bp)	Size (AA)	Trinity assembly contig(s)
Crustacea					
COX Blue crab	<i>Callinectes sapidus</i>	KM437921	3096	604	comp33861_c0_seq3
COX Lobster	<i>Homarus americanus</i>	KM437916	2865	611	comp58381_c0_seq9
COX <i>Halocaridinides</i>	<i>Halocaridinides trigonophthalma</i>	KM437937	1255	418	composite of 3 contigs: comp55042_c0_seq1 comp56439_c0_seq1 comp56921_c0_seq1
COX Alpha opae ula	<i>Metabetaeus lohena</i>	KM437938	2264	352	composite of 2 contigs: comp48128_c0_seq1 comp78816_c0_seq1
COX Opae ula	<i>Halocaridina rubra</i>	KM437920	2787	616	comp40632_c0_seq1
COX Lauan anchialine shrimp	<i>Antecaridina lauensis</i>	KM437919	1931	628	comp24395_c0_seq5
Annelida					
Putative COXa <i>Aspidosiphon</i>	<i>Aspidosiphon laevis</i>	KM437924	2043	578	comp66818_c0_seq1
Putative COXb <i>Aspidosiphon</i>	<i>Aspidosiphon laevis</i>	KM437925	575	170	comp844087_c0_seq1
Putative COX <i>Lithacrosiphon</i>	<i>Lithacrosiphon cristatus</i>	KM437923	2019	581	comp55520_c0_seq1
^a Putative COX <i>Phascolosoma</i>	<i>Phascolosoma agassizii</i>	KM437926	1921	392	comp21202_c0_seq1
COX <i>Arabella</i>	<i>Arabella sp.</i>	KM437918	2122	595	comp91133_c0_seq1
COXa <i>Polygordius</i>	<i>Polygordius sp.</i>	KM437915	683	227	comp135172_c0_seq1
COXb <i>Polygordius</i>	<i>Polygordius sp.</i>	KM437904	1117	372	comp63333_c0_seq1
COX Rock worm	<i>Eunice pennata</i>	KM437922	1327	441	comp32265_c0_seq1
COX <i>Paramphinome</i>	<i>Paramphinome jeffreysii</i>	KM437927	412	137	comp318435_c0_seq1
^a COX <i>Leitoscoloplos</i>	<i>Leitoscoloplos robustus</i>	KM437929	796	172	comp42677_c0_seq1
COX <i>Histriobdella</i>	<i>Histriobdella homari</i>	KM437917	1395	465	comp60317_c0_seq2
COXa <i>Phyllochaetopterus</i>	<i>Phyllochaetopterus prolifica</i>	KM437906	848	282	comp333845_c0_seq1
COXb <i>Phyllochaetopterus</i>	<i>Phyllochaetopterus prolifica</i>	KM437905	2180	588	comp82868_c0_seq3
^a COX <i>Protodriloides</i>	<i>Protodriloides symbioticus</i>	KM437941	601	139	comp157901_c0_seq1
COX <i>Mesochaetopterus</i>	<i>Mesochaetopterus taylora</i>	KM437907	2218	595	comp35185_c0_seq1
Nemertea					
^a COX Red ribbon worm	<i>Tubulanus polymorphus</i>	KM437930	2362	635	comp39895_c0_seq1
Brachiopoda					
^a COX <i>Novocrania</i>	<i>Novocrania anomala</i>	KM437928	369	122	comp59745_c0_seq1
Mollusca					
COXa <i>Prochaetoderma</i>	<i>Prochaetoderma californicum</i>	KM437908	2231	478	comp154570_c0_seq6
COXb <i>Prochaetoderma</i>	<i>Prochaetoderma californicum</i>	KM437909	2243	413	comp154570_c0_seq7
^a COX <i>Simrothiella</i>	<i>Simrothiella margaritacea</i>	KM437900	2475	599	comp42267_c0_seq1
COX <i>Entonomenia</i>	<i>Entonomenia tricarinata</i>	KM437914	2206	626	comp65102_c0_seq1
COX <i>Epimienia</i>	<i>Epimienia babai</i>	KM437913	832	170	comp317419_c0_seq1
COXa <i>Proneomenia</i>	<i>Proneomenia custodiens</i>	KM437931	965	321	comp58035_c0_seq1
COXb <i>Proneomenia</i>	<i>Proneomenia custodiens</i>	KM437932	3594	635	comp71711_c0_seq7
COXa <i>Neomenia</i>	<i>Neomenia carinata</i>	KM437911	322	107	comp1041492_c0_seq1
COXb <i>Neomenia</i>	<i>Neomenia carinata</i>	KM437910	2597	603	comp90940_c0_seq1
COX <i>Helluoherpia</i>	<i>Helluoherpia aegiri</i>	KM437903	1571	477	comp25483_c0_seq1
COX <i>Macellomenia</i>	<i>Macellomenia schanderi</i>	KM437902	4280	608	comp37681_c0_seq1
COX <i>Meiomenia</i>	<i>Meiomenia swedmarki</i>	KM437901	3036	601	comp48139_c0_seq10
^a COX Antarctic <i>Aplacophora sp.</i>	<i>Antarctic Aplacophora sp.</i>	KM437912	437	145	comp12349_c0_seq1

Table 1 continued

Sequence name	Organism	GenBank Accession #	Size (bp)	Size (AA)	Trinity assembly contig(s)
Cnidaria					
COX Yellow sea fan	<i>Eunicella cavolini</i>	KM437933	2157	597	comp20797_c0_seq1
COX Pink sea fan	<i>Eunicella verrucosa</i>	KM437934	2247	597	comp16963_c0_seq1
COX Purple sea fan	<i>Gorgonia ventalina</i>	KM437935	2272	593	comp79078_c2_seq12
^a COX Orange sea fan	<i>Leptogorgia sarmentosa</i>	KM437936	604	174	comp101993_c0_seq1
Porifera					
PXDN-like <i>Sympagella</i>	<i>Sympagella nux</i>	KM437939	2151	716	comp22862_c0_seq1
PXDN-like <i>Rossella</i>	<i>Rossella fibulata</i>	KM437940	2973	929	comp13695_c0_seq1

^a Indicates “rogue” sequences removed from selected phylogenetic analyses

Candidate sequences passing the above screenings were then combined with previously described COX sequences (defined as the “complete dataset”) for additional phylogenetic analyses using Maximum Likelihood (ML) and Bayesian Inference (BI). ML analyses were performed using the pthreads implementation of RAxML version 8.0.24 (Stamatakis 2006; Ott et al. 2007; Stamatakis 2014), with runs (performed in duplicate) consisting of 1000 rapid bootstrap replicates (Stamatakis et al. 2008) using the PROTGAMMALG4X model of protein evolution (as specified with the -m flag). PROTGAMMALG4X uses the Γ model of rate heterogeneity (Yang 1996) and the LG4X model of protein substitution, which has been shown to significantly outperform single matrix substitution models (Le et al. 2012). BI analyses were performed with PhyloBayes version 3.3 (Lartillot and Philippe 2004; Lartillot et al. 2009) and utilized two runs of four chains to infer phylogeny based on the same trimmed amino acid alignment as in ML analyses. The CAT model of protein evolution was used (as an alternative to the LG4X model, which is not available in PhyloBayes), with 15,000 cycles (corresponding to ~15,000,000 generation) and a 10 % burn-in (when likelihood scores had reached stability). Maximum discrepancy (i.e., max diff) between bipartitions was 0.128, suggesting that independent chains had reached a similar, stable point in treespace, and the number of generations was “acceptable” to give a good overview of the posterior consensus (Lartillot and Philippe 2004). The ML and BI analyses were conducted at the Alabama Supercomputer Center (ASC) in Huntsville, Alabama. In an effort to encourage future studies, all sequences and alignments are available publically via <http://www.auburn.edu/~santost/sequencedatasets.htm>.

