Ann. Limnol. - Int. J. Lim. 52 (2016) 65–74
© EDP Sciences, 2016
DOI: 10.1051/limn/2015023

Available online at:
www.limnology-journal.org

Identification and expression of proteolysis-response genes for *Macrophthalmus japonicus* exposure to irgarol toxicity

Kiyun Park[1], Chamilani Nikapitiya[1,2] and Ihn-Sil Kwak[1]*

[1] Faculty of Marine Technology, Chonnam National University, Yeosu 550-749, Korea
[2] Department of Aqualife Medicine, Chonnam National University, Yeosu 550-749, Korea

Received 12 January 2015; Accepted 29 June 2015

Abstract – *Macrophthalmus japonicus* is one of the most abundant macrobenthic animals in estuarine zone, which are transition areas between freshwater and marine environments. The crabs are potential bio-indicators reflecting aquatic sediment toxicity. So far, limited genetic research has focused on this group of animals largely due to the limited genomic information. In this study, we performed *de novo* transcriptome sequencing to produce the most comprehensive expressed sequence tag resource for *M. japonicus*, and identified stress response genes in *M. japonicus* exposed to antifouling biocides irgarol. Using 454 pyrosequencing, a total of 887 690 reads were obtained, which were assembled into 24 217 high-quality expressed sequence tags. These contigs were then clustered into 17 289 isotigs and further grouped into 12 923 isogroups. About 48% of the isogroups showed significant matches to known proteins based on sequence similarity. Moreover, irgarol toxicity induced up-regulation of stress response genes, associated with proteolysis and oxidation–reduction in biological process by Gene ontology analyses. Our data provide the most comprehensive transcriptome resource available for *M. japonicus*. This resource allowed us to identify genes associated with proteolysis processes, which facilitated the quantitative analysis of differential gene expression respond to irgarol toxicity. These data would provide foundation for future genetic and genomic studies of crustacean species and monitoring to freshwater as well as marine ecosystems.

Key words: Crustacean / irgarol / proteolysis / oxidation / 454 GS FLX

Introduction

An estuary is a partly enclosed coastal body of brackish water with one or more rivers or streams flowing into it, and with a free connection to the open sea. They form a transition zone between freshwater and marine environments. The areas are productive and dynamic systems with extreme and variable environmental conditions (Heip *et al.*, 1995). Heavy rainfall is well known that during floods, unusually large volumes of freshwater runoff from rivers can have severe impacts on marine habitats, particularly estuary habitats (Forbes and Cyrus, 1992; Van Woesik *et al.*, 1995). Various matters in freshwater may act as a pollutant in the sense of having detrimental effects on marine habitats. A convenient indicator system able to detect effects of freshwater runoff into marine habitats would be useful in assessing the impacts of disturbance stress on benthic organisms such as crustaceans. Crustaceans are an ubiquitous group of invertebrates inhabiting several aquatic habitats, from muddy lagoons to rocky shores, and they can accumulate contaminants from both water and food, to reaching levels of contamination in mussels higher than those measured in environment (Micheletti *et al.*, 2007). They are important species in coastal and estuarine areas which are often heavily impacted by natural and anthropogenic pollutants (Dam *et al.*, 2008).

Crabs are potential model species bio-indicators reflecting marine ecosystem, as well as being a commercially important species living along coastal areas in Asia (Park *et al.*, 2014). Crabs are capable of assisting the osmoregulatory mechanism by their behavior. Typically, it does not inhabit waters deviating much in salinity from normal sea water, but it can regulate osmotically in both dilute and concentrated sea water (Gross, 1955; Prosser *et al.*, 1995). Thus, crabs are a significant benthos for monitoring aquatic environments, such as freshwater as well as marine ecosystem. Despite this species abundance, potentiality for aquatic monitoring, and commercial

*Corresponding author: iskwak@chonnam.ac.kr, inkwak@hotmail.com

importance, biological information is not enough at the molecular level (Rothberg and Leamon, 2008; Ning *et al.*, 2013; Wang *et al.*, 2013a).

The *Macrophthalmus japonicus* crab is one of the most abundant macrobenthic animals of estuary mudflats and ubiquitously distributed in Australia as well as Asia regions involving in Korea (Kitaura *et al.*, 2002; Park *et al.*, 2014). It burrows on organic matter in the estuarine zone. Benthos of the tidal flats region play an important role in purification functions. Macrobenthos activity mediated decomposition of organic materials in sediments (Herman *et al.*, 1999). Thus, the crabs are marine invertebrate with an important role in material cycling and energy flow in ecosystems, as they are detritivores, feeding primarily on detritus which is sifted from the surface sediment (Otania *et al.*, 2010). More information on genetic aspects of *M. japonicus* would be valuable, considering its commercial importance and the model species for marine environmental monitoring.

The next-generation sequencing (NGS) using GS-FLX 454 pyrosequencer (Life Science/Roche) is well-suited for *de novo* transcriptome sequencing for the rapid production of sequence data with reduced time, labor and cost, and generates the longest reads (Hudson, 2008). Recent advances in NGS and bioinformatics show great potential for expanding EST databases for potentially any non-model organisms, thus paving the way for functional genomics on crustaceans. Therefore, the technologies have been favorably chosen for *de novo* transcriptome sequencing in some ecologically and economically important crustaceans species such as lobsters and shrimps (Kawahara-Miki *et al.*, 2011; Li *et al.*, 2012; O'Rorke *et al.*, 2012). However, investigations in *M. japonicus* have not been conducted, although the transcripts have been established from marine crustaceans with other technology (He *et al.*, 2013; Lv *et al.*, 2013; Zeng *et al.*, 2013).

Marine benthic environment involved an interface structure between water and substrate. It can adsorb toxicants, such as biocides, and accumulate with chronic effects in marine ecosystem (Hall *et al.*, 1999; Larras *et al.*, 2013; Mai *et al.*, 2013). Emerging booster biocides contamination raises particular attention in the marine ecosystem health (Ali *et al.*, 2013). A wide range of chemicals are used as antifouling biocides, which have very different physico-chemical properties and therefore differing environmental fates, behavior and effects (Thomas and Brooks, 2010). Irgarol is a common antifouling biocide and is highly toxic to non-target marine species at low $ng.L^{-1}$ concentrations (Sapozhnikova *et al.*, 2013).

