Host transcriptomic response to oyster oedema disease (OOD) in *Pinctada maxima*

Viet Hung Nguyen
BSc (Hons)

Submitted in accordance with the requirements for the degree of Master of Research at Macquarie University, Australia

Macquarie University
Faculty of Science and Engineering
Department of Biological Sciences
Sydney Institute of Marine Science

August 2015

This thesis is written in the form of a journal article from Fish and Shellfish Immunology.

Declaration

I wish to acknowledge the following assistance in the research detailed in this report:

Prof. David Raftos

Dr. David Jones

Dr. Timothy Green

Dr. Emma Thompson

All other research described in this report is my own original work.

Viet Hung Nguyen

05/08/2015

Table of Contents

Abstract

Pinctada maxima is the most commercially important pearl oyster species in Australia. However production is currently suffering from oyster oedema disease (OOD), the symptoms of which include oedema of interstitial tissues resulting in swelling of tissues such as the kidney and mantle. There have been two well documented mass mortality events linked to OOD, one in the summer of 2006 and another in 2013, but a definitive cause is yet to be identified.

Here, the goal was to analyse the host transcriptomic response to disease, to better classify and understand OOD, and to test whether anti-viral responses were evident in OOD-affected oysters. RNA-seq next generation nucleotide sequencing analysed the host transcriptome of OOD-affected and OOD-unaffected *P. maxima* and a comparison of transcriptional responses identified a number of up- and down-regulated genes. These genes were further classified into functional biological pathways regulated during the disease state, providing an insight into the causative agent of OOD. Differentially expressed genes were able to be segregated into two main categories – general stress response related genes and immune response/wound healing related genes. Although the response was not characteristic to a viral infection, therefore ruling out a viral stressor, the response profile was not able to further implicate what remaining stressor could be the cause of OOD.

1. Introduction

Worldwide aquaculture production has increased dramatically over the last few decades. It is now comparable to the aquatic wild-harvest industry, at a production rate of 90.43 million tonnes per annum [1]. The rapid development of the global aquaculture industry has largely been a response to the growing food crisis and inability to maintain sustainable wild populations of aquatic species [2]. Shellfish, such as pearl oysters, have multiple uses as aquaculture species. They are reared not solely for food, but also for the harvest of non-edible products such as pearls [3]. This separates shellfish from other typical aquaculture species, as they have a place in multiple markets. Worldwide, approximately 27% of total aquaculture production is attributed to shellfish aquaculture [4]. However, after a period of rapid growth and expansion, the shellfish aquaculture industry is experiencing a major reduction in production, both due to economic pressure and a lack of major advancements in management practices, resulting in issues such as mass mortalities from diseases [5].

High production aquaculture industries are especially prone to mass mortalities, which lead to substantial economic loss. For example, shrimp white-spot disease plagued South East Asia in the early 1990s, with losses amounting to several billion dollars [6]. One major contributor to these mass mortality events are diseases driven by pathogenic agents. This is especially true for the shellfish aquaculture industry. Jones and Creeper [7] tallied many pathogens implicated in production losses among shellfish species, including *Vibrio spp.* in pearl oysters, haplosporidan parasites in rock oysters, microsporidan parasites in mussels, nematodes in scallops; and trematodes and *Perkinsus spp.* in abalones. In addition to pathogenic agents, the diseases suffered by aquaculture shellfish can also be caused by environmental factors [8], poor management practices [9], or a combination of factors [10]. For example, in 1987, mass mortality events of the pearl oysters (*Pinctada maxima*) in Western Australia were attributed to mass infection by *Vibrio spp.* resulting from sub-optimal management practices [11].

The diversity of disease-causing agents and affected species pose a problem with respect to disease identification. In many situations, extensive studies may be required to elucidate and validate the cause of a disease. In France, the first attempt to identify the cause of summer mortalities of juvenile *Crassostrea gigas* in the Bay of Morlaix failed to conclusively implicate any infectious agents, leading to an entirely different approach

required to eventually identify a *Vibrio* strain as the infectious agent [12]. In other cases of shellfish diseases, the causes remain completely unidentified [13]. Oyster oedema disease (OOD) [14] in pearl oysters is one such uncharacterised disease. OOD has caused mass mortalities in populations of farmed *P. maxima* in various areas of northern Australia. However, no conclusive cause of OOD has been identified through traditional approaches [15, 16]. Without a clear understanding of the aetiology of the disease, the formulation of improved management practices, cures or preventions is difficult, and alternative methods to analyse the cause of disease that can overcome the drawbacks of the traditional methods are required.

