Published in final edited form as:

Mar Environ Res. 2010 ; 69S1: S1–S3. doi:10.1016/j.marenvres.2009.10.013.

Identification of CYP genes in *Mytilus* (mussel) and *Crassostrea* (oyster) species: First approach to the full complement of cytochrome P450 genes in bivalves

Juliano Zanette^{1,2}, Jared V. Goldstone¹, Afonso C. D. Bainy², and John J. Stegeman¹

¹Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA.

²Biochemistry Department, Biological Sciences Center, Universidade Federal de Santa Catarina, Florianopolis, SC, Brazil.

Abstract

Understanding the fate and effects of organic chemicals in animals requires knowledge of cytochrome P450 (CYP) genes, which thus far are poorly known in bivalve mollusks. We searched for *CYP* sequences in EST databases for *Mytilus* and *Crassostrea* species, lophotrochozoan representatives of the protostomes. From ESTs averaging ca. 924 bp, we identified 58 *CYP* genes in *Mytilus californianus* and 39 *CYP* genes in *Crassostrea gigas*. The sequences fell in all known animal *CYP* clans, and collectively they clustered in phylogenetic analysis with vertebrate *CYP* families 1, 2, 3, 4, 17, 20, 26 and 27. As in deuterostomes, a majority of the sequences fell in Clan 2. The *CYP* sequences found thus far in bivalves suggest a diversity consistent with that found in many other animal species. The present description of mollusk genes provides the overall framework for classification of any additional bivalve sequences. The sequences identified also will be useful in obtaining full-length sequences and in designing primers for analysis of expression of mussel and oyster *CYP* genes, or for recombinant protein expression to identify potential substrates for the bivalve *CYP* proteins, and understand their roles in xenobiotic detoxification and physiology of bivalves.

Keywords

P450; CYP; *Mytilus californianus*; *Crassostrea gigas*; bivalve; pollution; biomarker

Members of the cytochrome P450 (CYP) superfamily are involved in a variety of detoxification and endogenous functions. Increasing genomic information available for model species points to a diversity of CYPs in some invertebrates greater than that observed in vertebrates. The identification and regulation of CYPs in protostomes have been principally characterized in the ecdysozoa, especially insects (Rewitz et al., 2006), and CYPs are poorly known in bivalve mollusks. No genome sequence is available for any bivalve, hampering the identification of CYPs in these ecologically and economically important species. There is need for knowledge of molluscan *CYP* genes and their regulation, in part to identify potential biomarkers of

© 2009 Elsevier Ltd. All rights reserved.

Correspondence to: John J. Stegeman, Biology Department, MS 32, Woods Hole Oceanographic Institution, Woods Hole MA USA 02543, jstegeman@whoi.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

exposure in this important animal group, and to understand *CYP* gene evolution in the protostomes.

The induction of cytochrome P450 subfamily 1A (CYP1A) transcripts, protein or enzyme activity has long been employed as a biomarker of exposure of fish and other vertebrates to aryl hydrocarbon receptor agonists such as PCB or PAH contaminants in aquatic environments. A number of studies have searched for similar CYP responses in bivalve mollusks, as these animals possess biological features excellent for pollution monitoring. Total cytochrome P450 levels and monooxygenase activity have been studied in bivalves for decades (Stegeman, 1985; Livingstone et al., 1989). Studies also have attempted to detect CYP1-like proteins using antibodies to vertebrate CYP1As. However, results in many of these studies have been complicated by cross-reactivity of antibodies vertebrate CYP1As with non-P450 molluscan proteins, and/or low enzymatic activity with known CYP substrates (Grosvik et al., 2006; Rewitz et al., 2006). Thus, there is a critical need for the development of new tools for the study of CYPs in bivalves. Identification of *CYP* genes in bivalve EST libraries would allow for the use of gene-specific primers to evaluate gene expression using quantitative PCR, and cloning, heterologous expression and functional analysis of CYP that may be induced by contaminants.