The presence of irgarol biocides was revealed in coastal waters from Korea, USA, Singapore, Australia and Bermuda (Konstantinou and Albanis, 2004; Knutson *et al.*, 2012; Kim *et al.*, 2014). Irgarol was detected in wastewater with average concentrations up to 22 $ng.L^{-1}$ as well as in streams and small rivers with up to 14 $ng.L^{-1}$ (Luft *et al.*, 2014). In marine ecosystem, it is reported that irgarol measured up to 254 $ng.L^{-1}$ in water and up to 9 $ng.g^{-1}$ dry weight in sediments from Southern California recreational marinas (Sapozhnikova *et al.*, 2013). Irgarol detected in the Bay of Vilaine area (Brittany, France) exhibited a clear seasonal pattern, with highest concentrations recorded in June and July (Caquet *et al.*, 2013). Maximum concentration of detected irgarol was 2021 $ng.L^{-1}$ in coastal waters of Peninsular Malaysia (Ali *et al.*, 2013). Irgarol were also detected up to 67.64 $ng.L^{-1}$ in Korean coasts (Lee *et al.*, 2010). Irgarol may cause problems to the top coastal predators due to bio-accumulation through the food chain (Wang *et al.*, 2013a). It induced apoptosis through mitochondrial dysfunction and increased levels of intracellular reactive oxygen species (ROS) and an immediate ROS burst in cell line (Wang *et al.*, 2013b). Consequently, frequent detections of irgarol in sea water lead to the changes of molecular responses in marine organisms. It is necessarily need to establish genomic sequence resources of macrobenthic animals such as crabs for assessment of marine ecosystem health by irgarol toxicity. Mud crabs inhabiting in sediment, such as *M. japonicus*, are more suitable species as a bio-indicator for monitoring of toxic stress on marine benthic environment than swimming crabs.

In this study, we performed high-throughput sequencing of *M. japonicus* crabs using the 454 GS FLX platform methods. This provided the comparative transcriptome of *M. japonicus* under control and irgarol exposure condition. The aims of this study were to enrich the genomic resources for *M. japonicus*, and to identify potential genes inducing in irgarol toxicity. A further goal was to examine the patterns of gene expression as a biomarker to detect stress response in *M. japonicus* by antifouling biocides irgarol exposure. These results would provide an invaluable resource for future genetic and genomic studies of mud crab and related species. Furthermore, information generate from this study will be useful to identify candidate genes for monitoring irgarol toxicity in marine ecosystems.

Materials and methods

Sample preparation and irgarol exposure experiment

M. japonicus (Macrophthalmidae, Decapoda, Malacostraca) crabs, with an average length of about 3 cm, were collected from Yeosu Aquatic Products Market (Jeonnam, Korea). The crabs were placed in aerated glass containers filled with natural seawater at a density of 15 crabs per tank, and were fed daily with fish or clam meat for a week. Crabs were acclimated to a temperature of $20 \pm 1\,^{\circ}C$, salinity of $30 \pm 2‰$ and an alternating 12-h light–dark schedule for at least 1 week, and the water was changed daily and aeration was kept constant. Non-damaged mature specimens were then selected for testing.

The crabs were divided into two groups (20 crabs per group) and exposed to 10 $\mu g.L^{-1}$ irgarol and seawater (control) for 96 h. Analytical-grade Irgarol® (Irgarol 1051) was obtained from Sigma-Aldrich (USA). A stock solution of irgarol was prepared by dissolution of the solid in analytical grade acetone (100%), and the doses were

administered to obtain a final acetone concentration of $< 0.1\%$ in each treatment. It was stored at room temperature ($22 \pm 2\,^{\circ}C$) for the duration of the experiments. Working solutions were prepared by diluting the stock solution in sea water. During exposure, crabs were maintained in glass aquaria containing aerated sea water, in under same thermohaline conditions used in the acclimatization period. For exposure tests, water was changed daily, and fresh irgarol was added at predetermined concentrations. Exposure was conducted at a constant temperature ($20 \pm 1\,^{\circ}C$), and a 12:12 h light–dark photoperiod was used for all experiments. All experiments were conducted in triplicate using independent samples. For each group, the gill from total 60 crabs were collected after 96 h and samples were stored at $-80\,^{\circ}C$ until RNA extraction within 1 weeks.

RNA isolation and 454 sequencing

Total RNA from *M. japonicus* were obtained using TRIZOL® reagent (Invitrogen) in accordance with the manufacturer's instructions, and were cleaned using the Nucleospin RN-Clean up kit (Germany). RNA purity and concentration were checked using the NanoDrop2000 spectrophotometer (Thermo Scientific, USA) and 1% agarose gel electrophoresis. The purified RNA was stored at $-80\,^{\circ}C$ until further processing. Equal total 3 μg RNA were reverse-transcribed into cDNA using a SuperScript™ III RT kit (Invitrogen, CA, USA) as per the manufacturer's instructions. Irgarol treated and control samples were run on a sequencing plate. The 454 sequencing experiments were performed by the SolGent™ sequencing center (Korea).

Sequence assembly and functional annotation

The raw reads of 454 sequencing were first filtered to remove adaptor sequences and low-quality reads. All reads shorter than 40 bp were removed based on an assumption that small reads would fail to assemble and that they might represent sequencing artifacts. The preprocessed sequences of trimmed and size-selected reads were assembled using the Newbler v2.5 assembly program (Roche) (cDNA assembly mode) where all of the parameters were set to their default values. Default assembly parameters were used with the minimum overlap length of 40 bp and the minimum sequence identity of 90%. Singletons, *i.e.*, reads that were not incorporated into any contigs, were retained in the data set because many were likely to be fragments of low-expressed transcripts. Assembled contigs and singletons were pooled for subsequent analysis.

The assembled sequences were mapped against Swiss-Prot and the NCBI non-redundant (Nr) protein databases using BLASTX with an *E*-value threshold of 10^{-3}. Gene names were assigned to each assembled sequence based on the best BLAST hit. GoPipe was used to annotate the assembled sequences with gene ontology (GO) terms describing biological processes, molecular functions and cellular components (Chen *et al.*, 2005). GoPipe is a tool for integrating BLAST results into streamlined GO annotations for batched sequences.

Real-time RT–PCR confirmation of 454 GS FLX sequencing data

To validate induced expression of proteolysis genes in crabs exposed to irgarol, five differentially expressed genes were selected for quantitative RT–PCR analysis. The full names and sequence information of these genes are listed in Table 3. All primers in this study were chosen using Primer 3 software (Whitehead Institute for Biomedical Research, Cambridge, MA, USA). The mRNA level of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal control (KF804084) for RT–PCR.

For RNA samples, group containing 20 crabs were treated with $10\ \mu g.L^{-1}$ irgarol and another group containing 20 crabs were kept in seawater as a control group. Exposure period was 96 h and each group containing 20 crabs was utilized for subsequent analyses. All experiments were conducted in triplicate using independent samples. The solvent control ($< 0.1\%$ acetone) crabs that were used as controls were also measured in triplicate. Exposure experiment was conducted at a constant temperature ($20 \pm 1\,^{\circ}C$), and a 12:12 h light–dark photoperiod was used for all experiments. Total RNA isolation and cDNA synthesis from gill of *M. japonicus* were obtained using the same methods as for the transcriptome profiling in accordance with the manufacturer's instructions.