The analysis of host transcriptomic responses to disease is one such alternative method. Examination of the host transcriptome can help improve our understanding of disease states, identify common responses exhibited by different species, and help design more robust methods to identify and combat disease.

1.1 Management responses to disease

The prevalence of diseases result in constant shifts in management practices. Adapting practices to the presence of diseases to avoid mass mortality events involves three main strategies. The first is to improve animal husbandry techniques. For example, major improvements in survivability of pearl oysters came about after a mass mortality event in 1987 [11], when it was discovered that transportation practices back then placed undue stress on the oysters, resulting in greater susceptibility to infection. By altering transportation practices, the oysters were able to better cope with *Vibrio spp.* infections and therefore greatly reduced mortality.

The second management strategy involves the application of therapeutic chemicals as a prophylactic measure against infectious agents. This has been applied successfully to protect various aquaculture species, primarily fish, from serious biotic agents such as ectoparasites in *Piaractus mesopotamicus* [1] and *Streptococcus iniae* in hybrid striped bass [17]. However, therapeutic chemicals have yet to be widely explored in shellfish aquaculture [18]. Therapeutic chemicals in the aquatic environment can also negatively affect native organisms such as local microalga [19]. In addition, residual therapeutic chemicals can result in the development and retention of disease resistance in microbes, making disease outbreaks more difficult to treat [20-22].

The third strategy to control disease involves selective breeding for individuals with certain traits, such as disease resistance or accelerated growth. This is sometimes seen as a preferable alternative to therapeutic chemicals, because selection for disease-resistant individuals negates the necessity for repetitive usage of therapeutic chemicals [23, 24]. Selective breeding allows for disease resistance and other desirable traits to be established in the population and transferred from one generation to another, such as with the protection against the protistan parasite *Bonamia ostreae* by the oyster species *Ostrea edulis* in Rossmore, Ireland [25]. One disadvantage is the selection for particular traits often involves evolutionary trade-offs with other desirable characteristics. Long term inbreeding can, for example, cause reduction in yield and growth rate in the Pacific oyster *C. gigas* [26].

1.2. Disease identification and examination

In most cases, improvements in management practices need to be based on an understanding of the cause of a disease. The most common method used to identify the cause of a disease in shellfish aquaculture is histopathology. Here, the histopathological symptoms of the disease are used to identify potential causes, and each cause is investigated in turn. This can allow the direct identification of disease causing microbes. Histopathology has been varyingly successful in identifying the cause of disease in molluscs, dependent on the disease's aetiology. The major drawback of histopathology is that a lack of discovery may not indicate absence of pathogenic agent(s) [12].

Transcriptomic analysis of the host responses to disease now represents one of a number of alternatives to histopathology for disease identification. Many advancements have been made in genetic technologies over the last decade, making it possible to potentially identify the pathogenic agent/s responsible for the disease. Rather than identifying visible signs of a disease, whole transcriptomic analysis focuses instead on the host gene response to the disease, and how expression of specific genes changes under different stresses and disease states. Analysis of the host transcriptomic responses to diseases does not require focusing on investigating one cause at a time, and therefore has a higher potential for discovery of the cause of disease, regardless of whether the cause is a biological agent [27-29] or an abiotic factor [30]. This broadens the range of aetiological agents that can be identified using these methods. Whole transcriptome sequencing provides a broad understanding of the biology and aetiology of disease, based on host responses [31].

Finally, since shellfish are sessile or semi-sessile organisms, mass mortalities can occur very rapidly [32]. A diagnostic tool developed through understanding a host's transcriptional response to a disease allows for early detection of diseases, which is extremely beneficial to shellfish aquaculture industries.