Based on the examination of preliminary trees, nodal support values, and phylogenetic affinities from the FastTreeMP analyses, “rogue” sequences, or ones decreasing statistical support for otherwise robust clades, were

suspected in the dataset. “Rogues”, which can be due to incomplete sequences (e.g., 53 % of novel COX sequences identified here were incomplete), incomplete taxon sampling, accelerated substitution rates leading to long branches, or the possible misclassification of non-COX sequences as COX homologs (Sanderson and Shaffer 2002), were identified via the RogueNaRok webservice (Aberer et al. 2013; <http://rnr.h-its.org/>) under default parameters. Additional ML and BI phylogenetic analyses were then conducted using RAxML and PhyloBayes under the same parameters as above with these sequences excluded (designated the “rogues removed” dataset).

Phylogenetic placement of cnidarian COX sequences was particularly interesting based on the final topology (see below). Given this, alternate topological placements of cnidarian COX sequences were compared to the most likely topology using Shimodaira-Hasegawa (SH) tests (Shimodaira and Hasegawa 1999) as implemented in RaxML via the -f h flag.

Characterization of Novel COX Sequences

To further characterize novel genes identified in this study, those sequences used in final phylogenetic analyses (i.e., the “complete dataset”) were searched for conserved protein domains, motifs, and residues characteristic of COX functionality. Domains were characterized using InterProScan version 5 (Jones et al. 2014) and all features were manually checked by eye. Predicted protein features indicative of COX were derived from previous descriptions of known vertebrate COX homologs (Kulmacz et al. 2003; Simmons et al. 2004; Ishikawa et al. 2007; Havird et al. 2008) including residues for COX activity (tyrosine-385, histidine-388, and serine-530), peroxidase activity (glutamine-203 and histidine-207), and substrate binding (arginine-120), as well as heme-binding, membrane-binding, and dimerization domains.

Results

The “Complete Dataset”

Novel transcriptome resources resulted in 40 COX candidate sequences from metazoan lineages such as the Cnidaria, Mollusca, Brachiopoda, Nemertea, Annelida, and Crustacea (Table 1). Ninety-five additional COX candidate sequences were also analyzed from public databases, 25 of which had not been previously analyzed in a phylogenetic context (Online Resource 3). Thirteen MPO and PXDN genes that were used as the outgroup included two novel sequences from sponge transcriptomic data, while the other 11 were from public databases (Table 2). In total, the “complete dataset” consisted of 148 sequences.

Evolution of COX in the Metazoans

ML analyses of the “complete dataset” (Fig. 1) revealed the presence of COX in many metazoan lineages. Putative homologs clustered with well-characterized COX sequences with strong support to the exclusion of MPO and PXDN sequences. No putative COX homologs were found in Hemichordata, Echinodermata, or Chelicerata. Notably, COX homologs were also absent from genomes and transcriptomes of Ctenophora and Porifera, the hypothesized sister lineages to all other metazoans (see below). In contrast, putative COX homologs were recovered from select

taxa within Annelida, Mollusca, and Cnidaria. Within Cnidaria, all octocorallian (i.e., soft coral), but no hexacorallian (i.e., hard coral or anemone), datasets were found to harbor putative COX homologs. ML bootstrap analysis of the complete dataset identified many metazoan COX lineages as being weakly supported, leading to unresolved relationships in the 50 % majority-rule tree (Fig. 1). Similar results were obtained from BI analysis of the complete dataset (Fig. 2).

RogueNaRok identified 27 potential “rogue” sequences (identified in Figs. 1, 2; Tables 1, 2, and Online Resources 3 and 4), including a putative COX from the poriferan *Amphimdeon queenslandica* with an extremely long branch as well as sequences from the louse *Pediculus humanus*, the nemertean *Tubulanus polymorphus*, and the annelid *Protodriloides symbioticus*. Exclusion of such “rogues” reduced the dataset to 121 sequences, with ML (Fig. 3) and BI (Online Resource 5) analyses recovering greater support for monophyly of previously recognized COX lineages (e.g., bootstrap support for COX-1 of vertebrates increased from 45 to 93 when rogue sequences were excluded). Furthermore, “rogue” exclusion found Sipuncula COX sequences transitioning from an unresolved polytomy (Fig. 1) to being sister to other COX sequences (albeit with a relatively low bootstrap support of 72), while aphid sequences were still sister to all other COX sequences. Relationships were very similar between ML and BI analyses, with the exception of COX sequences from the aphid (*Acyrtosiphon pisum*).

Table 2 Peroxidase (PXDN) and myeloperoxidase (MPO) sequences used as outgroups

Sequence name	Organism	^a Accession #
Vertebrata		
MPO human	<i>Homo sapiens</i>	P05164.1
MPO alligator	<i>Alligator mississippiensis</i>	XP_006261742.1
MPO W. clawed frog	<i>Xenopus (Silurana) tropicalis</i>	XM_002935190.2
Hemichordata		
^b PXDN acorn worm	<i>Saccoglossus kowalevskii</i>	XP_002735321.1
Echinodermata		
MPO sea urchin	<i>Strongylcentrotus purpuratus</i>	SPU_016914
Placozoa		
^b PXDN-like <i>Trichoplax</i>	<i>Trichoplax adhaerens</i>	TriadT22758
Platyhelminths		
^b PXDN blood fluke	<i>Schistosoma mansoni</i>	Smp_123650.1
Mollusca		
^b MPO oyster	<i>Crassostrea gigas</i>	EKC40014.1
Cnidaria		
PXDN-like sea anemone	<i>Nematostella vectensis</i>	EDO42052
Ctenophora		
MPO sea gooseberry	<i>Pleurobrachia bachei</i>	sb_2651030
MPO sea walnut	<i>Mnemiopsis leidyi</i>	ML106611a