Quantitative RT–PCR amplification and measurements were conducted using an AB7300 Real Time PCR System (Applied Biosystems, Foster City, CA, USA) in accordance with the manufacturer's instructions. To quantify the cDNA, RT–PCR was conducted using a master mix with a final volume of 20 μL that contained 1 μL of cDNA template, 0.2 μM of each primer and $1 \times$ SYBR® Premix Ex Taq (Takara, Kyoto, Japan). The RT–PCR reactions were run with an initial denaturing at $95\,^{\circ}C$ for 5 min followed by 40 cycles of $95\,^{\circ}C$ for 15 s, $60\,^{\circ}C$ for 50 s and $72\,^{\circ}C$ for 1 min. The quality of the amplification was then assessed with an AB7300 Real Time PCR system (Applied Biosystems) to conduct dissociation curve analysis. The obtained cycle threshold (Ct) values were converted into mRNA copy numbers using standard plots of Ct *versus* log copy number by the $2^{-\Delta\Delta Ct}$ method, using AB7300 System SDS software.

All data were provided as relative mRNA expression reported as means ± standard error of the mean (SEM). The expression levels of *M. japonicus* GAPDH mRNA was used as an internal standard to normalize the expression levels (Nikapitiya *et al.*, 2014). For calculation of the relative mRNA expression, transcript levels of five regulated genes were normalized against the level of GAPDH in the same samples based on the standard curves. The normalized levels of the five gene transcripts in

Table 1. Summary of the sequencing and assembly of the intertidal mud crab, *Macrophthalmus japonicus*, transcriptome.

Raw reads (base pairs)	887 690 (368 697 206)
Clean reads	759 336
Contigs	24 217
Contig size N50	807
Average length of contigs	735
Isotigs	18 273
Isotig size N50	1029
Average length of isotigs	855
Isogroups	12 923
Singletons	128 354

the irgarol-treated group were then compared with those of the controls using one-way ANOVA test using SPSS 12.0KO (SPSS Inc., Chicago, IL, USA) at a significance level of $P < 0.05$. Each test consisted of at least three replicates.

Results

Sequencing analysis and assembly

Two types of cDNA samples, which represented control and irgarol treated crab tissues of *M. japonicus*, were prepared and sequenced using the 454 GS FLX platform. The two sequencing runs produced a total of 887 690 reads with an average length of 320 bases (Table 1). Newbler v2.6 was used with the default parameters to screen for adapter sequences and eliminate poor quality reads. After quality trimming and removal of adapter sequences, 759 336 (86% of the raw reads) reads remained in the assembly. Out of these, 630 982 (71.1%) reads assembled completely or partially into contigs and 128 354 (14.5%) reads remained as singletons. Further, redundants (758 reads; 0.08%), outliers (32 700 reads; 4.3%) and short reads (< 40 base pairs: 20 266 reads; 2.7%) were excluded during assembly.

Newbler's terminology for assembled reads comprised three elements: contigs (stretches of assembled reads that were free of branching conflicts), isotigs (continuous path through a set of contigs) and isogroups (groups of isotigs arising from the same set of contigs). For consistency, we use this terminology throughout this paper. Our data were assembled into 24 217 contigs, which grouped into 18 273 isotigs. Of these isotigs, 7153 (39.1%) contained only one contig, and the average number of contigs per isotig was 1.4. Assembly of the high-quality reads produced 24 217 contigs with an average length of 557 bp (N50 = 768 bp). Size distribution of these contigs is shown in Figure 1(A). Contigs were then assembled into 18 273 isotigs with an average length of 692 bp (N50 = 886 bp). The size distribution of isotigs is shown in Figure 1(B). The average contig coverage for each isotig was 1.4. The 18 273 isotigs fell into 12 923 isogroups, where 14 819 (77.4%) contained only one isotig (the average number of isotigs per isogoup was 1.2, Table 1). Although many singletons could represent useful lowly expressed transcripts, it is also possible that some are artifacts derived from cDNA synthesis, sequencing and contamination (Meyer *et al.*, 2009). PCR validation or re-sequencing is necessary to verify the validity of these singletons. Hence, these singletons were excluded from the following analyses.

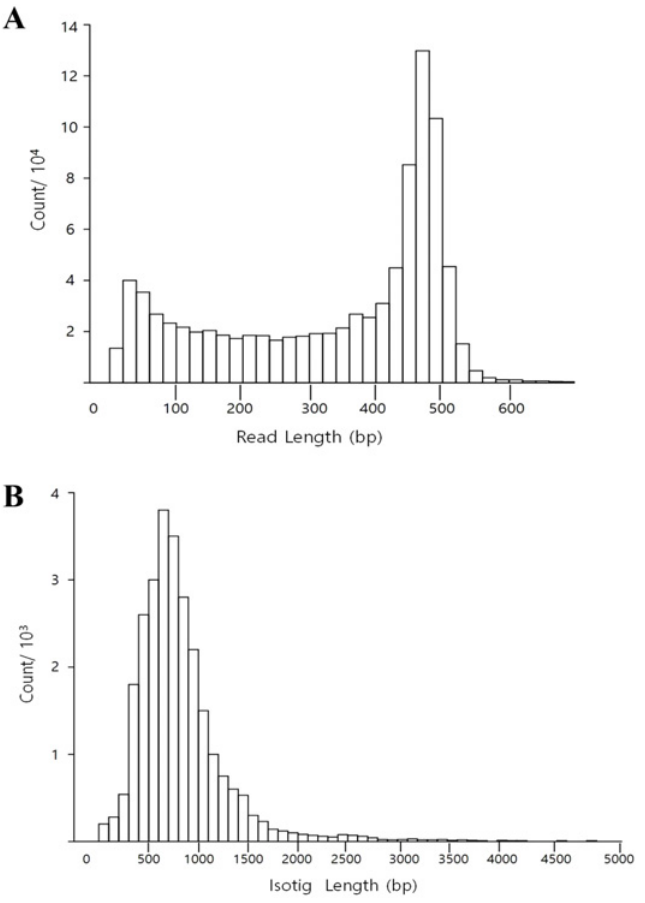

Fig. 1. Overview of *Macrophthalmus japonicus* transcriptome sequencing and assembly. (A) Size distribution of 454 sequencing after removal of adapter and short reads (< 40 bases). (B) Size distribution of isotigs.

Transcriptome annotation

Several complementary approaches were used to annotate the assembled sequences. First, we used BLASTX to map the 146 627 assembled sequences (18 273 isotigs + 128 354 singletons) against the entire RefSeq protein database with an *E*-value cut-off of 1×10^{-3}. To simplify the statistics, we only reported the results for the longest isotig per isogroup. Of the 12 923 isotigs, 6731(52.1%) had at least one hit and 6020 (46.6%) were matched to proteins with known functions. Out of the 128 354 singletons, 48 241 (37.6%) had hits and 39 124 (30.1%) were matched to known proteins.