1.3. Common technologies for host transcriptomic response analysis

Analysis of the host transcriptomic responses at various stages of different diseases has been done in numerous studies, and on a wide range of species, and therefore has broad applications. For instance in animals, the transcriptome analysis of the host response of non-human primates to an infection by the virus H5N1 has indicated that the severity of disease is dependent on differential gene expression during the early and late phases of disease [33]. In plants, the analysis of the *Rehmannia glutinosa* leaf transcriptome suffering from replanting disease found that there were excessive differential expression of genes involved in both the ethylene signalling and metabolism pathways [34].

Initially, transcriptome analyses were undertaken on a gene by gene basis. However, many multi-gene analytical methods have now been established. One well established tool for transcriptomic analysis is cDNA microarrays, which have the capability to quantify the expression levels of many genes of interest [35]. However, there are a number of disadvantages to utilizing microarrays. The design of microarray probes requires sequence information of genes known to be involved with the disease. Such tools requiring gene-specific probes are therefore only effective when specific sequence information is available. The amount of sequence data in the public domain for most shellfish aquaculture species is limited and therefore only a subset of genes can be analysed without additional sequencing projects. In addition, gene products are known to undergo substantial changes in terms of deletions, additions and realignments to form the final gene product. The purpose of these changes is to confer specific functions to the resultant strand that may not be available otherwise, such as for mediators of RNA-protein interactions [36]. Finally, the cost of transcriptome-wide screening by microarrays can be impractical, as the number of probes required can be very high. This also applies to similar technology such as in situ hybridization [37-39], which generally requires known expressed sequence tags [40] or cDNA sequences to act as a basis for hybridization. Additionally, there is a likelihood for cross-hybridization, which results in 'background noise' (visible cues for expression levels that are actually false readings) [41, 42].

Other methods that are commonly used for the transcriptomic analysis of disease are serial analysis of gene expression [43, 44], cap analysis of gene expression [45-47] and other technologies [24, 48, 49]. These approaches have the advantage of being high throughput and they allow for quantification of expression levels. However, they are based on Sanger sequencing technology, and so are generally costly and do not have the ability to distinguish between gene isoforms.

The most powerful and encompassing technology for host transcriptomic analysis is next-generation RNA-seq technology. This method is capable of transcriptome-wide analysis of gene responses without prior sequence information [50]. Next-generation RNA-seq technology is a novel high-throughput tool that produces large amounts of sequence data allowing transcriptome-wide analysis of RNA expression levels [50].

By sequencing the entire transcriptome, RNA-seq technology broadly identifies major differences in expression levels for many gene pathways between samples, providing substantially more comprehensive information on host transcriptomic responses than any other tool. In addition, RNA-seq allows for the discovery of novel genes and/or transcripts [51]. This insight into the biological machinery of a host substantially increases the amount of information available, including differences in expression levels of alternative splicing of transcripts [52] and isoforms [53, 54].

Its relative high cost is the singular limitation of utilizing RNA-seq technology. This forces researchers to be highly selective of the targets of their study, and generally they cannot select as many test subjects as desirable. This tends to limit the scope of a study, when experiments have to focus on a particular stage of a host's life cycle or a particular stage of a disease. Even when studies have been done on the same species at different stages of their life, the specifics may not be directly comparable [55, 56].

1.4. Response profiles

For RNA-seq technology to allow the aetiology of unknown diseases to be identified through the discovery of transcriptomic response of hosts, the transcriptomic response profile identified through RNA-seq analysis must be matched to response profiles to known stressors. A response profile to a specific stressor must contain a list of genes that are

differentially expressed in response to the stressor, so that a match to the response profile would then implicate the stressor as the cause of disease (and a limited or non-existent match would indicate otherwise). Response profiles must fulfil a number of requirements. Firstly, profiles must be distinct enough to allow for differentiation between stressors. The specificity will depend on the individual stressors, and whether or not genes exist that can be utilized across species to determine the source of stresses. Secondly, profiles must contain enough information to allow for comparison with results of existing transcriptomic studies, so that the tested response event can be determined as being consistent or inconsistent with previous response profiles.

