Journal of Shellfish Research, Vol. 30, No. 1, 127–132, 2011.

EXPRESSED SEQUENCE TAG ANALYSIS OF *PHYSA ACUTA*: A FRESHWATER PULMONATE IN KOREA

YONG SEOK LEE,[1] SEUNG-GYU LEE,[2] SE WON KANG,[1] JI EUN JEONG,[1] MOON KI BAEK,[1] SANG-HAENG CHOI,[3] SUNG-HWA CHAE,[4] YONG HUN JO,[5] YEON SOO HAN[5] AND HONG-SEOG PARK[3]*

[1]*Department of Parasitology, College of Medicine and UHRC, Inje University, 633-165 Gaegum-Dong, Busanjin-Gu, Busan, 614-735, Korea;* [2]*Division of Quality Control and Utilization of Animal Products, National Institute of Animal Science, RDA, 77 Chuksan gil, Kwonsun-Gu, Suwon, 441-706, Korea;* [3]*Genome Resource Center, Korea Research Institute of Bioscience and Biotechnology, 52 Oun-Dong, Yusong-gu, Daejeon, 305-806, Korea;* [4]*Research Institute, GnC BIO Co., LTD., 621-6 Banseok-dong, Yusong-gu, Daejeon, 305-150, Korea;* [5]*Department of Agricultural Biology, College of Agriculture and Life Science, Chonnam National University, 300 Yongbong-Dong, Puk-Gu, Gwangju, 336-745, Korea*

ABSTRACT *Physa acuta* (left-handed shell) have strong natural growth activity not only in lentic waters but also in eutrophic environments. Therefore, it has been considered one of the candidate species that could evaluate the degree of water pollution by physiological and biochemical methods. In this study, we constructed a *P. acuta* cDNA library using the 5′ oligo capping method, and determined the sequences of 2,282 clones by 5′ end-single path sequencing. After trimming, clustering, and assembling these sequences, we finally obtained 575 distinctly available transcripts that were 718 bp in average length. These transcripts were annotated using the BLASTX search and were classified by function using KOG analysis. After comparison with biomarker genes already known in several organisms, we identified 27 potential biomarker candidates that were categorized into two groups strongly related to stress and defense genes by their functions. To the best of our knowledge, this is the first report of massive profiling of cDNA sequences and the characterizing of potential biomarker genes in *P. acuta*. Our study offers valuable information to scientists for developing new environmental biomonitoring markers, and for scientists studying the physiology, growth and development, immunity, genetic identification, and evolutional diversity in *P. acuta*.

KEY WORDS: *Physa acuta*, EST, biomarker, cDNA library

INTRODUCTION

To find out an alternative way in measuring the level of contamination in grossly polluted sites, biomarkers have received considerable attention. The biomarkers can be defined as biochemical, physiological, histological, and morphological responses to environmental chemicals. Recently, many investigators have used an EST-based method to screen potential biomarkers by analyzing transcripts expressed from the species of bioindicator organisms (Davey et al. 2001, Gross et al. 2001, Jenny et al. 2002).

In this study, we investigated *Physa acuta*, a freshwater pulmonate with a sinistrally coiled shell, to determine its possible role as a biomarker. *P. acuta* is a ubiquitous species that can tolerate many habitats, even highly polluted waters. The family Physidae has been known to live longer in the aquatic world even after exposure to herbicides such as Diuron (Nebeker & Schuytema 1998), chlorpyrifos, lindane (Cuppen et al. 2002), and Paraquat (Bacchetta et al. 2002). Its resistance to butanol is 2–4 times higher than in fish (Stobaeus et al. 1990). Interestingly, this family is also known to have a higher egg-hatching rate when exposed to a strong toxic chemical (Leung et al. 2004), and can survive better than aquatic insects when exposed to heavy metals, chromium (Cr^{6+}), and arsenic (As^{3+}) (Canivet et al. 2001). These natural characteristics of *P. acuta* suggest that it has a high potential as an index organism for monitoring ecological environments polluted with heavy metals and harmful chemicals.

We constructed a cDNA library using mRNA extracted from the whole body of *P. acuta* and determined its EST sequence. This is the first report profiling the transcriptome in *P. acuta*. We will show valuable candidate genes with the potential as biomonitoring markers compared with those of other organisms.