^a Accession numbers are from GenBank, Ensemble, or individual genomic datasets

^b Indicates “rogue” sequences removed from selected phylogenetic analyses

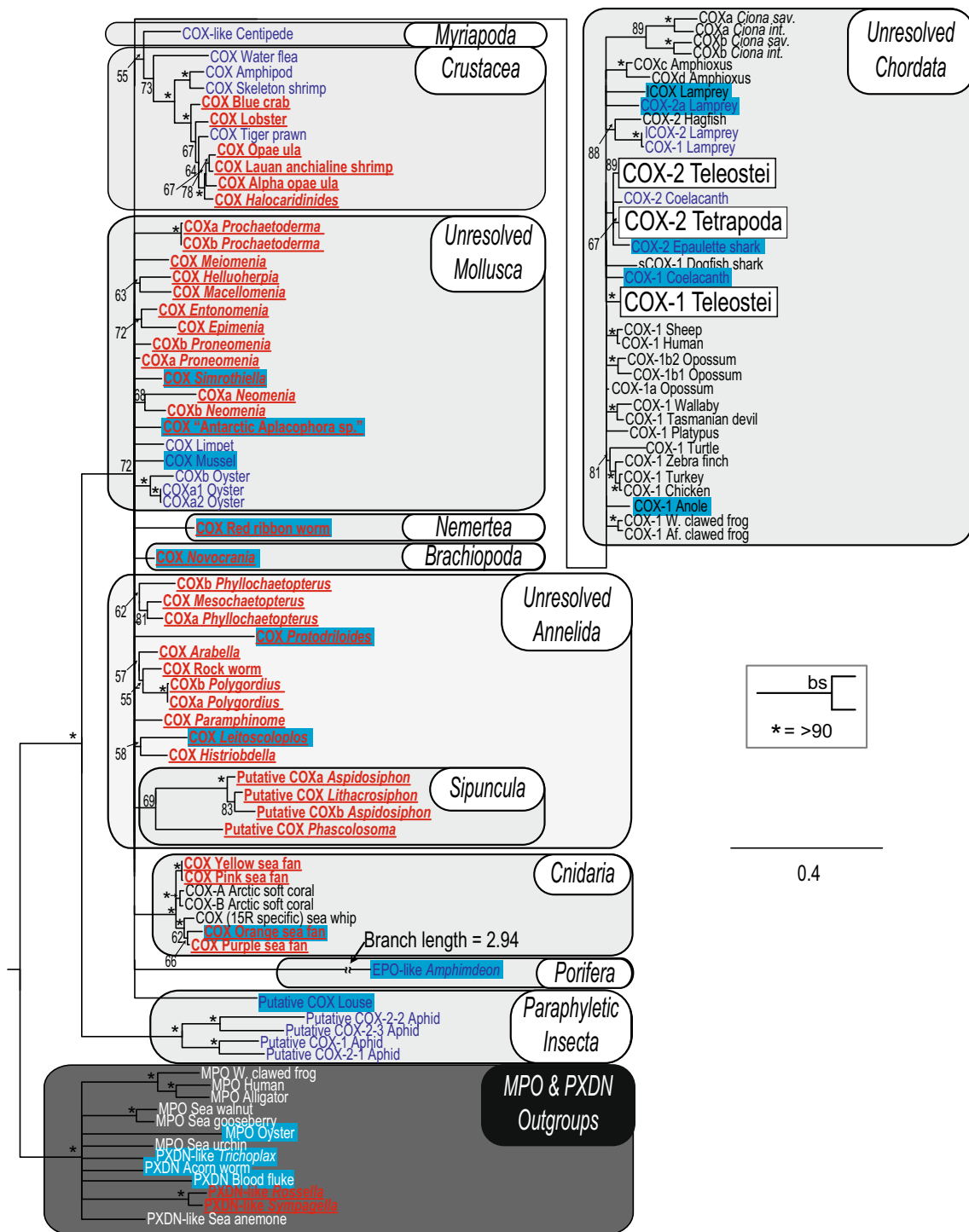


Fig. 1 Maximum likelihood (ML) phylogeny generated with RAXML for all 148 cyclooxygenase (COX) and outgroup amino acid sequences used in the current analysis based on majority rule (50 %). The split between COX and non-COX sequences was used to root the tree (as per Daiyasu and Toh 2000). Values at nodes represent bootstrap support based on 1000 rapid bootstraps conducted in RaxML. Nodes with >90 % bootstrap support values are noted with asterisks. Taxa are grouped based on phyla, although many phyla are unresolved. Sequences generated from novel transcriptomic data and

presented here for the first time are *bolded and underlined* (in red in the online version). Sequences obtained from public sources, but lacking in previous phylogenetic analyses of COX are presented in gray (ingroup taxa; presented in dark blue in the online version) or white (outgroup taxa). Rogue sequences that were subsequently removed from particular analyses are boxed (and shaded blue in the online version). Scale bar shows replacements/site. See Online Resource 6 for relationships among vertebrate COX sequences (Color figure online)