A transcriptomic response study will yield its own list of up- or down-regulated genes/gene products, and when matched against these response profiles, can determine or at least narrow down the stressor involved in uncharacterised diseases. Figure 1 illustrates the overlap between genes that are regulated in response to varying stressors by *C. gigas*. In this study, 80 unique genes were found to be expressed only in response to extremes of temperature, and not to altered salinity, metal contamination or air exposure. Therefore, a host response profile matching those 80 unique genes in diseased animals would implicate temperature as a contributing factor. This is similar for all other stressors.

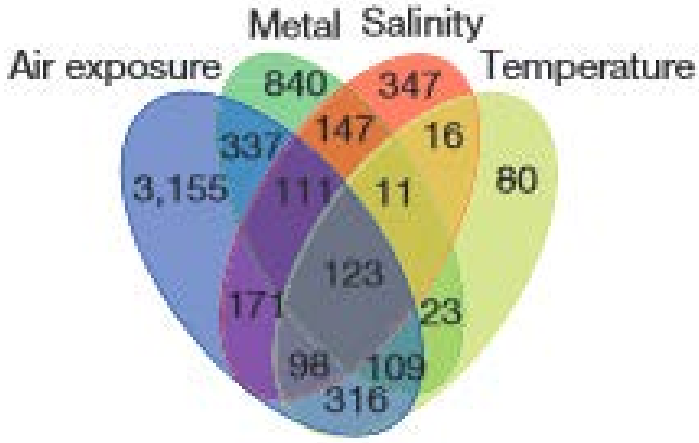


Figure 1: Adapted from Zhang et al. [57]. Venn diagram of common and unique genes expressed by *C. gigas* in response to temperature, salinity, air exposure and heavy-metal stress (zinc, cadmium, copper, lead and mercury), showing overlap of responses.

One current gap in our knowledge is that many specific response profiles remain unknown. The study by Zhang et al. on *C. gigas* only investigated four potential stressors [57], whereas there are many other stressors, including infectious agents that can affect productivity. In such cases, the number of unique genes found to be expressed in response to specific stressors will be limited.

Additionally, many response studies do not meet the criteria for forming response profiles. Studies that have managed to identify a large range of differentially expressed genes, or broad biological pathways, may not be specific enough to generate response profiles for diagnostic purposes. For example, Fu et al. was able to identify five major biological pathways differentially expressed in response to thermal stress in the Zhikong scallop *Chlamys farreri* [58]. However, they did not investigate specific genes/gene products in detail limiting the diagnostic value of the study. Many studies also lack information on the direction (up or down regulation) of the differential expression [59, 60], which may be important in dictating the presence of a disease state.

1.5. Transcriptomic response of shellfish to stressors

In the shellfish aquaculture industry, the focus of transcriptomic response studies has been on high value species, including oysters, mussels, scallops and abalones [61]. Overall, the research has been focused on gene discovery and broader expression differences, as opposed to uncovering specific genes that can act as diagnostic markers. Table 1 lists examples of commonly studied stressors, and the response of the host to the stressors. Although there have been studies of responses to both abiotic and biotic factors, the focus has been on abiotic stressors, predominantly environmental contaminants [62-65] and changing water quality parameters [58, 66-68].

Table 1: Differentially expressed genes in host responses to various commonly studied stressors of shellfish. Expression is marked with either a downward or upward arrow, to respectively identify down/up-regulation of the gene, gene family or biological pathway involved (if it was stated in the study). Those that were unclear/unstated are marked as '-'.