MATERIALS AND METHODS

Sample Collection

The specimens of *P. acuta* were collected from Lake Dogo located in Asan (clear-water area), Chungnam Province, South Korea, and were maintained in aerated natural lake water. *P. acuta* belongs to the family Physidae, the order Basommatophora, the superorder Pulmonata, the class Gastropoda, and the phylum Mollusks.

cDNA Library Construction

Total RNA was isolated with Trizol, and mRNA was purified from the total RNA using an mRNA purification kit. A cDNA library was generated using the SMART cDNA Library Construction Kit (Clontech, Palo Alto, CA). In brief, cDNA was ligated into a TriplEx2 vector and packaged using the Gigapack Gold I (Stratagene, La Jolla, CA) packaging system. The library contained 1.6×10^5 plaque forming units. Forty random plaques were chosen from each library for PCR amplification and were cut with the *Sfi*I enzyme. Average insert

*Corresponding author. E-mail: hspark@kribb.re.kr
First and second authors contributed equally.
DOI: 10.2983/035.030.0119

size was determined by agarose gel (0.8%) electrophoresis run at 180 V and compared with a standard ladder of 100 bp. The average insert size for the library was 1.2 kb.

DNA Sequencing

The lambda TripleEx2 vector (linear form) was converted to the pTriplEx2 vector (circular form) with *Escherichia coli*, which utilizes the Cre-loxP recombination system (Clontech, Mountain View, CA). Plasmid DNA isolated by the standard alkaline lysis method was amplified using the dye-terminator cycling method (BigDye v3.1) using 5′ universal primers in the vector (Stratagene), and single-pass sequencing using the ABI3730 XL capillary sequencer.

Sequence Analysis

The chromatogram files obtained from the sequencer were initially submitted to Phred (Ewing & Green 1998, Ewing et al. 1998) for base calling and quality assignment. The trace files were trimmed with trim-alt 0.12, and sequences shorter than 100 bp were removed. In addition, vector trimming was conducted with cross-match software (http://www.phrap.org), and poly-A was removed using the trimest program of Emboss package (Rice et al. 2000). The prepared Multi-FASTA formatted data were clustered and assembled using the TGICL (Pertea et al. 2003) package. Last, the contigs and singletons were analyzed with NCBI local BLAST (Altschul et al. 1990). Sequences were searched against the following databases: (1) Mollusks amino acid database (Lee et al. 2004), (2) NCBI nr database, and (3) NCBI KOG database. Matches with an E-value of 1e-5 for BLASTX were considered as putative hits. All 1,196 *P. acuta* EST sequences analyzed in this study were deposited in the DNA data bank of DDBJ/EMBL/NCBI (accession nos. BW985220–BW986415).

RESULTS

EST Sequence Compilation, BLAST Analyses, and Functional Clustering of Genes

A total of 2,286 clones were randomly selected from the *P. acuta* cDNA library, and their 5′ end sequences were determined by single-pass sequencing. All the sequences were trimmed from the vector region and the low-quality sequences (see Materials and Methods). We collected 1,196 high-quality ESTs with an average 718 bp in length, as shown in Table 1. After clustering and assembling the EST sequences using TGICL (Pertea et al. 2003), we finally obtained 575 distinct sequences composed of 504 singletons and 71 contigs in 62 clusters. The information about these 1,196 sequences, 575 distinct sequences, 71 contigs, 504 singletons, and BLAST results are provided as supplementary data at our website (http://blast.inje.ac.kr/~physa).

First we compared the 1,196 distinct sequences with an exclusive molluscs database (http://blast.inje.ac.kr/mollusks/ (Lee et al. 2004)) using the BLASTX program (<1e-10). No-hit sequences from the exclusive database were compared with the nonredundant database in the NCBI (nr) using the BLASTX programs as shown in Figure 1. To eliminate any mismatched sequences, an E-value of less than 1e-5 was used as the threshold against BLASTX. Based on the top Blast hits, 651 distinct sequences were putatively identified; 404 sequences aligned with protein sequences of the exclusive mollusc database and 247 sequences hit on the NCBI (nr) database. The remaining 575 of 1,196 sequences did not identify with sequences in the databases based on this criterion (Table1). The 360 putative sequences identified through BLAST searching were clustered by orthologous groups for eukaryotic complete genomes (KOG) (Tatusov et al. 2003). They were classified into 22 categories as represented in Figure 2: translation, ribosomal structure, and biogenesis (24.4%); RNA processing and modification (1.11%); replication, recombination, and repair (0.28%); chromatin structure and dynamics (1.11%); nuclear structure (0.28%); defense mechanisms (1.11%); signal transduction mechanisms (2.22%); cell wall/membrane/envelope biogenesis (0.28%); cell motility (0.28%); cytoskeleton (1.67%); extracellular structures (1.39%); intracellular trafficking, secretion, and vesicular transport (1.39%); posttranslational modification, protein turnover,

TABLE 1.

General characteristics of *P. acuta* ESTs.