GO assignments were used to classify the functions of the predicted genes. Based on sequence homology,

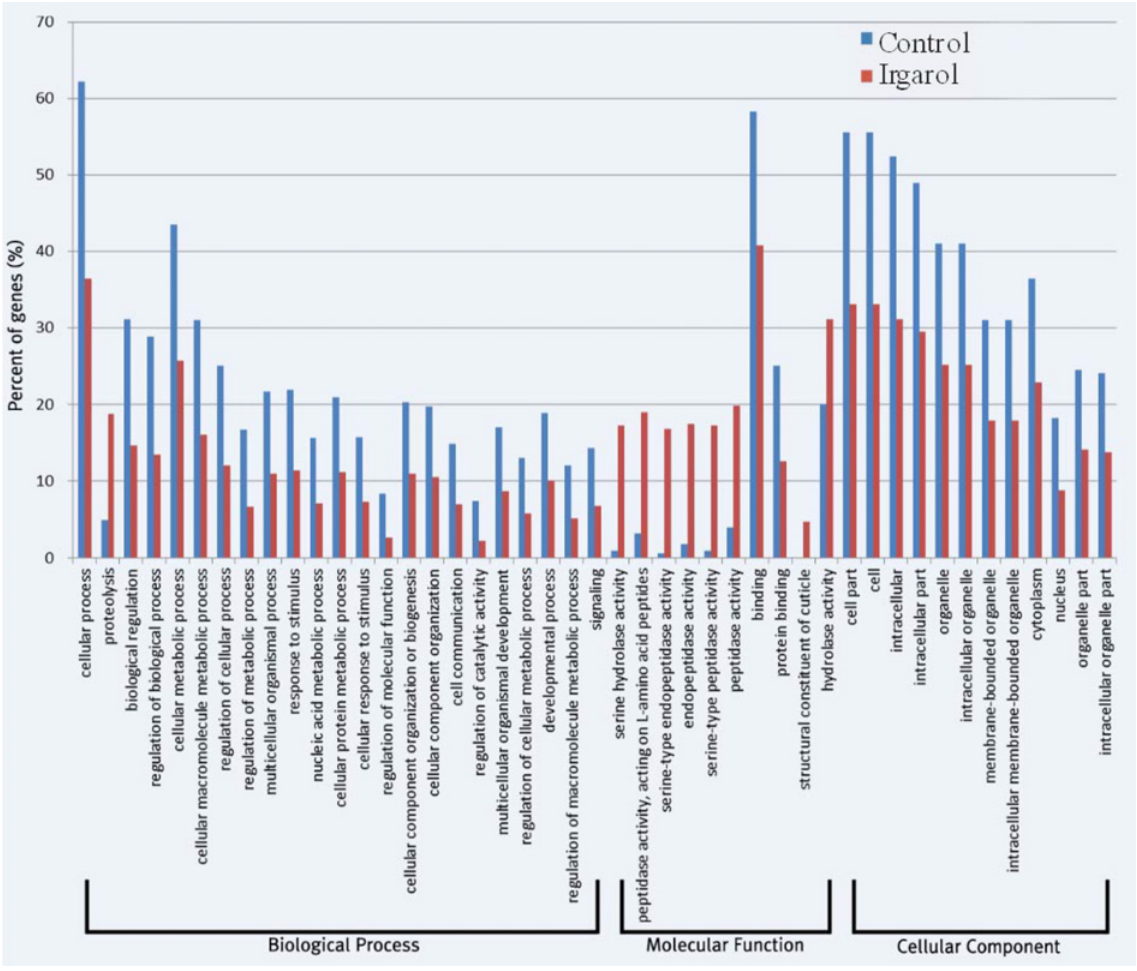

Fig. 2. GO term distribution of BLAST hits from the *Macrophthalmus japonicus* transcriptome. Selected GO categories are shown within the top-level divisions of Biological Process, Molecular Function, and Cellular Component. The relative percentages of genes falling into GO categories are comparable between our *M. japonicus* transcriptome (blue) and the *M. japonicus* exposed to irgarol genome (brown).

sequences can be categorized into 45 functional groups; the best hits from this query were extracted for GO classification using Uniprot2GO; each sequence was assigned at least one GO term. Second-level GO terms were used to classify the sequences in terms of their involvement as cellular components, in molecular functions, and in biological processes (Fig. 2). Of the annotated sequences, 17 905 were assigned GO terms. In total, 110 295 GO terms were obtained, which included 51.1% related to biological processes, 22.2% to molecular functions and 26.7% to cellular components (Fig. 2).

Selected GO categories in control crab samples were shown within the top-level divisions of Biological Process, Molecular Function, and Cellular Component. In generally, percentages of genes in control samples were higher than irgarol exposed samples in each GO categories. However, nine GO terms were assigned to annotate categories of up-regulated genes in irgarol exposure condition. Genes involved in the proteolysis categories (GO:0006508) was highly represented in biological process. Within molecular function, the most represented GO categories were serine hydrolase activity (GO:0017171), acting on L-amino acid peptides (GO:0070011), serine-type endopeptidase activity (GO:0004252), endopeptidase activity (GO:0004175), serine-type peptidase activity (GO:0008236), peptidase activity (GO:0008233), hydrolase activity (GO:0016787) and structural constituent of cuticle (GO:0042302) (Fig. 2).

In main species distribution matched against Nr database, 12.3% of the matched unigenes showed similarities with *Daphnia pulex*, followed by *Eriocheir sinensis* (11.4%), *Litopenaeus vannamei* (9.3%), *Portunus pelagicus* (6.4%), *Tribolium castaneum* (5.2%), *Scylla paramamosain* (4.8%), *Callinectes sapidus* (2.7%), *Nasonia vitripennis* (2.1%), *Marsupenaeus japonicus* (1.3%) and 44.5% of other species (Fig. 3).

Differentially expressed genes

GO enrichment analysis conducted to compare differently expressed genes between control and irgarol