Stressor	**Differentially expressed genes/gene products**	**Species**
Abiotic		
Thermal	Apoptosis regulation-related genes (↓) mRNA binding-related genes (↓) Mitochondrial envelope formation-related genes (↓) Oxidation reduction-related genes (↓) Cytoskeletal protein binding/chaperones (↑)	*Chlamys farreri* [58]
	Translation-related genes (↓) Protein folding related genes (↑) Genes involved in chitin metabolism (↑)	*Mytilus galloprovincialis (Lam.)* [67]
Copper	*NADH2* gene (↓ 1-4 days, ↑ day 6 and 8) Heat shock proteins (molecular chaperones) (↑) GPx (a peroxidase enzyme) (↑) Cavortin (↑) Pernin (↑) Ferritin (↑) *Phr1* gene (↑) *IGF1* gene (↑) Alpha tubulin (↑) *EF1A* gene (↑) Tributylin binding protein type 1 (↑) Cellulase (↓)	*Argopecten purpuratus* (post-larvae) [65]
	Toll-like receptors (-) NOD-like and RIG-like receptors (-) Apoptosis pathway (-) Lysosome and C-type lectin (-)	*Mizuhopecten yessoensis* (gills) [63]

Copper (continued)	NADH dehydrogenase subunit 4 (↑ <24h, ↓ 24-168h) Ferritin (↓) Laminin (↓) Senescent protein (↓) *EF1α* gene (↓) EF2 genes (↓) Calponin 2 (cytoskeleton structural protein) (↓) *vdg3* gene (↓) LPr1 (signalling/multifunctional scavenger molecule) (↓)	*Haliotis rufescens* [64]
Cadmium	ABC, HSP and CYP protein families (↑) ADH dehydrogenase (↑) ATPase (↑) cytochrome c oxidase-related genes (↑) cytochrome P450 families (↑)	*M. yessoensis* [62]
Ocean Acidification (CO2)	Chitinase (catalyze the degradation of chitin polymers) (↓) Calponin-like protein (↓) TYR1 (a tyrosinase) (↑) F-ATPase subunits (from F.sub.O] and the [F.sub.1] complex) (↓) EFalpha (support binding of aminoacyl tRNA to ribosomes) (↑)	*Mytilus edulis* [66]
Salinity	*LTrpC-8* gene (↓) Na/Pi-cotransporter (↓) C-type lectin (↑) thioester-containing protein (↑) C1q domain containing proteins (↑) Molecules related to antimicrobial activity (↓) Heat shock protein 70 (↑) Ca^{2+}-binding proteins (↓) Apoptosis (↓) Ankyrins (↑) Tubulin (↑) Actin (↑) MEGF10 (↓)	*Crassostrea gigas* [68]

Biotic		
Bacterial	Scavenger receptors cysteine-rich (SRs) (-) C1qDC proteins (-) C-type lectins (-) Serine protease inhibitor-2 (↑) Glutathione s-transferase (-) Cytochrome p450 (-) Heat shock proteins (molecular chaperones) (-) Inhibitor of apoptosis (IAP) (-) GTPase of the immunity-associated protein (GIMAP) (-) epididymal secretory protein E1 (-) Cadherin (-) Legumain (-) vdg3 gene (-) Dermatopontin 2 (-) Apextrin (-) Furin (↑) Interleukin 17 (IL-17) (-) Arginase (nitric oxide modulator) (-)	*Crassostrea virginica* [59]
Viral	Myeloid differentiation 88 (MyD88) (↑) Interferon induced protein 44 (IFI44) (↑) Glypican (Gly) (↓) Inhibitor of nuclear factor kappaB kinase beta (IkB2) (↑) Inhibitor of Apoptosis (IAP) (↑)	*C. gigas* [69]
Combined Stressors		
Heat and nickel	Proteolysis-related genes (↑) Genes encoding small molecular chaperones (↑) Genes involved in the chitin metabolic process (↑)	*M. galloprovincialis* [70]
Heat and copper	Translation-related genes (↑) Genes encoding heat shock proteins (molecular chaperones) (↑) "Microtubule-based movement" proteins (↑)	*M. galloprovincialis (Lam.)* [67]
Heat and salinity	Molecular chaperones (↑) NADH dehydrogenase (-) Arginine kinase (-) Actin-binding regulatory proteins (-) GTPases (-) Mitogen-activated protein kinases (-)	Mytilus spp. [60]