Description	*n*
Total no. of cDNA analyzed	1,196
Average EST length	718
EST clusters	
Cluster	62
No. of unique sequences	575
Contigs	71
Singletons	504
Significant blast hit (KOG DB)	360/1,196
Total significant blast hit	651*/1,196
[A]Significant blast hit (mollusk amino acid database)	404/1,196
[B]Significant blast hit (NCBI NR database)	247/1,196

* A total of 651 distinct sequences contain 404 with a significant blast hit via the mollusc amino acid database[A] and 247 from the NR database of NCBI[B]. These are different from the number of EST clusters here.

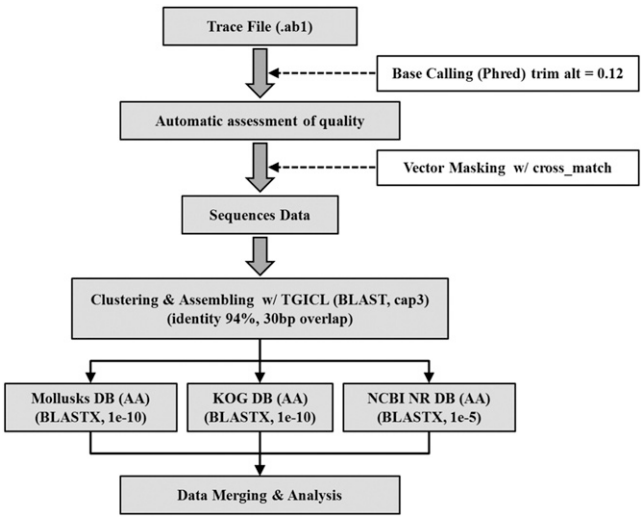

Figure 1. Schematic diagram of the EST analysis process.

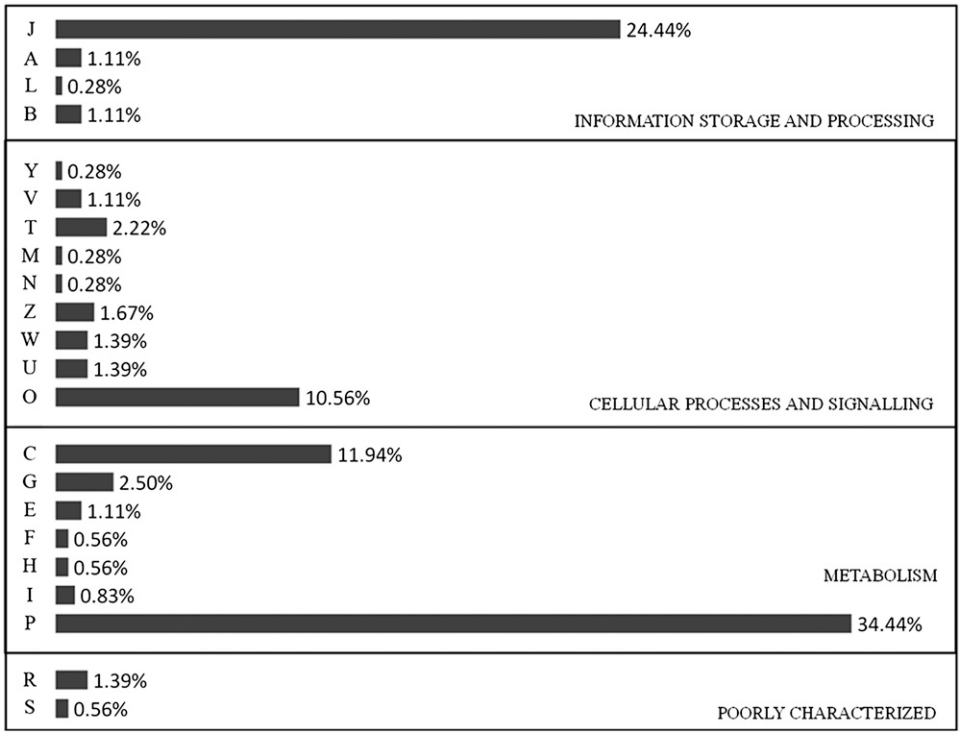

Figure 2. KOG analysis of EST sequences of *P. acuta*. Code descriptions of KOG: J, translation, ribosomal structure, and biogenesis; A, RNA processing and modification; K, transcription; L, replication, recombination, and repair; B, chromatin structure and dynamics; D, cell cycle control, cell division, and chromosome partitioning; Y, nuclear structure; V, defense mechanisms; T, signal transduction mechanisms; M, cell wall/membrane/envelope biogenesis; N, cell motility; Z, cytoskeleton; W, extracellular structures; U, intracellular trafficking, secretion, and vesicular transport; O, posttranslational modification, protein turnover, and chaperones; C, energy production and conversion; G, carbohydrate transport and metabolism; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport, and catabolism; R, general function prediction only; S, function unknown.

and chaperones (10.56%); energy production and conversion (11.94%); carbohydrate transport and metabolism (2.50%); amino acid transport and metabolism (1.11%); nucleotide transport and metabolism (0.56%); coenzyme transport and metabolism (0.56%); lipid transport and metabolism (0.83%); inorganic ion transport and metabolism (34.44%); general function prediction only (1.39%); and function unknown (0.56%).