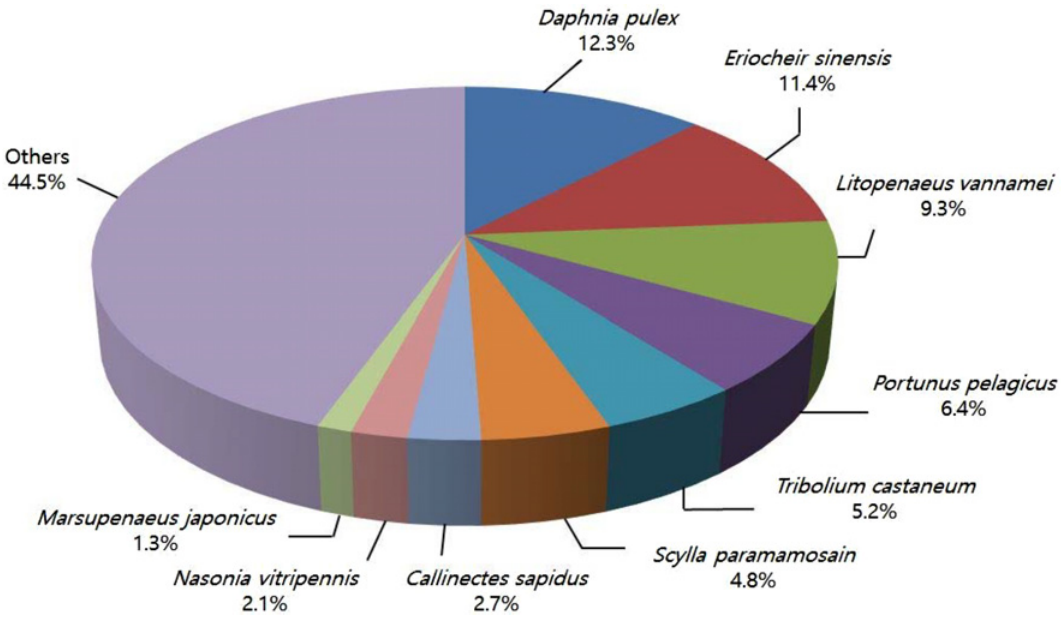

Fig. 3. Percentage of *Macrophthalmus japonicus* pyrosequence ESTs showing significant similarity with sequences in GenBank searched against NCBI. NCBI: National Center for Biotechnology Information, EST: expressed sequence tag.

treatment libraries. Comparing annotated GO terms, we identified over/under represented functional categories. GO terms with significant differences were deduced from statistical results in comparison with two groups by fisher's exact test. Comparison of gene expression showed that 4110 isotigs were differentially expressed between control and irgarol exposure samples. Total 40 genes involving in proteolysis of biological process were selected (Table 2). Selected genes were assigned to annotated categories of up-regulated genes in irgarol exposure condition.

To validate our 454 sequencing results, five differentially regulated genes were selected for quantitative real-time RT–PCR analysis. These genes involved caspase, chymotrypsin, trypsin, serine protease (SP) and heat shock protein 90 (HSP90). The sequences of the primers used are listed in Table 3. We investigated the mRNA responses of the five genes in *M. japonicus* exposed to 10 $\mu g.L^{-1}$ irgarol for 96 h (Fig. 4). The caspase gene expression significantly increased in *M. japonicus* as three times compare with solvent control. Expressions of chymotrypsin and HSP90 mRNA induced two times more in *M. japonicus* exposed to irgarol than the control group. Up-regulation of SP gene also observed in *M. japonicus* after irgarol exposure. The greatest increase in trypsin gene expression was observed in response to treatment with irgarol at 10 $\mu g.L^{-1}$ ($P < 0.01$). The expression levels of the five genes validated using real-time RT–PCR confirmed the robustness of the 454 sequencing results (Fig. 4).

Discussion

Organotin compounds, used as antifouling biocides on ships, marine structures and fishing nets, has become a problem, because of their toxicity and environmental accumulation (Cleary and Stebbing, 1987). Irgarol has been widely used as a booster biocide in copper-based antifoulant paints for controlling fouling organisms on the hulls of recreational and commercial watercraft (Bao *et al.*, 2013). As the consumption of these biocides increases, it gives a negative impact on physiology through the retardation of the development of eggs and embryos of aquatic species, leading to an increased mortality rate (Tsunemasa and Yamazaki, 2014). The sediment layers were thought to indicate the chronology of the biocide's accumulation.

The *M. japonicus*, inhabiting muddy flats of estuary ecosystem, are an excellent bio-indicator for monitoring of toxic stress on estuary sediment. Due to limited knowledge of genomic resources necessary for mechanistic study, the effect of toxic stress at the genomic level was rarely studied. However, there are no enough transcripts information of *M. japonicus* species to identify the molecular effects of *M. japonicus* following irgarol toxic stress. This work describes the first analysis of *M. japonicus* transcriptome using 454 pyrosequencing technology, and obtained a significant portion of its transcriptome. Moreover, we identified clusters of genes involved in a broad range of functional processes and characterized proteolysis response genes to antifouling biocides irgarol toxicity in *M. japonicus*.

The developing molecular techniques have been led a vast accumulation of genomic information in various marine species (Li *et al.*, 2012; O'Rorke *et al.*, 2012; Lv *et al.*, 2013). The former studies on *M. japonicus* transcriptome were performed using traditional cDNA library and sanger sequencing methods with RNA from various organs; however, it remains insufficient for the comprehensive understanding of *M. japonicus*

Table 2. Selected proteolysis process genes identified in the *Macrophthalmus japonicus* transcriptome.

Biological process (proteolysis)	Number of isotig	Length (range)
Aminopeptidase	8	514–1191
ATP-dependent metalloprotease FtsH	1	499
Carboxypeptidase	24	457–793
Calpain	11	477–2164
Caspase	3	422–2317
Cathepsin	29	349–1948
Chymotrypsin	2	140–1918
Chymotrypsin-like proteinase	8	151–375
Clip domain serine protease (SP)	19	476–1352
Collagenolytic SP	5	142–986
Cystein protease	3	483–1131
DN-damage inducible protein	2	454–566
Heat shock protein HtpX	1	480
Hepatopancreas trypsin	10	482–522
Heat shock protein 90	13	303–1289
Masquerade-like protein	32	463–1156
Metalloendopeptidase	1	459
Metalloprotease	1	464
Metalloproteinase	3	493–500
Methionine aminopeptidase	1	500
Progesterone-like protein	4	449–1074
Prophenoloxidase-activating enzyme III	2	463–871
Prophenoloxidase-activating factor	3	182–1176
Proteasome	4	764–1408
Proteasome subunit	9	486–880
Serine collagenase 1 precursor	33	357–780
SP	15	485–890
Serine proteinase	7	512–1548
Serine proteinase-like protein	1	680
Tolloid-like protein 2	1	1743
Tripeptidyl-peptidase 2	2	462–500
Trypsin	3507	257–540
Trypsin-like SP	277	283–864
Trypsinogen 1	13	109–522
UDP-glucose 4-epimerase	2	492–1123
Ubiquitin carboxyl-terminal hydrolase	7	649–1817
Ubiquitin-specific peptidase	4	485–578
Ubiquitin carboxyl-terminal esterase L3	1	455
Zinc metalloproteinase	31	403–507
Zinc proteinase Mpc1	10	284–616

transcriptome. Comparing with the transitional sanger-based methods, the high-throughput NGS techniques provided sufficient transcriptome data of crustaceans and prompted the genome studies of crustaceans. In our study, sequencing runs produced a total of 887 690 reads and then, we generated 24 217 contigs of *M. japonicus* transcriptome based on the NGS techniques (Table 1). Further GO analysis suggested to categorize for searching a potential genes responding to irgarol exposure on the *M. japonicus*.