Oysters have been more intensively studied than other molluscs, with *C. gigas* being a model organism for many genome-wide studies. These studies have delved into the interplay between differentially expressed genes in response to various stressors. They have revealed that genes can be differentially expressed in response specifically to a particular stressor, or in response to multiple stressors (Figure 1). Pearl oysters in particular are sought after for their pearls, more so than many other molluscs. Therefore, in addition to stressors, there have also been transcriptomic studies of pearl oyster species to understand the mechanics behind biomineralization, and its relationship to both biotic and abiotic cues [71-73]

There are a number of genes, gene products and pathways that have been identified in multiple response studies. Molecular chaperones (heat shock proteins) are differentially expressed in response to thermal [58], salinity [60], copper [67], nickel [70] and bacterial [59] stressors, indicating that they play a wide, generalistic stress-response role. In almost all cases, there is an up-regulation of molecular chaperones in response to stress.

Apoptosis-related genes and gene products are also commonly differentially expressed, being involved in response to thermal [58], copper [63], salinity [60], and viral [69] stress., with viral apoptosis-related genes specifically belong to a singular family (IAP), characterized by amino-terminal baculovirus IAP repeats (BIRs) [74].

NADH dehydrogenases are among the most common differentially expressed genes in response to stress, even though their expression is temporally variable. In response to copper contamination, expression of NADH2 in post-larval *A. purpuratus* was found to decrease over the first four days of exposure, before increasing once again between day six and eight post-copper contamination [65]. In contrast, expression of NADH dehydrogenase subunit 4 in *Haliotis rufescens* under the same environmental conditions has been shown to increase within the first 24 hours, before decreasing expression between 24 and 168 hours [64]. This highlights the significance of addressing the various stages of both the host and disease cycle.

Anderson et al. [75] undertook a meta-analysis of the effects of environmental stress on edible to identify the most commonly affected genes by stressors. Overall, in response to stress, they found more up-regulation than down-regulation of genes. The most commonly differentiated genes were also identified, and the top ten are listed in Figure 2. Anderson et

al. also identified that in response to certain broad categories of stressors, biological pathways are affected differently. Genes involved in transcription and/or translation were more commonly differentially expressed in response to temperature than other stressors. In contrast, metabolism-related genes were more commonly differentially expressed in response to hypoxia and contamination than the other two stressors.

These authors were also able to identify a general response pathway, whereby response to stress induces energy production, which produces by-products (reactive oxygen species) that could lead to apoptotic cell death, if not controlled by molecular chaperones and the anti-oxidant system.

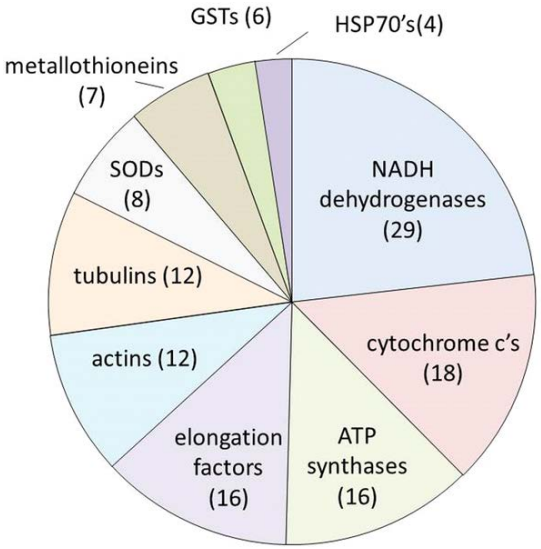


Figure 2: Adapted from Anderson et al. [75]. The top ten most commonly identified differentiated genes (not including those encoding ribosomal proteins) in response to ten different stressors. Numbers in brackets denote the number of times a gene was identified as differentially expressed in all treatments.