DISCUSSION

Novel Biomarker Candidates of P. acuta

A variety of biomarkers using the EST approach has been successfully developed in several organisms (Davey et al. 2001, Gross et al. 2001, Jenny et al. 2002), but they have not yet been used in *P. acuta*. One of the important purposes of this EST study was to identify candidate genes with the potential of biomonitoring markers. We investigated these annotated genes through sequence alignment with already known biomarkers identified in other organisms, as summarized in Table 2. We identified 27 potential transcripts as biomarkers already validated in several other organisms (references are indicated in Table 2). They are classified into 3 functional groups: stress-related genes (9 transcripts), defense-related genes (12 transcripts), and others related to several functions (6 transcripts) as noted in the following paragraphs.

Stress-Related Genes

Stress-related proteins classified into biomarkers are known to have a functional role in detoxification processes triggered by stress. We found four representative gene categories with an E-value score less than 1e-5 with other organisms.

Generally, superoxide dismutase (SOD) is known to protect cells against cytotoxic effects of reactive oxygen intermediates produced during phagocytosis. It has been proposed that extracellular SOD may also participate in immunity by mediating or regulating hemocyte adhesion and phagocytosis (Johansson 1999). Ferritin has a metal-binding capability and is involved in iron metabolism and regulation. Recently, it has been suggested that ferritin may also be considered an acute-phase protein, and its sequestration may be a component of the invertebrate immune response (Beck et al. 2002). The third cluster, metallothionein, is a metal-binding protein thought to be involved in the detoxification of heavy metals, in free radicals scavenging, and also in inflammatory response (Kanekiyo et al. 2002). Several studies support the validity of metallothionein as a biomarker for a variety of cellular stress (Pedersen et al. 1997, Cosson 2000). Two heat-shock proteins (HSPs)—HSP 90 and HSP 24.1—were identified in the EST study. HSPs were validated previously as biomarkers for cellular stress by several studies (Bauman et al. 1993, Matranga et al. 2000). Based on these data, metallothionein expression in the absence of cadmium was also studied by PCR in our previous study (Jo et al. 2009).

TABLE 2.

Potential biomarkers of *P. acuta* related functionally to those of other organisms.