In GO Enrichment analysis, percentage of genes in control samples were generally higher than samples exposed to irgarol in categorized 45 functional groups (Fig. 2). However, comparing with control percentage, high percentage of genes in irgarol samples observed in proteolysis categories of the biological process components (Fig. 2). The result suggests that irgarol toxicity may induce regulation of proteolysis on the crab *M. japonicus*. Proteolysis is a cleavage of proteins into smaller polypeptides or amino acids, typically either activate or impair the functionality of cellular proteins which occurs *in vivo* (Pham *et al.*, 2014). It is most commonly achieved by the activity of cellular enzymes called proteases. Many classes of proteases that cover a variety of physiological roles such as coagulation, metabolism and cell death has been reported (Kojro and Postina, 2009; Wysocka and Lesner, 2013). Proteolysis process can occur during cellular homeostasis or can be induced due to external stress stimuli such as heat, biological or chemical disturbance (Pham *et al.*, 2014).

In the present study, the gene percentages of eight GO categories were higher in irgarol samples than the control samples within molecular function (Fig. 2). Eight GO categories in molecular function were serine hydrolase activity, peptidase activity/acting on L-amino acid peptides, serine-type endopeptidase activity, endopeptidase activity, serine-type peptidase activity, peptidase activity and hydrolase activity, and structural constituent of cuticle. Moreover, irgarol treatment on *M. japonicus* significantly induced up-regulation of proteolysis related genes, such as caspase, chymotrypsin, trypsin, SP and HSP90 (Fig. 4). SPs are the dominant class of proteolytic

Table 3. Primers of specific genes for real-time RT–PCR confirmation.

Gene	Primer sequences	Amplified size (bp)	GenBank Accession no.
Caspase	5′-CGTCTGCGATAAGAACGTCA-3′ 5′-GTAGTCCGGTCTCGTCCTC-3′	241	HM483599**
Chymotrypsin	5′-CTCACCAGGTGGCTCTCTTC-3′ 5′-CAGCCTGATGAGGGAAATGT-3′	242	KT221087**
Trypsin	5′-CCCAACACAAGACTGGGACT-3′ 5′-CAACAAGTGGTTGGGTGTTG-3′	208	KJ653261
SP	5′-CGCGAGATAGTGTTCAACGA-3′ 5′-CTCTCCTTGGCAGCAACTTC-3′	235	KT221089**
HSP90	5′-GAGAACCTGCAACAGCACAA-3′ 5′-ACGTACTCGGAAAGGGAGGT-3′	233	KT221088**
GAPDH	5′-AGGTTGTGGCTGTGAATGAC-3′ 5′-CTTGGAGTACTTCTCAAGG-3′	340	KF804084

**Accession #s to be deposited in GenBank.

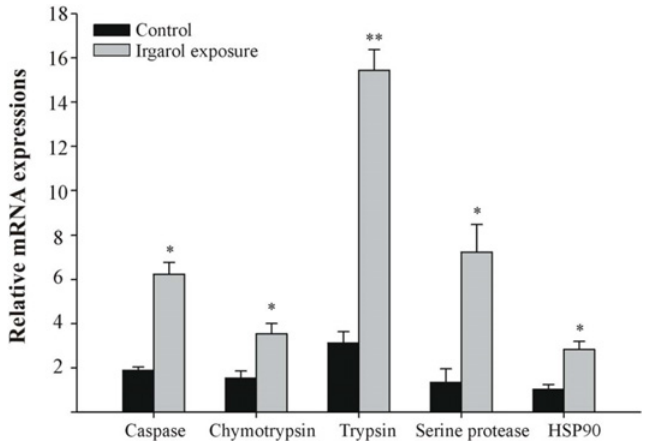

Fig. 4. Real-time PCR validation of differentially expressed genes may be involved in proteolysis. The values were normalized against GAPDH. Significant levels were $*P < 0.05$, $**P < 0.01$ and values indicate the mean ± SD.

enzymes in many insect species (Terra and Ferreira, 1994; Terra *et al.*, 1996; Park and Kwak, 2008), and constitute one of the largest families of enzymes in the animal kingdom (Barrett and Rawlings, 1995; Hedstrom, 2002). They play important roles in immune response, food protein digestion, embryogenesis and other biological progress in insects (Herrero *et al.*, 2005; O'Connell *et al.*, 2006; Sveinsdóttir *et al.*, 2006; Broehan *et al.*, 2007). The chymotrypsin-like SP is the enzyme involved in skin penetration. Previous studies have confirmed that chymotrypsin-like SP has the ability to degrade skin elastin, and the three sites of cleavage within elastin have been used to help define a new family of elastases (Salter *et al.*, 2000). Trypsin is an alkaline proteolytic enzyme which belongs to the SP family and acts as a digestive enzyme in many organisms including crustaceans which are involved in hydrolysis of lysine and arginine (Kishimura *et al.*, 2007). Irgarol induces apoptosis through mitochondrial dysfunction and oxidative stresses in human cell line (Wang *et al.*, 2013b). Irgarol exposure also induces oxidative damage, either by increasing the generation of ROS, decreasing anti-oxidant defenses, or both on the coral *Madracis mirabilis* (Downs and Downs, 2007). Caspases, aspartate-specific cysteine proteases, have fate-determining roles in many cellular processes, including apoptosis, differentiation, neuronal remodeling, inflammation and immunity in multicellular organisms (Yuan and Kroemer, 2010; Chávez-Mardones and Gallardo-Escárate, 2014). Additionally, HSP90 constitutes targets of caspase-mediated proteolysis during apoptosis triggered by ionizing radiation (Prasad *et al.*, 1998). Proteolysis is an irreversible posttranslational modification that directs the fate of cells in apoptosis, protein degradation, blood coagulation and many other processes (Morgan *et al.*, 2014). Results in the present study suggest that the proteolytic genes have significant roles for cellular fate determination during apoptosis or immune response induced by irgarol toxicity. Irgarol is essentially stable and persistent in the aquatic environment with an estimated half-life of about 100 days in seawater (Scarlett *et al.*, 1997; USEPA, 2010). Consequently, life cycle of marine organisms was influenced by chronic effects of irgarol toxicity. Comparison of gene expression between the control and irgarol treatment groups in the current experiment is helpful for identification of candidate genes underlying response to irgarol stress in *M. japonicus*. These transcriptome information can make use of environmental monitoring for chronic effects of irgarol toxicity on estuary ecosystem.

Acknowledgements. This study was supported by the National Research Foundation of Korea, which was funded by the Korean Government [grant no. NRF-2013-R1A2A2-01004914].

References

Ali H.R., Arifin M.M., Sheikh M.A., Mohamed Shazili N.A. and Bachok Z., 2013. Occurrence and distribution of antifouling biocide Irgarol-1051 in coastal waters of Peninsular Malaysia. *Mar. Pollut. Bull.*, *70*, 253–257.

Bao V.W., Leung K.M., Lui G.C. and Lam M.H., 2013. Acute and chronic toxicities of Irgarol alone and in combination with copper to the marine copepod *Tigriopus japonicus*. *Chemosphere*, *90*, 1140–1148.

Barrett A.J. and Rawlings N.D., 1995. Families and clans of serine peptidases. *Arch. Biochem. Biophys.*, *318*, 247–250.