1.6. Shellfish response profiles to stressors

One clear finding of existing transcriptomic studies is that there are relatively unique genes/gene products that can be used to construct disease specific response profiles, particularly for viral diseases. Of the genes/gene products identified in Table 1, response to viral infection by *C. gigas* involves the up-regulation of two specific genes, myeloid differentiation factor 88 (MyD88) and interferon induced protein 44 (IFI44) [69]. These two genes are of significance, as they have been previously identified to be specific to viral infections [76-78]. Green and Montagnani [79] reported in addition toll-like receptor (TLR)

and protein kinase R (PKR) as also up-regulated in response to viral infection. A different study also found the inhibitor of Rel/NF-KB (IKB) and interferon-inhibiting cytokine (I), are also up-regulated during viral infections [80]. None of the genes/gene products involved in response to viral infection have been classified as up- or down-regulated in response to other stressors, including biotic stresses such as bacterial infections (Table 1). In non-shellfish models, interferon has also been implicated in responses to viral infections. For example, interferon is part of the response of *Marsupenaeus japonicus* to viruses such as white spot syndrome virus [81] and Taura syndrome virus [82]. However, the interferon involved in response to white spot syndrome virus by *M. japonicas* has been found to not be an alpha-interferon, but rather a mitochondrial F0-ATP synthase [83]. This could be the case in other organisms as well. Nonetheless, these data suggest that the current and main value of transcriptomic analyses of disease in molluscs is in the differentiation of viral diseases from those caused by other types of stress.

Table 2: A collated list of anti-viral responses in molluscs (*C. gigas* as a model organism), as reported by various authors.

Affected gene/gene products	**Response**	**Species**
Toll-like receptor (TLR)	Up-regulated	*Crassostrea gigas* [79]
MyD88	Up-regulated	*C. gigas* [69, 79]
Interferon regulation/induction (IRF/IFI44)	Up-regulated	*C. gigas* [69, 75, 79]
Protein kinase R (PKR)	Up-regulated	*C. gigas* [79]
Glypican (Gly)	Down-regulated	*C. gigas* [69, 75]
Inhibitor of nuclear factor kappaB kinase beta (IkB2)	Up-regulated	*C. gigas* [69, 80]
Inhibitor of Apoptosis (IAP)	Up-regulated	*C. gigas* [69]

On the other hand, bacterial infections and abiotic stressors induce very similar responses, as can be seen by response to bacterial infection by *Crassostrea virginica* in Table 1. Here, a C-type lectin, cytochrome p450, molecular chaperones, GIMAP, vdg3 gene and arginase are all shown to be differentially expressed mainly in response to thermal, metal and salinity stressors. Therefore, at the current state of knowledge, only transcriptional responses to viral infections are distinct enough for molluscs to generate a definitive response profile. The key genes involved in this antiviral response profile are shown in Table 2.

1.7. Transcriptomic response profile to OOD

In the current study, the Australian South Sea pearl oyster, *P. maxima*, has been chosen as the focus of differential gene expression analysis to develop a response profile for the currently uncharacterized disease, OOD. The goal is to determine whether OOD may be caused by viral infection.

Mass mortalities in populations of *P. maxima* have been reported in Western Australia since 2006 and no definitive causative agent has been identified to date. These mass mortalities have been attributed to oyster oedema disease (OOD), which Jones et al. [7] characterise by the reported loss of epithelial cells, mantle retraction and mild watery swelling (oedema) of the mantle tissues and palps, digestive gland, pericardial region and intestines of diseased *P. maxima* specimens. Affected oysters are slow to close their shells and high mortality (80 to 100%) can be evident, with thousands of freshly dead oyster shells (valves still attached at the hinge) present on cultured panels. No mortality was reported for other species settled on the affected panels (including other *Pinctada* species such as *P. margaritifera*).

Whilst OOD can be characterized by gross signs and histopathology, the cause of the disease remains unknown. No study has been able to show conclusively that the disease is caused by environmental or infectious factors (or a combination of the two) [15, 16]. One implication of a histopathological study that yields no conclusive results, is that the causative agent is not visible upon examination – such as is the case with many viruses. The enigmatic nature of OOD provides a valuable opportunity to functionally annotate and group transcriptomic data available for *P. maxima* suffering from this disease to identify potential gene families or pathways that can be informative in regards to identification of a causative agent. Specifically, comparison to known viral infection response profiles from other species provides the opportunity to test whether OOD is a viral disease.

Hence, the aims of the current study are to:

- Determine changes in gene expression in *P. maxima* associated with OOD.
- Identify specific pathways regulated during the disease state that may provide an indication of the causative agent of OOD, and specifically,
- Determine whether the transcriptomic response of *P. maxima* to OOD is characteristic of a viral infection or otherwise.