Category	Related Gene Name	Source (species)	NCBI ID	E-Value	Reference
Stress	Cd-metallothionein isoform	*Helix pomatia*	AAK84863.1	7.00E-24	Dallinger et al. (2004)
	Heat-shock protein 90	*Chlamys farreri*	AAR11781.1	4.00E-99	Buckley et al. (2001), Arts et al. (2004)
	Metallothionein	*Crassostrea angulata*	AF349907.1	1.00E-10	Dallinger et al. (2004)
	Small heat-shock protein 24.1	*Branchiostoma lanceolatu*	CAE83570.1	7.00E-17	Buckley et al. (2001), Arts et al. (2004)
	Superoxide dismutase	*Aplysia californica*	AAM44291.1	7.00E-14	Bebianno et al. (2004)
	Apoptosis-linked gene 2	*Suberites domuncula*	CAF74916.1	8.00E-50	Subramanian et al. (2004), Chen and Sytkowski (2005)
	Ferritin	*Lymnaea stagnalis*	AAB24081.1	1.00E-64	Nakada et al. (1984)
	Ferritin GF1	*Crassostrea gigas*	AAP83793.1	2.00E-46	Nakada et al. (1984)
	Thioredoxinlike protein p19	*Homo sapiens*	AAN34781.1	2.00E-22	Taskov et al. (2005)
Defense	Basic proteinase inhibitor	*Lymnaea stagnalis*	A59204	1.00E-11	Dietl and Tschesche (1975)
	Cathepsin L	*Mytilus galloprovincialis*	AAT39505.2	7.00E-48	Wang et al. (2005)
	Chymotrypsin-like serine proteinase precursor	*Haliotis rufescens*	P35003	8.00E-17	Groppe and Morse (1993)
	Cysteine protease	*Cercopithecus aethiops*	AAG35605.1	5.00E-14	Mikes and Man (2003)
	Homologue of Sarcophaga 26, 29-kDa proteinase	*Periplaneta americana*	BAA86911.1	1.00E-33	Ishikawa et al. 1992
	Ink toxin 1	*Aplysia punctata*	AAR14185.1	2.00E-23	Butzke et al. (2004)
	Mucin 2 74B10	*Homo sapiens*	NP_002448.1	1.00E-19	Marin et al. (2000)
	Selectin 1	*Biomphalaria glabrata*	AAC14140.1	2.00E-15	Duclermortier et al. (1999)
	Serine protease alpha	*Biomphalaria glabrata*	AAG40233.1	1.00E-36	Panyutich et al. (1997), Oliver et al. (1999), Menth et al. (2005)
	Serpin 6	*Ctenocephalides felis*	AAN73320.1	3.00E-11	Mandich et al. (2005)
	ppg3	*Leishmania major*	AAK31375.1	2.00E-12	Panyutich et al. (1997)
	Escapin precursor	*Aplysia californica*	AAT12273.1	2.00E-29	Butzke et al. (2004)
Other	Endo-1,4-beta-glucanase	*Mytilus edulis*	CAC59694.1	1.00E-46	Xu et al. (2000), Xu et al. (2001)
	Octopus dofleini hemocyanin mRNA, 3′ end	*Octopus dofleini*	J02835.1	7.00E-11	Cuff et al. (1990)
	Peptidoglycan recognition protein	*Argopecten irradians*	AAR92030.1	5.00E-45	Choe et al. (2005)
	QM protein	*Pinctada fucata*	AAN85578.1	9.00E-49	Zhanget al. (2004)
	Selenoprotein W1	*Danio rerio*	AAO86696.1	2.00E-15	Panyutich et al. (1997), Oliver et al. (1999), Menth et al. (2005)
	Ubiquitin	*Biomphalaria glabrata*	AAG49540.1	2.00E-70	Hegde et al. (2000), Buckley et al. (2001)

The cadmium was exposed in a time-dependent manner. As a result, we found that the expression of metallothionein was increased at 4 h and 8 h (Jo et al. 2009).

Defense-Related Genes

Defense systems present in every living species are involved in the elimination of reactive chemical species of endogenous or exogenous origin, neutralization of their effects, repair of initial lesions, and compensation of deficient metabolic pathways. Therefore, genes related to these functions have been actively studied as useful biomarkers. This is indicated in the reference column of Table 2. In this study, we identified 12 transcripts as potential biomarkers related to a defense function. They were comprised of proteinase and its inhibitors (7 transcripts), antimicrobial peptides (2 transcripts), lectin (1 transcript), and others (2 transcripts). Representative proteinase cathepsin-L is a cysteine proteinase belonging to the papainlike cysteinase family. The amphioxus cathepsin-L proteinase was reported to have a possible functional role in the inflammatory reaction in amphioxus (Wang et al. 2005). The protease inhibitors are known to be involved in the humoral immune response of invertebrates by inactivating proteases that are produced by invading pathogens (Kanost 1999). For example, the serine protease inhibitor functions in the control of the antifungal response in *Drosophila* (Levashina et al. 1999). Two protease inhibitors—basic proteinase inhibitors and serpin 6—were identified in this study by homology research. These 2 protease inhibitors may function to protect against microbial or fungal infections in their natural habitats.

We have identified 2 clusters—proprotegrin-3 and antimicrobial protein escapin—as being homologous to antimicrobial peptides. Such antibacterial peptides in invertebrates are known to be important components of the innate immune response (Charlet et al. 1996). One gene with a strong similarity to

selectin, a kind of lectin, was also identified in this study. Selectin is an important humoral component of innate immunity and is implicated in recognition mechanisms between immune cells and foreign bodies (Duclermortier et al. 1999).

Other Biomarker Genes

In this study, we identified 6 kinds of genes implicated in critical biological functions. Their putative functions are listed on the basis of the references in Table 2, although it remains to be proved experimentally whether unique and new environmental biomarkers can be identified from *P. acuta* in both laboratory and natural conditions.

CONCLUSIONS

Development of biomarkers indicating exposure to pollutants such as chemical or heavy metals is becoming increasingly important in toxicology and human health. The ecological and physiological characteristics of *P. acuta* show that it has the potential to be an index organism for monitoring ecological environments polluted with heavy metals and harmful chemicals. In this context, we identified 27 potential biomarker transcripts by analyzing the ESTs for *P. acuta*. To our knowledge, this study is the first to demonstrate high-throughput data analysis of the transcriptome for *P. acuta*. We believe that all the potential biomarkers are useful candidates for further studies to validate biomarkers in various environmental conditions. Furthermore, this approach will greatly expand our understanding about the physiological response to *P. acuta* to pollutants at the molecular level and will provide an alternative way to measure the level of water pollution.

ACKNOWLEDGMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Education, Science and Technology (2009-0073975). We also appreciate the use of the system development for application of genomic sequence information program supported by the Korea government (MEST).

LITERATURE CITED

Altschul, S. F., W. Gish, W. Miller, E. W. Meyers & D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.

Arts, M. J., R. O. Schill, T. Knigge, H. Eckwert, J. E. Kammenga & H. R. Kohler. 2004. Stress proteins (hsp70, hsp60) induced in isopods and nematodes by field exposure to metals in a gradient near Avonmouth, UK. *Ecotoxicology* 13:739–755.

Bacchetta, R., P. Mantecca & G. Vailati. 2002. Oocyte degeneration and altered ovipository activity induced by Paraquat in the freshwater snail *Physa fontinalis* (Gastropoda: Pulmonata). *J. Molluscan Stud.* 68:181–186.

Bauman, J. W., J. Liu & C. D. Klaassen. 1993. Production of metallothionein and heat-shock proteins in response to metals. *Fundam. Appl. Toxicol.* 21:15–22.

Bebianno, M. J., F. Geret, P. Hoarau, M. A. Serafim, M. R. Coelho, M. Gnassia-Barelli & M. Romeo. 2004. Biomarkers in *Ruditapes decussatus:* a potential bioindicator species. *Biomarkers* 9:305–330.

Beck, G., T. W. Ellis, G. S. Habicht, S. F. Schluter & J. J. Marchalonis. 2002. Evolution of the acute phase response: iron release by echinoderm (*Asterias forbesi*) coelomocytes, and cloning of an echinoderm ferritin molecule. *Dev. Comp. Immunol.* 26:11–26.

Buckley, B. A., M. E. Owen & G. E. Hofmann. 2001. Adjusting the thermostat: the threshold induction temperature for the heat-shock response in intertidal mussels (genus *Mytilus*) changes as a function of thermal history. *J. Exp. Biol.* 204:3571–3579.

Butzke, D., N. Machuy, B. Thiede, R. Hurwitz, S. Goedert & T. Rudel. 2004. Hydrogen peroxide produced by *Aplysia* ink toxin kills tumor cells independent of apoptosis via peroxiredoxin I sensitive pathways. *Cell Death Differ.* 11:608–617.

Canivet, V., P. Chambon & J. Gibert. 2001. Toxicity and bioaccumulation of arsenic and chromium in epigean and hypogean freshwater macroinvertebrates. *Arch. Environ. Contam. Toxicol.* 40:345–354.

Charlet, M., S. Chernysh, H. Philippe, C. Hetru, J. Hoffmann & P. Bulet. 1996. Innate immunity: isolation of several cysteine-rich antimicrobial peptides from the blood of a mollusc, *Mytilus edulis*. *J. Biol. Chem.* 271:21808–21813.

Chen, C. & A. J. Sytkowski. 2005. Apoptosis-linked gene-2 connects the raf-1 and ask1 signalings. *Biochem. Biophys. Res. Commun.* 333:51–57.

Choe, K. M., H. Lee & K. V. Anderson. 2005. Drosophila peptidoglycan recognition protein lc (pgrp-lc) acts as a signal-transducing innate immune receptor. *Proc. Natl. Acad. Sci. USA* 102:1122–1126.

Cosson, R. P. 2000. Bivalve metallothionein as a biomarker of aquatic ecosystem pollution by trace metals: limits and perspectives. *Cell. Mol. Biol.* 46:295–309.

Cuff, M. E., W. A. Hendrickson, J. Lamy, J. N. Lamy, K. I. Miller & K. E. van Holde. 1990. Crystals of the carboxyl-terminal functional unit from *Octopus dofleini* hemocyanin. *J. Mol. Biol.* 213: 11–15.

Cuppen, J. G., S. J. Crum, H. H. Van den Heuvel, R. A. Smidt & P. J. Van den Brink. 2002. Effects of a mixture of two insecticides in freshwater microcosms: I. Fate of chlorpyrifos and lindane and responses of macroinvertebrates. *Ecotoxicology* 11:165–180.

Dallinger, R., B. Lagg, M. Egg, R. Schipflinger & M. Chabicovsky. 2004. Cd accumulation and Cd-metallothionein as a biomarker in *Cepaea hortensis* (Helicidae, Pulmonata) from laboratory exposure and metal-polluted habitats. *Ecotoxicology* 13:757–772.