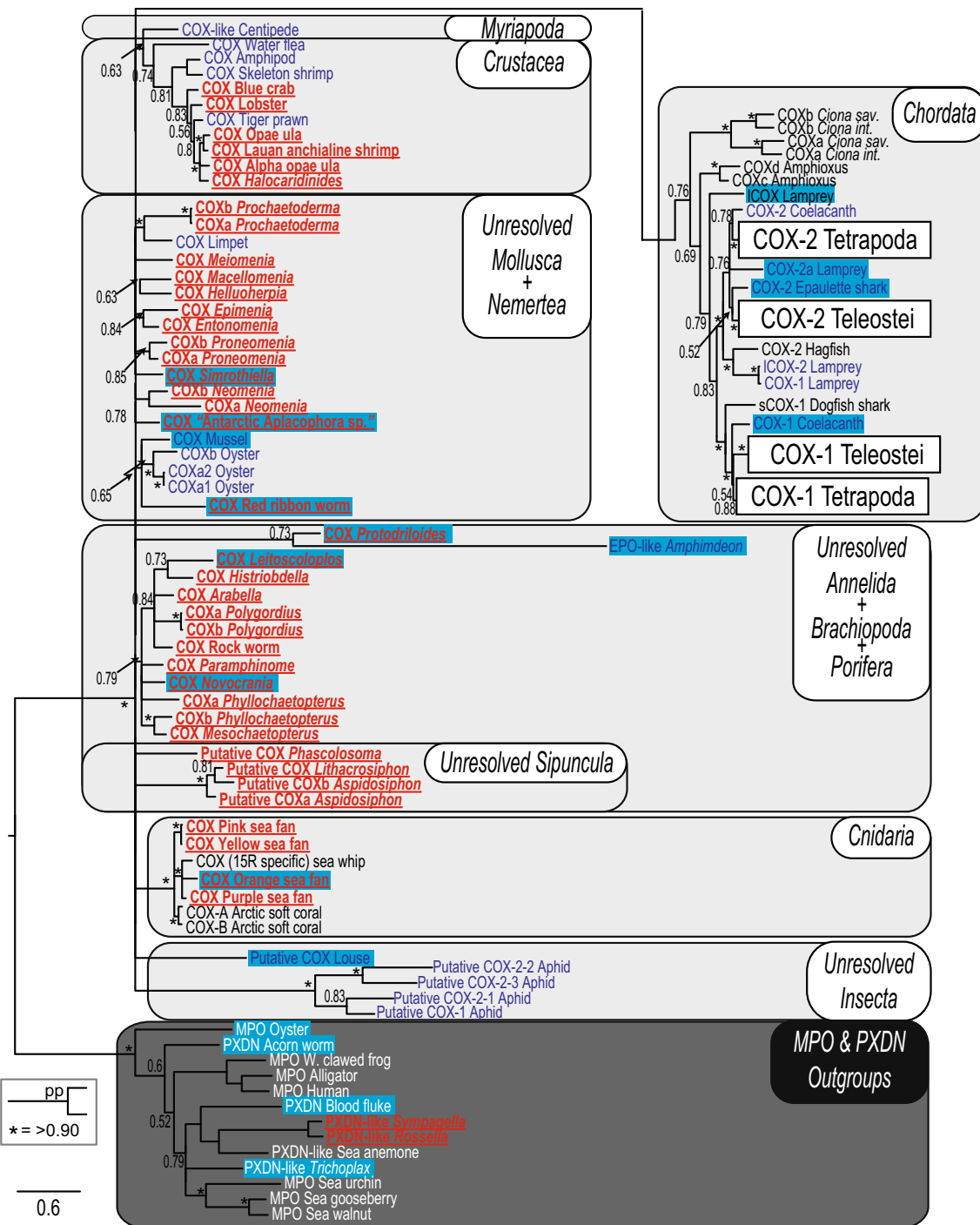


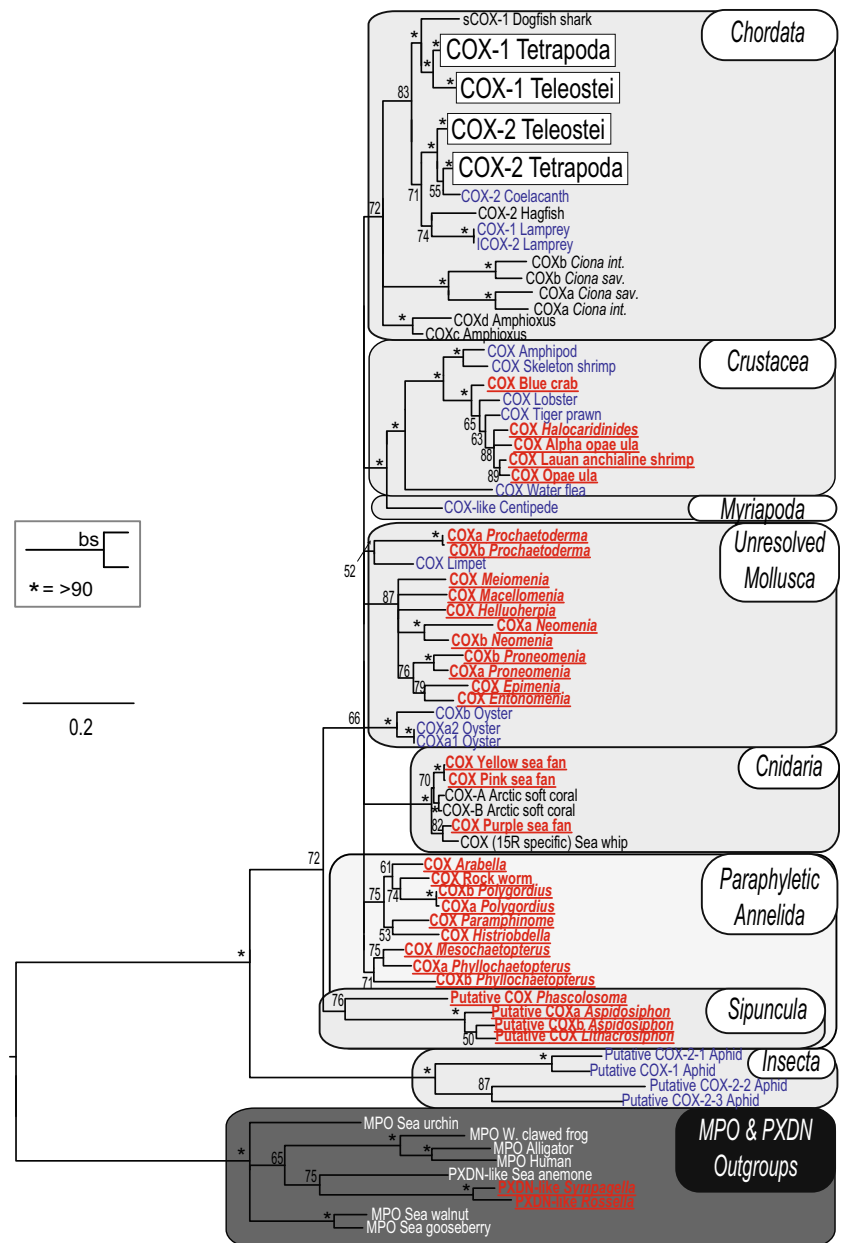
Fig. 2 Bayesian inference (BI) phylogeny generated with PhyloBayes for all 148 cyclooxygenase (COX) and outgroup amino acid sequences used in the current analysis based on majority rule (50 %). Values at nodes represent posterior probabilities based on 15,000

cycles (~15,000,000 generations) conducted in PhyloBayes. Nodes with >90 % posterior probability support values are noted with asterisks. Color schemes, shading, and scale bar as in Fig. 1. See Online Resource 6 for relationships among vertebrate COX sequences

Relationships among COX sequences of chordates have been described previously (Havird et al. 2008; Havird and Miyamoto 2010; Kawamura et al. 2014) and our analyses largely support those previously inferred topologies and inferences (Online Resources 6 and 7).

The position of octocorallian COX sequences (both previously described sequences and novel sequences included here) within the resulting topologies is particularly notable because 1) they represent the earliest metazoan clade (i.e., Cnidaria) with well-defined COX genes shown

Fig. 3 Maximum likelihood (ML) phylogeny generated with RAXML for 121 non-rogue cyclooxygenase (COX) and outgroup amino acid sequences used in the current analysis based on majority rule (50 %). Run parameters, color schemes, shading, and scale bar as in Fig. 1. See Online Resource 7 for relationships among vertebrate COX sequences (Color figure online)



biochemically to convert arachidonic acid to PGs (Koljak et al. 2001; Valmsen et al. 2001, 2004), 2) octocorallians were the only cnidarian group found to contain COX, and 3) a previous COX phylogenetic analysis placed octocorallian sequences sister to Chordata to the exclusion of Crustacea (Varvas et al. 2009). Because placement of octocorallian COX sequences was unresolved in our analyses, topologies with alternative placements of these sequences were explored (Fig. 4). Other clades were allowed to move freely or were constrained in accordance with the most likely topology (Fig. 4e). Although $\Delta L_n L$ values tended to increase with more divergent phylogenetic placement of this clade (Fig. 4i), topologies with octocorallian COXs as

sister to other COX sequences did not significantly differ from that of the most likely octocorallian/chordate relationship ($P > 0.05$, SH tests). However, topologies with the octocorallian COX clade within Chordata generally resulted in significantly worse values of $\Delta L_n L$ (Fig. 4i, $P < 0.05$).

Conserved COX Residues and Domains Among Metazoans

InterProScan analysis of protein domains (confirmed manually by inspection of the untrimmed amino acid alignment) generally supported inferred COX identities via