Broehan G., Zimoch L., Wessels A., Ertas B. and Merzendorfer H., 2007. A chymotrypsin-like serine protease interacts with the chitin synthase from the midgut of the tobacco hornworm. *J. Exp. Biol.*, *210*, 3636–3643.

Caquet T., Roucaute M., Mazzella N., Delmas F., Madigou C., Farcy E., Burgeot T., Allenou J.P. and Gabellec R., 2013. Risk assessment of herbicides and booster biocides along estuarine continuums in the Bay of Vilaine area (Brittany, France). *Environ. Sci. Pollut. Res. Int.*, *20*, 651–666.

Chávez-Mardones J. and Gallardo-Escárate C., 2014. Immune response of apoptosis-related cysteine peptidases from the red abalone *Haliotis rufescens* (HrCas8 and HrCas3): molecular characterization and transcription expression. *Fish Shellfish Immunol.*, *39*, 90–98.

Chen Z., Xue C., Zhu S., Zhou F., Ling X.B., Liu G. and Chen L., 2005. GoPipe: streamlined gene ontology annotation for batch anonymous sequences with statistics. *Prog. Biochem. Biophys.*, *32*, 187–191.

Cleary J.J. and Stebbing A.R.D., 1987. Organotin on surface microlayer and subsurface water of southwest England. *Mar. Pollut. Bull.*, *8*, 238–246.

Dam E., Rewitz K.F., Styrishave B. and Andersen O., 2008. Cytochrome P450 expression is moult stage specific and regulated by ecdysteroids and xenobiotics in the crab *Carcinus maenas*. *Biochem. Biophys. Res. Commun.*, *377*, 1135–1140.

Downs C. and Downs A., 2007. Preliminary examination of short-term cellular toxicological responses of the coral *Madracis mirabilis* to acute Irgarol 1051 exposure. *Arch. Environ. Contam. Toxicol.*, *52*, 47–57.

Forbes A.T. and Cyrus D.P., 1992. Impact of a major cyclone on a southern African estuarine lake system. *Neth. J. Sea Res.*, *30*, 265–272.

Gross W.J., 1955. Aspects of osmotic regulation in crabs showing the terrestrial habit. *Amer. Nat.*, *89*, 205–222.

Hall L.W. Jr., Giddings J.M., Solomon K.R. and Balcomb R., 1999. An ecological risk assessment for the use of Irgarol 1051 as an algaecide for antifoulant paints. *Crit. Rev. Toxicol.*, *29*, 367–437.

He L., Jiang H., Cao D., Liu L., Hu S. and Wang Q., 2013. Comparative transcriptome analysis of the accessory sex gland and testis from the Chinese mitten crab (*Eriocheir sinensis*). *PLoS ONE*, *8*, e53915.

Hedstrom L., 2002. Serine protease mechanism and specificity. *Chem. Rev.*, *102*, 4501–4524.

Heip C.H.R., Goosen N.K., Herman p.m.J., Kromkamp J., Middleburg J.J. and Soetaert K., 1995. Production and consumption of biological particles in temperate tidal estuaries. *Oceanogr. Mar. Biol. Ann. Rev.*, *33*, 1–146.

Herman P.M.J., Middelburg J.J., Van de Koppel J. and Heip C.R., 1999. Nutrients in estuaries. *In:* Nedwell D.B. and Raffaelli D.G. (eds.), Advances in Ecological Research: Estuaries, Academic Press, San Diego, CA, 195–240.

Herrero S., Combes E., Van Oers M.M., Vlak J.M., de Maagd R.A. and Beekwilder J., 2005. Identification and recombinant expression of a novel chymotrypsin from *Spodoptera exigua*. *Insect Biochem. Mol. Biol.*, *35*, 1073–1082.

Hudson M.E., 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Resour.*, *8*, 3–17.

Kawahara-Miki R., Wada K., Azuma N. and Chiba S., 2011. Expression profiling without genome sequence information in a non-model species, Pandalid shrimp (*Pandalus latirostris*), by next-generation sequencing. *PLoS ONE*, *6*, e26043.

Kim N.S., Shim W.J., Yim U.H., Hong S.H., Ha S.Y., Han G.M. and Shin K.H., 2014. Assessment of TBT and organic booster biocide contamination in seawater from coastal areas of South Korea. *Mar. Pollut. Bull.*, *78*, 201–208.

Kishimura H., Tokuda Y., Yabe M., Klomklao S., Benjakul S. and Ando S., 2007. Trypsin from the pyloricceca of jacopever (*Sebastes schlegelii*) and elkhorn sculpin (*Alcichthys alcicornis*): isolation and characterization. *Food Chem.*, *100*, 1490–1495.

Kitaura J., Nishida M. and Wada K., 2002. Genetic and behavioral diversity in the *Macrophthalmus japonicus* species complex (Crustacea: Brachyura: Ocypodidae). *Mar. Biol.*, *140*, 1–8.

Knutson S., Downs C.A. and Richmond R.H., 2012. Concentrations of irgarol in selected marinas of Oahu, Hawaii and effects on settlement of coral larval. *Ecotoxicology*, *21*, 1–8.

Kojro E. and Postina R., 2009. Regulated proteolysis of RAGE and AbetaPP as possible link between type 2 diabetes mellitus and Alzheimer's disease. *J. Alzheimers Dis.*, *16*, 865–878.

Konstantinou I.K. and Albanis T.A., 2004. Worldwide occurrence and effects of antifouling paint booster biocides in the aquatic environment: a review. *Environ. Int.*, *30*, 235–248.

Larras F., Montuelle B. and Bouchez A., 2013. Assessment of toxicity thresholds in aquatic environments: does benthic growth of diatoms affect their exposure and sensitivity to herbicides? *Sci. Total Environ.*, *463–464*, 469–477.

Lee S.E., Won H.S., Lee Y.W. and Lee D.S., 2010. Study on the new antifouling compounds in Korean coasts. *Bull. Environ. Contam. Toxicol.*, *85*, 538–543.

Li C., Weng S., Chen Y., Yu X., Lü L., Zhang H., He J. and Xu X., 2012. Analysis of *Litopenaeus vannamei* transcriptome using the next-generation DNA sequencing technique. *PLoS ONE*, *7*, e47442.

Luft A., Wagner M. and Ternes T.A., 2014. Transformation of biocides irgarol and terbutryn in the biological wastewater treatment. *Environ. Sci. Technol.*, *48*, 244–254.

Lv J., Liu P., Wang Y., Gao B., Chen P. and Li J., 2013. Transcriptome analysis of *Portunus trituberculatus* in response to salinity stress provides insights into the molecular basis of osmoregulation. *PLoS ONE*, *8*, e82155.

Mai H., Morin B., Pardon P., Gonzalez P., Budzinski H. and Cachot J., 2013. Environmental concentrations of irgarol, diuron and S-metolachlor induce deleterious effects on gametes and embryos of the Pacific oyster, *Crassostrea gigas*. *Mar. Environ. Res.*, *89*, 1–8.