Davey, G. C., N. C. Caplice, S. A. Martin & R. Powell. 2001. A survey of genes in the Atlantic salmon (*Salmo salar*) as identified by expressed sequence tags. *Gene* 263:121–130.

Dietl, T. & H. Tschesche. 1975. Trypsin–kallikrein isoinhibitor k (type kunitz) from snails (*Helix pomatia*): purification and characterization. *Eur. J. Biochem.* 58:453–460.

Duclermortier, P., V. Lardans, E. Serra, F. Trottein & C. Dissous. 1999. *Biomphalaria glabrata* embryonic cells express a protein with a domain homologous to the lectin domain of mammalian selectins. *Parasitol. Res.* 85:481–486.

Ewing, B. & P. Green. 1998. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res.* 8:186–194.

Ewing, B., L. Hillier, M. C. Wendl & P. Green. 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res.* 8:175–185.

Groppe, J. C. & D. E. Morse. 1993. Molluscan chymotrypsin-like protease: structure, localization, and substrate specificity. *Arch. Biochem. Biophys.* 305:159–169.

Gross, P. S., T. C. Bartlett, C. L. Browdy, R. W. Chapman & G. W. Warr. 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the pacific white shrimp, *Litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. *Dev. Comp. Immunol.* 25:565–577.

Hegde, A. N., B. M. Broome, M. Qiang & J. H. Schwartz. 2000. Structure and expression of the *Aplysia* polyubiquitin gene. *Brain Res. Mol. Brain Res.* 76:424–428.

Ishikawa, M., T. Kubo & S. Natori. 1992. Purification and characterization of a diptericin homologue from *Sarcophaga peregrina* (flesh fly). *Biochem. J*. 287:573–578.

Jenny, M. J., A. H. Ringwood, E. R. Lacy, A. J. Lewitus, J. W. Kempton, P. S. Gross, G. W. Warr & R. W. Chapman. 2002. Potential indicators of stress response identified by expressed sequence tag analysis of hemocytes and embryos from the American oyster, *Crassostrea virginica*. *Mar. Biotechnol*. 4:81–93.

Jo, Y. H., M. K. Baek, S. W. Kang, J. B. Lee, I. S. Byun, S. H. Choi, S. H. Chae, J. H. Kang, Y. S. Han, H. S. Park & Y. S. Lee. 2009. Molecular cloning and expression pattern of metallothionein gene from the left-handed shell, *Physa acuta*. *Korean J. Malacol*. 25:223–230.

Johansson, M. W. 1999. Cell adhesion molecules in invertebrate immunity. *Dev. Comp. Immunol*. 23:303–315.

Kanekiyo, M., N. Itoh, A. Kawasaki, A. Matsuyama, K. Matsuda, T. Nakanishi & K. Tanaka. 2002. Metallothionein modulates lipopolysaccharide-stimulated tumour necrosis factor expression in mouse peritoneal macrophages. *Biochem. J*. 361:363–369.

Kanost, M. R. 1999. Serine proteinase inhibitors in arthropod immunity. *Dev. Comp. Immunol*. 23:291–301.

Lee, Y. S., Y. H. Jo, D. S. Kim, D. W. Kim, M. Y. Kim, S. H. Choi, J. O. Yon, I. S. Byun, B. R. Kang, K. H. Jeong & H. S. Park. 2004. Construction of BLAST server for mollusks. *Korean J. Malacol*. 20:165–169.

Leung, K. M. Y., N. J. Morley, E. P. M. Grist, D. Morritt & M. Crane. 2004. Chronic toxicity of tributyltin on development and reproduction of the hermaphroditic snail *Physa fontinalis*: influence of population density. *Mar. Environ. Res*. 58:157–162.

Levashina, E. A., E. Langley, C. Green, D. Gubb, M. Ashburner, J. A. Hoffmann & J. M. Reichhart. 1999. Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science* 285:1917–1919.

Mandich, A., E. Benfenati, M. T. D. Cronin, A. Goksoyr, B. E. Grosvik, W. Kloas, A. van Cauwenberge & L. Vigano. 2005. Environmental agent susceptibility assessment using existing and novel biomarkers as rapid noninvasive testing methods. *Ann. N. Y. Acad. Sci*. 1040:381–386.

Marin, F., P. Corstjens, B. de Gaulejac, E. de Vrind-De Jong & P. Westbroek. 2000. Mucins and molluscan calcification: molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel *Pinna nobilis* (Bivalvia, Pteriomorphia). *J. Biol. Chem*. 275:20667–20675.