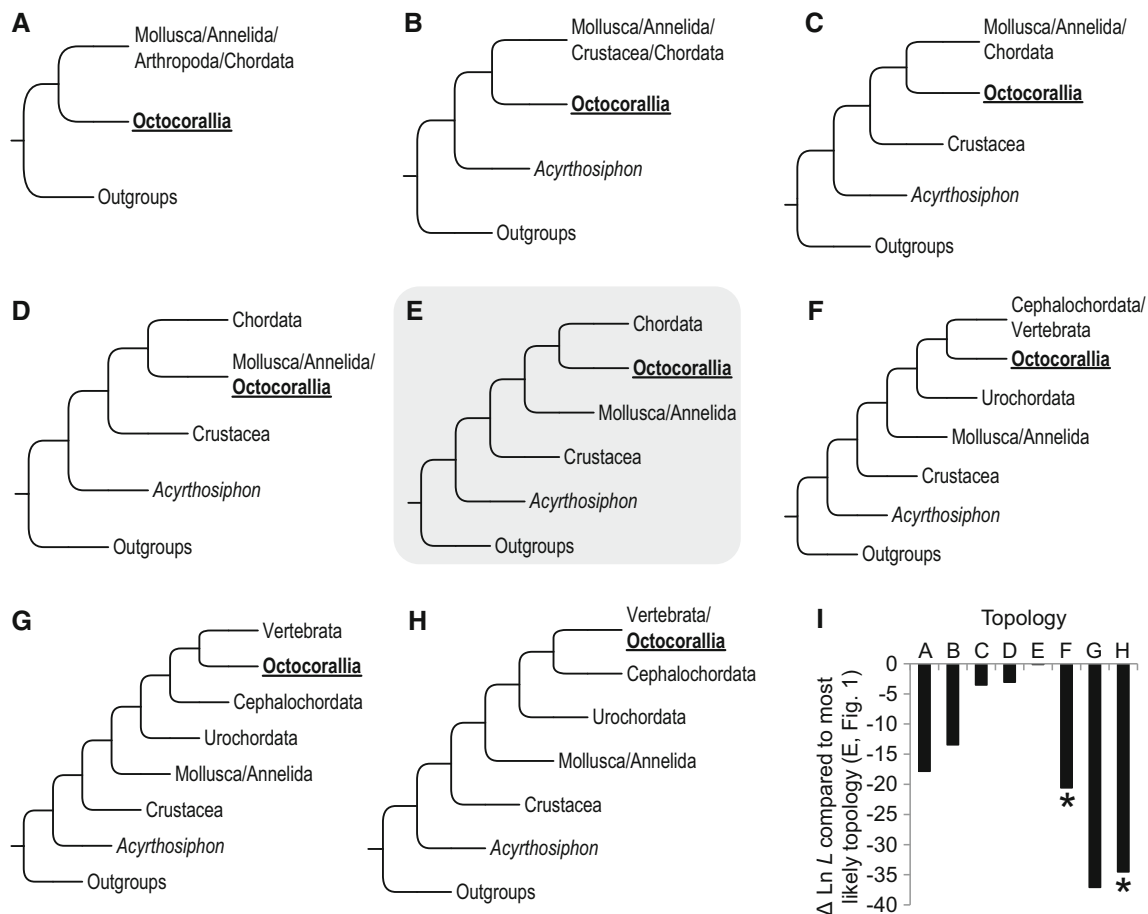


Fig. 4 Tests of alternative topologies for hypotheses on the origin of cyclooxygenase (COX) in Metazoa. In the most likely topology (E, shaded, via RAxML maximum likelihood analyses), COX sequences from Cnidaria/Octocorallia grouped sister to Chordata COX sequences, albeit with low support (hence the polytomy in Fig. 1). Alternative topologies with either earlier (A–D) or latter (F–H) branching positions of Cnidarian/Octocorallia were tested against the

most likely topology using Shimodaira-Hasegawa (SH) tests as implemented in RAxML. Decreases in log-likelihood ($\Delta \text{Ln } L$) from the most likely topology are presented in (I), with *asterisks* indicating those topologies that were significantly worse (according to SH tests). Basal placement of Cnidaria/Octocorallia is reasonable, but more derived placements are significantly worse than the most likely topology ($P < 0.05$)

phylogenetic analyses. For example, sequences identified from Mollusca, Annelida (excluding Sipuncula), and Crustacea which grouped most closely with chordate COX possessed conserved and characteristic amino acid residues for chordate COX and peroxidase activity along with major recognized protein domains (Table 3). On the other hand, those sequences identified as being “rogues” (e.g., *A. queenslandica*, Fig. 1), or those that grouped more distantly to other sequences (Insecta, Fig. 1; Sipuncula, Fig. 3), possessed only a subset of these features. For instance, the putative COX homolog from the poriferan *Amphimdeon* (1) had the longest branch in the ML analysis (Fig. 1); (2) was identified by RogueNaRok as being the least stable (Online Resource 4); and (3) lacked nearly all domains and residues characteristic of COX (Table 3). Along with this, the putative COX genes from insects, sipunculans, and the single

myriapod represent intermediate cases, possessing some, but not all, residues for vertebrate COX-specific functionality. For insects, louse COX grouped closely with Chordata COX and possessed 1 of 3 and 1 of 2 residues for COX and peroxidase activity, respectively, while aphid sequences grouped more distantly to Chordata COX (Figs. 1, 3) and lacked all COX and peroxidase residues. When rogues were excluded, sipunculan sequences grouped more distantly to Chordata COX (Fig. 3) and also lacked functional residues (Table 3). Myriapod COX grouped strongly with the crustacean COX sequences (bootstrap support of 95 in the “rogues removed” dataset, Figure 3), but had 1 of 3 and 2 of 2 conserved residues for COX and peroxidase activity, respectively.

Functional analyses of chordate COX sequences were performed previously (Havird et al. 2008). Although the

Table 3 Conservation of known vertebrate cyclooxygenase functional domains and residues

Domain or residue	AA Residue ^a		Presence (+) or absence (-)										
	Amphimdeon	Insecta	Brachiopoda ^b	Nemertea	Annelida ^c	Mollusca	Sipuncula	Myriapoda	Crustacea	Cephalochordata ^d			
Cyclooxygenase = active residues	Y-385	-	+	+	+	+	±	-	+	+			
	H-388	-	±	+	+	+	±	-	+	+			
	S-530	-	-	?	+	+	-	+	±	+			
	Q-203	-	±	?	+	+	-	+	+	+			
	H-207	-	-	?	+	+	±	+	+	+			
Substrate-binding site	R-120	-	-	?	+	+	-	+	+	±			
	596-599	-	-	?	+	+	-	+	-	+			
Endoplasmic reticulum (ER) retention signal	I-24	-	±	?	+	+	+	+	+	±			
	71-115	-	±	?	+	+	+	+	+	±			
Membrane-binding domain	N-68	-	±	?	+	±	+	-	±	±			
	N-104	-	±	?	+	±	+	+	+	±			
N-glycosylation residues	N-144	-	±	?	+	+	+	+	+	±			
	N-410	-	±	+	-	±	+	-	±	±			
Dimerization domains	N-593	-	-	?	-	±	±	-	±	±			
	33-70	-	±	?	+	+	+	+	+	±			
Haem-binding domain	124-142	-	±	?	+	+	-	+	+	±			
	312-316	+	±	?	+	+	+	+	+	±			
N-terminal insertions	305-314	-	±	?	+	+	+	+	+	±			
	23-32	-	±	?	+	±	+	+	+	±			
C-terminal insertions	573-588	-	-	?	-	-	-	-	±	±			
	R-513	-	-	?	+	±	±	-	-	-			
Vert. inhibitor target residues	V-523	+	-	?	+	±	-	+	+	±			

^a Based on sheep COX-1 numbering (GenBank accession# NP_001009476)

^b Many features in Brachiopoda are unknown (?) due to an incomplete (20 %) sequence being the only representative from this phylum

^c Excluding Sipuncula sequences

^d See Havird et al. (2008) for conservation of domains and residues in Cnidaria and other Chordata

level of conservation among residues in the Cephalochordata was unknown at that time due to incomplete sequences from the cephalochordate *Branchiostoma lanceolatum*, new complete sequences from *B. floridae* (GenBank accession# XP_002586987.1 and XP_002613340.1) provide a more comprehensive understanding. Specifically, cephalochordates possess conserved functional residues and domains characteristic of vertebrate COX.

Discussion

Phylogenetic analyses presented here for COX sequences demonstrate occurrences of this gene across multiple animal lineages, including newly described instances from major invertebrate phyla. Previously, discovery of COX among soft corals (i.e., members of the phylum Cnidaria) raised the possibility that this enzyme spanned the breadth of animal lineages (Koljak et al. 2001). Our sampling of lineages across Metazoa (Online Resources 1–3), including genomes from Ctenophora, the putative sister lineage to all other metazoans (Ryan et al. 2013; Moroz et al. 2014), yielded members of the Cnidaria (Octocorallia) as the earliest animal lineage in which COX genes with notable sequence homology were found (Valmsen et al. 2001; 2004). Furthermore, statistical tests of alternative topologies suggest a basal position of octocorallian COX is reasonable.