An increasing number of expressed sequence tags (ESTs) have been generated recently for some bivalves, including several *Crassostrea* and *Mytilus* species (Venier et al., 2003; Quilang et al., 2007; Tanguy et al., 2008; Venier et al., 2009). ESTs are a valuable resource for uncovering and identifying genes (e.g. (Seve et al., 2004; Boutrot et al., 2008)). In the present study, we identified and classified *CYPs* from sequences in mussel and oyster EST databases, as a first effort to assess the diversity and to provide a benchmark for future studies of *CYP* genes in bivalves.

All ESTs available as of May, 2009 for *M. californianus*, *M. galloprovincialis*, *C. gigas* and *C. virginica* were downloaded from GenBank, vectors and adaptors were removed, and contigs were constructed using CAP3 (Huang & Madan, 1999). EST sequences averaging 924 bp (including an average of 173 predicted amino acids), were then searched for significant homology to CYPs using a hidden Markov model technique implemented as the ESTWise function of GeneWise (Birney et al., 2004). ESTWise performs a sensitive and accurate search by comparing a library of EST sequences to a hidden Markov model (or library). This technique is less biased than BLAST (Basic Local Alignment Search Tool), which is influenced by the overrepresentation of vertebrate sequences in most BLAST databases used for searching. The identified *CYP* sequences were translated, aligned using Clustalw, and maximum likelihood phylogenetic analyses were performed using RAxML (v 7.0.4) (Stamatakis, 2006; Stamatakis et al., 2008) with the WAG model (Whelan and Goldman) (Whelan & Goldman, 2001) of amino acid substitution. The similarity of single ESTs and contigs with previously annotated CYP sequences was also examined using tBLASTx.

The numbers of *CYPs* we identified in mussel and oyster EST databases are listed in Table 1. The length of the ESTs examined, together with the phylogenetic analyses and tBLASTx results, allow us to classify the CYPs identified in the database searches at least to the level of clan. Sequences from all four major *CYP* clans were identified in *M. californianus* ESTs and contigs, and in *C. gigas* ESTs. All *M. californianus* EST sequences that are listed in Table 1 were submitted by Gracey et al. (2007–2009, unpublished) and were generated using cDNA of organisms that were exposed to a variety of environmental challenges, including heat, cold, hypoxia, emersion, heavy metal and salinity stress. The *CYPs* identified for *C. gigas* are from EST sequences from many different sources (Boutet et al., 2004; Hedgecock et al., 2007; Tanguy et al., 2008; Roberts et al., 2009), but with one exception (CB617459; (Boutet et al., 2004)), none had been previously annotated as CYPs.

The mussel *Mytilus californianus* and the oyster *Crassostrea gigas* were chosen for further analysis as they currently have the largest number of sequenced ESTs available (see Table 1). The relationship of *M. californianus* ESTs or contigs to human CYPs is shown in Figure 1. Alignment and molecular phylogeny identified sequences similar to vertebrate genes in *CYP* families 1, 2, 3, 4, 20, 26 and 27 in *M. californianus*. Sequences similar to vertebrate *CYP* families 1, 2, 3, 4, 17 and 27 were identified in *C. gigas* (not shown). As in deuterostomes, a majority of the molluscan sequences fell in Clan 2, which includes the *CYP1* and *CYP2* families, and most of these genes may be *CYP2*-like genes. Our provisional assignment of sequences to known gene families, including families that are involved in xenobiotic responses in other species, suggests that many of these bivalve CYPs could be involved in defenses against chemicals.

This is the first study to begin approaching the full complement of *CYP* sequences expressed in mussels and oysters. The *CYP* sequences we uncovered present a foundation for more detailed characterization of *CYP* functions and expression in bivalve mollusks. For example, using these sequences and other techniques, we recently cloned full-length *CYP1*-like, *CYP3*-like and *CYP26*-like sequences from *M. edulis*, and established the patterns of expression in different tissues (unpublished). Additional sequences and further phylogenetic studies should support an accurate annotation and reveal the full diversity of bivalve *CYPs*, helping to understand the evolution of *CYP* genes in this branch of the lophotrochozoa.