Meyer E., Aglyamova G.V., Wang S., Buchanan-Carter J., Abrego D., Colbourne J.K., Willis B.L. and Matz M.V., 2009. Sequencing and *de novo* analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics*, *10*, 219.

Micheletti C., Critto A. and Marcomini A., 2007. Assessment of ecological risk from bioaccumulation of PCDD/Fs and dioxin-like PCBs in a coastal lagoon. *Environ. Int.*, *33*, 45–55.

Morgan C.W., Julien O., Unger E.K., Shah N.M. and Wells J.A., 2014. Turning on caspases with genetics and small molecules. *Methods Enzymol.*, *544*, 179–213.

Nikapitiya C., Kim W.S., Park K. and Kwak I.S., 2014. Identification of potential markers and sensitive tissues for low or high salinity stress in an intertidal mud crab (*Macrophthalmus japonicus*). *Fish Shellfish Immunol.*, *41*, 407–416.

Ning J., Wang M., Li C. and Sun S., 2013. Transcriptome sequencing and *de novo* analysis of the copepod *Calanus sinicus* using 454 GS FLX. *PLoS ONE.*, *8*, e63741.

O'Connell A.R., Lee B.W. and Stenson-Cox C., 2006. Caspase-dependant activation of chymotrypsin-like proteases mediates nuclear events during Jurkat T cell apoptosis. *Biochem. Biophys. Res. Commun.*, *345*, 608–616.

O'Rorke R., Lavery S., Chow S., Takeyama H., Tsai P., Beckley L.E., Thompson P.A., Waite A.M. and Jeffs A.G., 2012. Determining the diet of larvae of western rock lobster (*Panulirus cygnus*) using high-throughput DNA sequencing techniques. *PLoS ONE*, *7*, e42757.

Otania S., Kozukib Y., Yamanakab R., Sasaokac H., Ishiyamac T., Yoshihito O., Sakaic H. and Fujiki Y., 2010. The role of crabs (*Macrophthalmus Japonicus*) burrows on organic carbon cycle in estuarine tidal flat, Japan. *Est. Coast Shelf. Sci.*, *86*, 434–440.

Park K. and Kwak I.S., 2008. Expression of *Chironomus riparius* serine-type endopeptidase gene under di(2-ethylhexyl) phthalate (DEHP) exposure. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.*, *151*, 349–354.

Park K., Kwak T.S. and Kwak I.S., 2014. Vitellogenin Gene characterization and expression of Asian Paddle Crabs (*Charybdis japonica*) following endocrine disrupting chemicals. *Ocean Sci. J.*, *49*, 127–135.

Pham V.C., Anania V.G., Phung Q.T. and Lill J.R., 2014. Complementary methods for the identification of substrates of proteolysis. *Methods Enzymol.*, *544*, 359–380.

Prasad S., Soldatenkov V.A., Srinivasarao G. and Dritschilo A., 1998. Identification of keratins 18, 19 and heat-shock protein 90 beta as candidate substrates of proteolysis during ionizing radiation-induced apoptosis of estrogen-receptor negative breast tumor cells. *Int. J. Oncol.*, *13*, 757–764.

Prosser C.L., Green S.W. and Chow T.S., 1955. Ionic and osmotic concentrations in blood and urine of *Pachygrapsus crassipes* acclimated to different salinities. *Biol. Bull.*, *109*, 99–107.

Rothberg J.M. and Leamon J.H., 2008. The development and impact of 454 sequencing. *Nat. Biotechnol.*, *26*, 1117–1124.

Salter J.P., Lim K.C., Hansell E., Hsieh I. and McKerrow J.H., 2000. Schistosome invasion of human skin and degradation of dermal elastin are mediated by a single serine protease. *J. Biol. Chem.*, *275*, 38667–38673.

Sapozhnikova Y., Wirth E., Schiff K. and Fulton M., 2013. Antifouling biocides in water and sediments from California marinas. *Mar. Pollut. Bull.*, *69*, 189–194.

Scarlett A., Donkin M.A., Fileman T.W. and Donkin P., 1997. Occurrence of the marine antifouling agent Irgarol 1051 within the Plymouth Sound locality: implications for the green macroalga *Enteromorpha intestinalis*. *Mar. Pollut. Bull.*, *34*, 645–651.

Sveinsdóttir H., Thorarensen H. and Gudmundsdóttir Á., 2006. Involvement of trypsin and chymotrypsin activities in Atlantic cod (*Gadus morhua*) embryogenesis. *Aquaculture*, *260*, 307–314.

Terra W.R. and Ferreira C., 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol. A*, *109*, 1–62.

Terra W.R., Ferreira C., Jordao B.P. and Dillon R.J., 1996. Digestive enzymes. *In:* Bilingsley P. and Lehane M. (eds.), Biology of the Insect Midgut, Chapman and Hall, London, 159–194.

Thomas K.V. and Brooks S., 2010. The environmental fate and effects of antifouling paint biocides. *Biofouling*, *26*, 73–88.

Tsunemasa N. and Yamazaki H., 2014. Concentration of antifouling biocides and metals in sediment core samples in the northern part of Hiroshima bay. *Int. J. Mol. Sci.*, *15*, 9991–10004.

USEPA, 2010. Study of Discharges Incidental to Normal Operation of Commercial Fishing Vessels and Other Non-Recreational Vessels Less than 79 Feet (EPA 833-R-10–005), Accessed online 30 April 2012, http://www.epa.gov/npdes/pubs/vesselsreporttocongress_all_final.pdf.

Van Woesik R., Bevantier L.M. and Glayebrook J.S., 1995. Effects of cyclone Joy on nearshore coral communities of the Great Barrier Reef. *Mar. Ecol. Pro. Ser.*, *128*, 261–270.

Wang S., Hou R., Bao Z., Du H., He Y., Su H., Zhang Y., Fu X., Jiao W., Li Y., Zhang L., Wang S. and Hu X., 2013a. Transcriptome sequencing of Zhikong scallop (*Chlamys farreri*) and comparative transcriptomic analysis with Yesso scallop (*Patinopecten yessoensis*). *PLoS ONE*, *8*, e63927.

Wang L., Liang B., Li L. and Liu W., 2013b. Induction of HepG2 cell apoptosis by Irgarol 1051 through mitochondrial dysfunction and oxidative stresses. *Toxicol. In Vitro*, *27*, 1771–1779.

Wysocka M. and Lesner A., 2013. Future of protease activity assays. *Curr. Pharm. Des.*, *19*, 1062–1067.

Yuan J. and Kroemer G., 2010. Alternative cell death mechanisms in development and beyond. *Genes Dev.*, *24*, 2592–2602.

Zeng D., Chen X., Xie D., Zhao Y., Yang C., Li Y., Ma N., Peng M., Yang Q., Liao Z., Wang H. and Chen X., 2013. Transcriptome analysis of Pacific white shrimp (*Litopenaeus vannamei*) hepatopancreas in response to Taura syndrome virus (TSV) experimental infection. *PLoS ONE*, *8*, e57515.