Although a candidate was identified from the genome of the sponge *Amphimdeon*, this sequence is likely not a “true” COX homolog since it lacks all of the critical amino acid residues typical of this gene (Table 3). Instead, its grouping with COX in phylogenetic analyses is likely due to long branch attraction (Felsenstein 1978). Assuming the sequenced ctenophore and *Amphimdeon* genomes are representative of their phyla and inferred topologies of metazoan evolution based on phylogenomic data (Ryan et al. 2013; Moroz et al. 2014) are correct, this places the last common ancestor of Cnidaria and Bilateria as the earliest animal lineage with a confirmed COX. Along with this, the presence of COX in the red alga *Gracilaria vermiculophylla* suggests it originated prior to the Metazoa and may have been lost in the sponges and ctenophores. Importantly, the COX genes from octocorallians not only show sequence homology to Chordata COX-1/2, but also have been documented to biochemically produce PGs from arachidonic acid (Koljak et al. 2001; Valmsen et al. 2001, 2004).

Although all but the earliest branches of the animal tree likely possessed a functional COX, a recognizable homolog was not found in many lineages examined here. For example, all Octocorallia datasets possessed a COX homolog, with the exception of the *Gersemia antarctica*

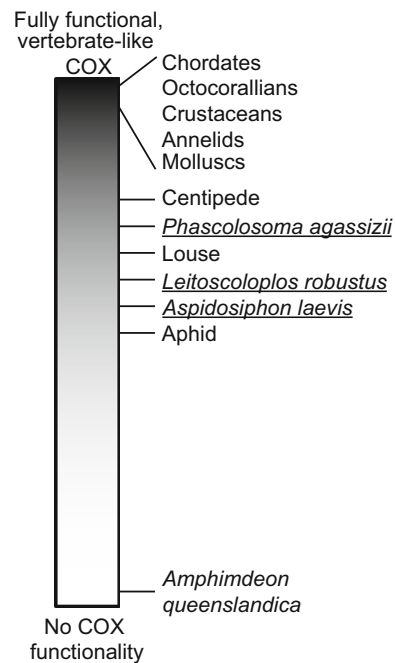


Fig. 5 Schematic representative of the spectrum of COX enzymes across known animal COX sequences, from well-characterized chordate COX to the non-COX *Amphimdeon* sequence described here. This spectrum is based on conserved COX functional residues and domains (Table 3) as well as phylogenetic affinity with known COX homologs. Sipuncula sequences are *underlined*, showing variation within this group

transcriptome, which is likely a sampling artifact given the relatively few ($n = 6516$) contigs in the dataset. In contrast, COX homologs were absent from the nine Hexacorallia (i.e., hard coral) transcriptomes, the single Scyphozoa transcriptome, and genomes from *Nematostella vectensis* and *Acropora digitifera*. As the massive amounts of PGs produced by some octocorallians act primarily to discourage predators (Gerhart 1986; Coll 1992), selective pressures may not have acted toward retaining COX-mediated PG synthesis in other cnidarian lineages (while remaining strong in the soft corals), possibly due to alternative predation prevention mechanisms in these other lineages.

Putative COX homologs appear to represent a spectrum (Fig. 5): ranging from the well-characterized, fully functional enzymes found in chordates to the candidate *Amphimdeon* COX, which is almost certainly not a functional COX, but more likely an EPO, peroxidase, or other non-COX member of the MPO superfamily (based on BLASTp searches to GenBank). Sequences that show close phylogenetic affinity to Chordata COX sequences and contain all functionally important amino acid residues include those from Mollusca, Crustacea, and non-sipunculan Annelida. In between these extremes are the putative COX sequences from Sipuncula, Insecta, and Myriapoda. Therefore, these sequences are designated as “COX-like” or “putative”

until further biochemical characterizations are conducted. Notably, COX functional residues are based on activities in chordate sequences and the identities of these residues may not extend to distantly related taxa such as Sipuncula, Insecta, and Myriapoda.

Supporting this, as well as providing evidence for the possibility of functional COX genes in taxa where homologs were not identified based on sequence homology, is the previously mentioned case of the red alga *Gracilaria vermiculophylla* (Kanamoto et al. 2011; Varvas et al. 2013). BLASTp searches from other unicellular eukaryote genomes on Ensembl failed to return COX homologs with appreciable sequence identity, although the red alga COX sequences (including one from *Coccotylus truncatus*) have low-sequence identity with other COX sequences and lack several functional domains/residues, despite functioning biochemically as COX enzymes (Kanamoto et al. 2011; Varvas et al. 2013). Furthermore, when red alga sequences are included in a phylogenetic analysis with the sequences analyzed here (Online Resource 8), they form a clade with other COX sequences that is sister to the aphid sequences. This suggests that the Sipuncula, Insecta, and Myriapoda sequences may indeed be functional COX enzymes, despite a lack of chordate functional residues. To further investigate this possibility, structural similarity was investigated between several of these sequences and the experimentally derived sheep COX-1 structure (accession# 1DIY in the RCSB protein data bank), as was previously done for the red alga COX (Varvas et al. 2013). Novel and reference COX structures were modeled against the sheep COX-1 structure using MODELLER 9.14 (Eswar et al. 2006) and resulting modeled COX structures were compared using TM-score (Zhang and Skolnick 2004). Based on these results (Online Resources 9 and 10), it seems reasonable that despite a lack of sequence homology, COX sequences from Sipuncula, Insecta (specifically the louse), and Myriapoda may possess adequate structural homology to qualify as functional COX enzymes (i.e., they had as much or more structural similarity to sheep COX-1 as red alga COX did). Moreover, even though COX homologs were not found in the ctenophore and sponge genomes, there may be functional COX genes in these genomes with low-sequence homology to known COXs. Future searches based on structural homology or hidden Markov model (HMM) profiles may prove more successful for identifying such divergent COX enzymes.

Although this study represents the first documentation of COX in annelids, putative homologs were only found in 9 of the 152 (6 %) annelid transcriptomes searched, in spite of possessing appreciable numbers (averaging ~65,267 contigs per transcriptome) of reasonably long (averaging 704 bp) contigs (Online Resource 2). The genes may be absent because they have been lost from the genome, or

because transcriptome data were used, they may not have been expressed at recoverable levels at the time of sampling. Another possibility in light of the above discussion is that COX homologs are present, albeit with low-sequence homology. However, for molluscs, COX appears to be present in most major classes, with homologs described here for Bivalvia, Polyplacophora, and Aplousobranchia (both Solenogastres [=Neomeniomorpha] and Caudofoveata [=Chaetodermomorpha]) possessing functional residues and domains. Moreover, although no COX homologs have been identified from Cephalochordata or Scaphopoda, there is a general lack of high-quality genomic and transcriptomic resources for these lineages. Furthermore, COX sequences were found in the majority of examined crustacean transcriptomes but not those of insects. Thus, despite evidence of PG synthesis (Stanley 2006; Stanley and Miller 2006; Stanley and Kim 2011), our analysis continues to offer no clear precedent for a chordate-like COX in Insecta. Lastly, vertebrate COX-1/2 have N-/C-terminal insertions and inhibitor target residues exhibiting specificity to single isoforms (i.e., either COX-1 or COX-2). Interestingly, sequences in many invertebrate lineages possess some combination of these features, giving them characteristics of both COX-1 and COX-2 and identities as potential functional precursors to vertebrate COX-1/2, as hypothesized earlier (Koljak et al. 2001).