Acknowledgments

This study was supported in part by an NIH grant to JJS (5R01-ES015912). JZ was a Guest Student at the Woods Hole Oceanographic Institution and was supported by a CAPES Ph.D. Fellowship and CNPq Ph.D. Sandwich Fellowship, Brazil. ACDB was recipient of the CNPQ Productivity Fellowship, Brazil. Study sponsors had no involvement in the studies reported here or in the decision to submit this paper for publication.

References

- Birney E, Clamp M, Durbin R. Genome research 2004;14:988–995. [PubMed: 15123596]
- Boutet I, Tanguy A, Moraga D. Gene 2004;329:147–157. [PubMed: 15033537]
- Boutrot F, Chantret N, Gautier MF. BMC genomics 2008;9:86. [PubMed: 18291034]
- Grosvik BE, Jonsson H, Rodriguez-Ortega MJ, Roepstorff P, Goksoyr A. Aquatic toxicology (Amsterdam, Netherlands) 2006;79:334–340.
- Hedgecock D, Lin JZ, DeCola S, Haudenschield CD, Meyer E, Manahan DT, Bowen B. Proceedings of the National Academy of Sciences of the United States of America 2007;104:2313–2318. [PubMed: 17277080]
- Huang X, Madan A. Genome research 1999;9:868–877. [PubMed: 10508846]
- Livingstone DR, Kirchin MA, Wiseman A. Xenobiotica; the fate of foreign compounds in biological systems 1989;19:1041–1062.
- Nelson DR. Comparative biochemistry and physiology 1998;121:15–22. [PubMed: 9972448]
- Quilang J, Wang S, Li P, Abernathy J, Peatman E, Wang Y, et al. BMC genomics 2007;8:157. [PubMed: 17559679]
- Rewitz KF, Styrisshave B, Lobner-Olsen A, Andersen O. Comp Biochem Physiol C Toxicol Pharmacol 2006;143:363–381. [PubMed: 16769251]
- Roberts S, Goetz G, White S, Goetz F. Marine biotechnology (New York, N.Y 2009;11:24–44.
- Seve M, Chimienti F, Devergnas S, Favier A. BMC genomics 2004;5:32. [PubMed: 15154973]
- Stamatakis A. Bioinformatics (Oxford, England) 2006;22:2688–2690.
- Stamatakis A, Hoover P, Rougemont J. Systematic biology 2008;57:758–771. [PubMed: 18853362]
- Stegeman JJ. Marine Biology 1985;89:21–30.
- Tanguy A, Bierre N, Saavedra C, Pina B, Bachere E, Kube M, et al. Gene 2008;408:27–36. [PubMed: 18054177]

- Venier P, De Pitta C, Bernante F, Varotto L, De Nardi B, Bovo G, et al. BMC genomics 2009;10:72. [PubMed: 19203376]
- Venier P, Pallavicini A, De Nardi B, Lanfranchi G. Gene 2003;314:29–40. [PubMed: 14527715]
- Whelan S, Goldman N. Molecular biology and evolution 2001;18:691–699. [PubMed: 11319253]

Table 1

Number of CYPs identified in bivalve EST data.

Species	Number of ESTs	Number of identified CYPs	CYP Clans ¹ of identified EST
<i>M. californianus</i>	42,354	<u>58</u>	<u>2, 3, 4, mito</u>
<i>M. galloprovincialis</i>	19,699	<u>12</u>	<u>2, 4</u>
<i>C. gigas</i>	29,573	<u>39</u>	<u>2, 3, 4, mito</u>
<i>C. virginica</i>	14,560	<u>14</u>	<u>2, 3, mito</u>

¹Clan 2 includes CYP families 1,2,17, and 21; Clan 3 includes CYP families 3,5,6, and 9; Clan 4 includes CYP family 4; and the mitochondrial ("mito") clan includes CYP families 11,24, and 27(Nelson 1998).