Matranga, V., G. Toia, R. Bonaventura & W. E. Muller. 2000. Cellular and biochemical responses to environmental and experimentally induced stress in sea urchin coelomocytes. *Cell Stress Chaperones* 5: 113–120.

Menth, M., C. Schmutzler, B. Mentrup, C. Hoang-Vu, K. Takahashi, T. Honjoh & J. Kohrle. 2005. Selenoprotein expression in Hurthle cell carcinomas and in the human Hurthle cell carcinoma line xtc.Uc1. *Thyroid* 15:405–416.

Mikes, L. & P. Man. 2003. Purification and characterization of a saccharide-binding protein from penetration glands of *Diplostomum pseudospathaceum*: a bifunctional molecule with cysteine protease activity. *Parasitology* 127:69–77.

Nakada, H., S. Matsuura, T. Sawamura & Y. Tashiro. 1984. Topology of asialoglycoprotein receptor in rat liver subcellular fractions: a ferritin immunoelectron microscopic study. *Cell Struct. Funct*. 9: 391–406.

Nebeker, A. V. & G. S. Schuytema. 1998. Chronic effects of the herbicide Diuron on freshwater cladocerans, amphipods, midges, minnows, worms, and snails. *Arch. Environ. Contam. Toxicol*. 35: 441–446.

Oliver, J. L., T. D. Lewis, M. Faisal & S. L. Kaattari. 1999. Analysis of the effects of *Perkinsus marinus* proteases on plasma proteins of the Eastern oyster (*Crassostrea virginica*) and the Pacific oyster (*Crassostrea gigas*). *J. Invertebr. Pathol*. 74:173–183.

Panyutich, A., J. Shi, P. L. Boutz, C. Zhao & T. Ganz. 1997. Porcine polymorphonuclear leukocytes generate extracellular microbicidal activity by elastase-mediated activation of secreted proprotegrins. *Infect. Immunol*. 65:978–985.

Pedersen, S. N., A.- K. Lundebye & M. H. Depledge. 1997. Field application of metallothionein and stress protein biomarkers in the shore crab (*Carcinus maenas*) exposed to trace metals. *Aquat. Toxicol*. 37:183–200.

Pertea, G., X. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai & J. Quackenbush. 2003. *Tigr* gene indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19:651–652.

Rice, P., I. Longden & A. Bleasby. 2000. Emboss: the European molecular biology open software suite. *Trends Genet*. 16:276–277.

Stobaeus, J. K., G. E. Heath, R. M. Parkhurst, W. O. Jones & J. E. Webster. 1990. A laboratory study of the toxicity of the butanol extract of endod (*Phytolacca dodecandra*) on two species of freshwater fish and two species of aquatic snails. *Vet. Hum. Toxicol*. 32:212–216.

Subramanian, L., J. W. Crabb, J. Cox, I. Durussel, T. M. Walker, P. R. van Ginkel, S. Bhattacharya, J. M. Dellaria, K. Palczewski & A. S. Polans. 2004. Ca2+ binding to ef hands 1 and 3 is essential for the interaction of apoptosis-linked gene-2 with alix/aip1 in ocular melanoma. *Biochemistry* 43:11175–11186.

Taskov, K., C. Chapple, G. V. Kryukov, S. Castellano, A. V. Lobanov, K. V. Korotkov, R. Guigo & V. N. Gladyshev. 2005. Nematode selenoproteome: the use of the selenocysteine insertion system to decode one codon in an animal genome? *Nucl. Acids Res*. 33:2227–2238.

Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin & D. A. Natale. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41.

Wang, Y., S. Zhang, Z. Liu, H. Li & L. Wang. 2005. Characterization and expression of amphicl encoding cathepsin l proteinase from amphioxus *Branchiostoma belcheri tsingtauense*. *Mar. Biotechnol*. 7:279–286.

Xu, B., U. Hellman, B. Ersson & J. C. Janson. 2000. Purification, characterization and amino-acid sequence analysis of a thermostable, low molecular mass endo-beta-1,4-glucanase from blue mussel, *Mytilus edulis*. *Eur. J. Biochem*. 267:4970–4977.

Xu, B., J. C. Janson & D. Sellos. 2001. Cloning and sequencing of a molluscan endo-beta-1,4-glucanase gene from the blue mussel, *Mytilus edulis*. *Eur. J. Biochem*. 268:3718–3727.

Zhang, Y., J. Huang, Q. Meng, T. Jiang, L. Xie, Z. Wang & R. Zhang. 2004. Molecular cloning and expression of a pearl oyster (*Pinctada fucata*) homologue of mammalian putative tumor suppressor qm. *Mar. Biotechnol*. 6:8–16.