Are COX duplications common in animal lineages with these genes? As noted by Kawamura et al. (2014), the presence of independently duplicated COX isoforms in vertebrates, urochordates, cephalochordates, and the octocorallian *Gersemia fruticosa* suggested this may be the case. However, analyses here do not support the universality of this trend: while species such as *Crassostrea gigas* and *Polygordius* sp. possess duplicates of COX individually, members of the phyla Mollusca, Annelida, Crustacea, and Cnidaria more-or-less harbor single, non-duplicate COX sequences, as noted earlier for Crustacea (Varvas et al. 2009). Although further searching of ever improving genomic and transcriptomic resources or COX-directed sequencing may reveal the presence of other COX isoforms, concluding that most invertebrate lineages possess just a single COX gene may be appropriate. Reasons for this could be a narrower range of COX functions in invertebrates relative to vertebrates and/or the history of genome duplications in the vertebrates (Meyer and Schartl 1999; Dehal and Boore 2005; Putnam et al. 2008). A combination of these two hypotheses may be the most reasonable scenario, where genome duplications led to multiple COX genes in the chordates and retention was driven by selection for subfunctionalization (Lynch and Conery 2000). If this is the case, closely related taxa with both duplicate and single COX genes (e.g., *C. gigas* and *Mytilus edulis*) provide a system to further investigate the evolution and subfunctionality of COX genes.

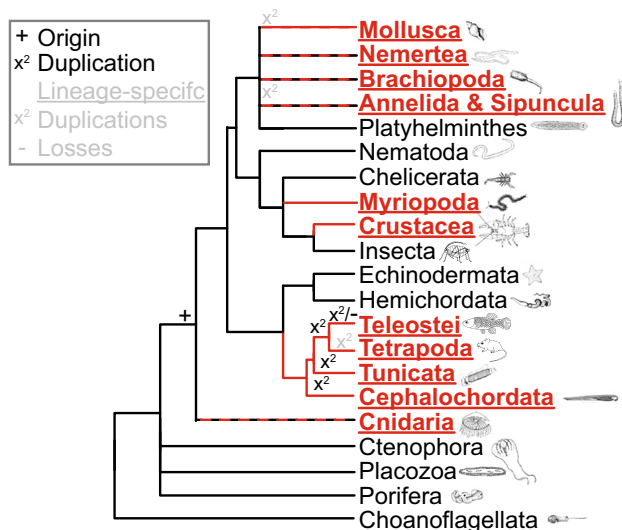


Fig. 6 Proposed hypothesis for the evolution of cyclooxygenase (COX) in Metazoa, with history of duplications among phyla. Relationships between metazoan lineages are based on Kocot et al. (2010). The origin of COX is noted with a “+”, while subsequent whole-lineage duplications are indicated by “ x^2 ”. Losses or duplications specific to particular lineages within phyla are indicated by gray symbols. Lineages with known COX sequences are *underlined and bolded* (and presented in *red* in the online version). Lineages where COX has not been found are presented in black and not underlined or bolded. Branches leading to lineages where some, but likely not all members of the clade have COX (e.g., Cnidaria) or COX identity is not well known from several taxa (e.g., Brachiopoda) are presented in *red/black dashed lines* in the online version (Color figure online)

Notably, the phylogenetic analyses presented here also allow for commentary on COX naming schemes, which have been inconsistent across the literature. Importantly, the COX-1 and COX-2 of aphids are not homologous to the COX-1 and COX-2 of chordates. Therefore, these putative COX homologs should be named COXa and COXb, following the convention established for naming COX homologs in amphioxus and *Ciona* (Havird et al. 2008). A similar convention is followed in naming the novel invertebrate COX homologs described here, with COXa and COXb being used in cases where duplicate homologs were found. When more than two homologs were found (e.g., oyster), they were named COXa1, COXa2, etc., according to their phylogenetic relationships.

Conclusion

Based on the phylogenetic analyses presented here, an evolutionary hypothesis summarizing our current knowledge of COX in Metazoa can be presented (Fig. 6), with duplications being common. Although “true” COX genes have only been found in metazoans and two species of red alga, related genes that oxidize fatty acids are found in

plants and fungi (Daiyasu and Toh 2000; Noverr et al. 2003) and COX-independent PG synthesis pathways have been described in these lineages (Morrow and Roberts 1996; Hornsten et al. 1999; Marks 1999; Noverr et al. 2003). Similarly, novel COX-independent PG pathways may yet be described from animal lineages that apparently lack COX genes. As a popular target for the pharmaceutical industry (i.e., non-steroidal anti-inflammatory drugs like aspirin target COX), COX and COX-like PG synthesis pathways and the biochemistry of the new COX homologs presented here could lead to development of novel pharmaceutical products or improvements on those from natural sources. Supporting this, PGs from the octocorallian *Plexaura homomalla* have been previously targeted for pharmaceutical applications (Bayer and Weinheimer 1974). Both COX-targeted examinations of non-vertebrate lineages and ever improving genomic resources across all animal taxa will allow for continued reevaluation of our hypothesis (Fig. 6) as well as clarify PG synthesis pathways and their evolution in the Metazoa.

Acknowledgments This work was supported by National Science Foundation grants Assembling the Tree of Life (DEB #1036537) and Antarctic Organisms and Ecosystems (ANT#1043745) to K.M.H. and S.R.S. and National Aeronautics and Space Administration grant NNX13AJ31G to K.M.H. and K.M.K. This represents contributions #127 and #32 to the Auburn University (AU) Marine Biology Program and Molette Biology Laboratory for Environmental and Climate Change Studies, respectively.

References

- Aberer AJ, Krompass D, Stamatakis A (2013) Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Syst Biol* 62:162–166
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Bayer FM, Weinheimer AJ (1974) Prostaglandins from *Plexaura homomalla*: ecology, utilisation and conservation of a major medical marine resource. A Symposium. Studies in Tropical Oceanography. University of Miami Press, Miami
- Cannon JT, Kocot KM, Waits DS, Weese DA, Swalla BJ, Santos SR, Halanych KM (2014) Phylogenomic resolution of the hemichordate and echinoderm clade. *Curr Biol* 24:2827–2832
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552
- 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
- Coll JC (1992) The chemistry and chemical ecology of octocorals (Coelenterata, Anthozoa, Octocorallia). *Chem Rev* 92:613–631
- Daiyasu H, Toh H (2000) Molecular evolution of the myeloperoxidase family. *J Mol Evol* 51:433–445
- Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol* 3:e314

- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U, Sali A (2006) Comparative protein structure modeling using modeller. *Curr Protoc Bioinformatics*. doi:10.1002/0471140864.ps0209s50 Chapter 2: Unit 2.9
- Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27:401–410
- Flicek P, Amode MR, Barrell D et al (2014) Ensembl 2014. *Nucleic Acids Res* 42:D749–D755
- Flippin JL, Huggett D, Foran CM (2007) Changes in the timing of reproduction following chronic exposure to ibuprofen in Japanese medaka, *Oryzias latipes*. *Aquat Toxicol* 81:73–78
- Genomic Resources Development Consortium, Havird JC, Santos SR (2014) Genomic resources notes accepted 1 June 2014–31 July 2014. *Mol Ecol Resour* 14:1322
- Gerhart DJ (1986) Prostaglandin-A₂ in the Caribbean gorgonian *Plexaura homomalla*—evidence against allelopathic and antifouling roles. *Biochem Sys Ecol* 14:417–421
- Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27:221–224
- Havird JC, Miyamoto MM (2010) The importance of taxon sampling in genomic studies: an example from the cyclooxygenases of teleost fishes. *Mol Phylogenet Evol* 56:451–455
- Havird JC, Miyamoto MM, Choe KP, Evans DH (2008) Gene duplications and losses within the cyclooxygenase family of teleosts and other chordates. *Mol Biol Evol* 25:2349–2359
- Hornsten L, Su C, Osbourn AE, Garosi P, Hellman U, Wernstedt C, Oliw EH (1999) Cloning of linoleate diol synthase reveals homology with prostaglandin H synthases. *J Biol Chem* 274:28219–28224
- Ishikawa TO, Griffin KJ, Banerjee U, Herschman HR (2007) The zebrafish genome contains two inducible, functional cyclooxygenase-2 genes. *Biochem Biophys Res Commun* 352:181–187
- Jarving R, Jarving I, Kurg R, Brash AR, Samel N (2004) On the evolutionary origin of cyclooxygenase (COX) isozymes: characterization of marine invertebrate COX genes points to independent duplication events in vertebrate and invertebrate lineages. *J Biol Chem* 279:13624–13633
- Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236–1240
- Kanamoto H, Takemura M, Ohyama K (2011) Identification of a cyclooxygenase gene from the red alga *Gracilaria vermiculophylla* and bioconversion of arachidonic acid to PGF₂α in engineered *Escherichia coli*. *Appl Microbiol Biotechnol* 91:1121–1129
- Kawamura M, Inaoka H, Obata S, Harada Y (2014) Why do a wide variety of animals retain multiple isoforms of cyclooxygenase? *Prostaglandins Other Lipid Mediat* 109–111:14–22
- Kocot KM, Cannon JT, Halanych KM (2010) Elucidating animal phylogeny: Advances in knowledge and forthcoming challenges. In: DeSalle R, Sheirwater B (eds) *Key Transitions*. CRC Press, Florida, pp 16–35
- Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, Santos SR, Schander C, Moroz LL, Lieb B, Halanych KM (2011) Phylogenomics reveals deep molluscan relationships. *Nature* 477:452–456
- Koljak R, Järving I, Kurg R, Boeglin WE, Varvas K, Valmsen K, Ustav M, Brash AR, Samel N (2001) The basis of prostaglandin synthesis in coral: molecular cloning and expression of a cyclooxygenase from the Arctic soft coral *Gersemia fruticosa*. *J Biol Chem* 276:7033–7040
- Kulmacz RJ, van der Donk WA, Tsai AL (2003) Comparison of the properties of prostaglandin H synthase-1 and -2. *Prog Lipid Res* 42:377–404
- Lartillot N, Philippe H (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol Biol Evol* 21:1095–1109
- Lartillot N, Lepage T, Blanquart S (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286–2288
- Le SQ, Dang CC, Gascuel O (2012) Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Mol Biol Evol* 29:2921–2936
- Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–1155
- Marks F (1999) Arachidonic acid and companions: an abundant source of biological signals. In: Marks F, Furstenberger G (eds) *Prostaglandins, Leukotrienes, and Other Eicosanoids*. Wiley-VCH, Weinheim, pp 1–46
- Meyer A, Schartl M (1999) Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* 11:699–704
- Moroz LL, Kocot KM, Citarella MR et al (2014) The ctenophore genome and the evolutionary origins of neural systems. *Nature* 510:109–114
- Morrow JD, Roberts LJ 2nd (1996) The isoprostanes: current knowledge and directions for future research. *Biochem Pharmacol* 51:1–9
- Noverr MC, Erb-Downward JR, Huffnagle GB (2003) Production of eicosanoids and other oxylipins by pathogenic eukaryotic microbes. *Clin Microbiol Rev* 16:517–533
- Ott M, Zola J, Aluru S, Stamatakis A (2007) Large-scale maximum likelihood-based phylogenetic analysis on the IBM BlueGene/L. In: *ACM/IEEE Supercomputing Conference 2007*
- Paredes JF, Vera LM, Martinez-Lopez FJ, Navarro I, Sanchez Vazquez FJ (2014) Circadian rhythms of gene expression of lipid metabolism in Gilthead Sea bream liver: synchronisation to light and feeding time. *Chronobiol Int* 31:613–626
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490
- Putnam NH, Butts T, Ferrier DE et al (2008) The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453:1064–1071
- Ryan JF, Pang K, Schnitzler CE et al (2013) The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science* 342:1242592
- Sanderson MJ, Shaffer HB (2002) Troubleshooting molecular phylogenetic analyses. *Annu Rev Ecol Syst* 33:49–72
- Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* 16:1114–1116
- Simmons DL, Botting RM, Hla T (2004) Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. *Pharmacol Rev* 56:387–437
- Singh G, Triadafilopoulos G (1999) Epidemiology of NSAID induced gastrointestinal complications. *J Rheumatol* 26:18–24
- Sorbera LA, Asturiano JF, Carrillo M, Zanuy S (2001) Effects of polyunsaturated fatty acids and prostaglandins on oocyte maturation in a marine teleost, the European sea bass (*Dicentrarchus labrax*). *Biol Reprod* 64:382–389
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313

- Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol* 57:758–771
- Stanley D (2006) Prostaglandins and other eicosanoids in insects: biological significance. *Annu Rev Entomol* 51:25–44
- Stanley D, Kim Y (2011) Prostaglandins and their receptors in insect biology. *Front Endocrinol (Lausanne)* 2:105
- Stanley DW, Miller JS (2006) Eicosanoid actions in insect cellular immune functions. *Entomol Exp Appl* 119:1–13
- Valmsen K, Järving I, Boeglin WE, Varvas K, Koljak R, Pehk T, Brash AR, Samel N (2001) The origin of 15R-prostaglandins in the Caribbean coral *Plexaura homomalla*: molecular cloning and expression of a novel cyclooxygenase. *Proc Natl Acad Sci U S A* 98:7700–7705
- Valmsen K, Boeglin WE, Järving I, Schneider C, Varvas K, Brash AR, Samel N (2004) Structural and functional comparison of 15S- and 15R-specific cyclooxygenases from the coral *Plexaura homomalla*. *Eur J Biochem* 271:3533–3538
- Vane JR, Bakhle YS, Botting RM (1998) Cyclooxygenases 1 and 2. *Annu Rev Pharmacol Toxicol* 38:97–120
- Varvas K, Kurg R, Hansen K, Järving R, Järving I, Valmsen K, Löhelaid H, Samel N (2009) Direct evidence of the cyclooxygenase pathway of prostaglandin synthesis in arthropods: genetic and biochemical characterization of two crustacean cyclooxygenases. *Insect Biochem Mol Biol* 39:851–860
- Varvas K, Kasvandik S, Hansen K, Järving I, Morell I, Samel N (2013) Structural and catalytic insights into the algal prostaglandin H synthase reveal atypical features of the first non-animal cyclooxygenase. *Biochim Biophys Acta* 1831:863–871
- Yang Z (1996) Among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol Evol* 11:367–372
- Zhang Y, Skolnick J (2004) Scoring function for automated assessment of protein structure template quality. *Proteins* 57:702–710