2. Materials and Methods

2.1. Sampling, RNA extraction and sequencing

A total of 20 *P. maxima* (adult production stock approximately 80-90 mm in size) were collected from two sites in north-western Australia in October 2013. The two sites were separated by approximately 600km of coastline. Additional site-specific information is confidential and cannot be disclosed by request of the pearling industry. Of these specimens, ten from one site fitted the formal case definition for OOD as defined by AusVet Animal Health Services and had obvious histopathological signs consistent with the criteria for this disease described in Jones et al. [16] (henceforth referred to as 'OOD-positive' specimens). The ten oysters from the other site did not fit the case definition, and did not have obvious histopathological signs of OOD (henceforth referred to as 'OOD-negative' specimens).

Tissue from five major organs (gills, gonads, mantle, foot and digestive tract) were collected from all 20 specimens. Immunocompetent cells in oysters are preliminarily hemocytes [84], which generally reside in high numbers in oyster gills [85, 86]. As such, the gills used for sample preparation would have contained large numbers of hemocyte suitable for analysis. Half of each tissue was stored in 4% formaldehyde in seawater for histopathological analysis to determine if oysters had clinical signs of OOD, and the other half was stored in *RNAlater*® (Ambion) at 4 °C overnight, then at -20 °C until RNA extraction. A modified Tri-Reagent (Sigma-Aldrich) protocol was used to extract RNA from the five tissues of each specimen [87].To ensure that the highest possible quality RNA was submitted for sequencing, a number of stringent quality control (QC) measures were undertaken including NanoDrop quantification and Illumina Bioanalyzer (Agilent) analysis.

Of the 20 specimens, four OOD-positive and four OOD-negative specimens were randomly selected and RNA from the five tissue types was pooled in equimolar amounts before being sent for library preparation and sequencing on an Illumina HiSeq 2000 platform at the Ramaciotti Centre for Genomics at the University of New South Wales. During library preparation using a TruSeq RNA Access kit, polyadenylation selection of mRNA molecules was undertaken and mRNA was fragmented to 200-300bp in length. mRNA fragments were then barcoded and pooled to be sequenced on one lane of Illumina's HiSeq 2000 platform (100bp paired-end reads).

2.2. Quality control

FastQC was used to access the quality of all eight libraries (4 OOD-positive specimens and 4 OOD-negative specimens), through an evaluation of various parameters. The ‘per base sequence quality’, ‘per sequence quality score’, ‘per base sequence content’ and ‘per base GC content’ were investigated to determine the quality of raw sequence data. For the eight libraries examined, a large portion of the reads were of high quality (bases with a quality score of 28 or above). A number of reads had quality scores that fell below the quality score of 28 and therefore trimming the reads based on their quality score was necessary.

QC of all sequences was conducted in Trimmomatic Version 0.32 [88] using the following parameters: a sliding window of 4:25, leading quality score trimming of 3, trailing quality score trimming of 3 and a minimum read length of 30. Trimming at this stringency ensured that the remaining reads were of high quality. The leading and trailing trimming scores were used to remove all bases of qualities 1 or 2, which are common unwanted artefacts from sequencing. This was verified by utilizing FastQC after quality trimming, to double-check the improved quality of the reads within each library.

Trimmomatic retained paired-end information of reads during sequence clean-up and removed entire sequences that post-enhancement consisted of lower than 30 base pairs. Some reads were retained post-enhancement, but were left without a complementary read. As sequence data is most valuable when existing as pairs, these sequences were ignored in further steps. Therefore, only paired, high quality sequences were kept for further analysis.

2.3. De novo assembly

As *P. maxima* does not have existing whole transcriptome or genome information available, the construction of the *P. maxima* transcriptome was performed *de novo* using the CLC Genomics workbench. First, the optimal kmer length was determined. The initial

assembly used an automated determination of the optimal kmer length. Once the kmer length provided by the automated software was obtained, kmer lengths both lower and higher in steps of 5 were tested manually (lengths 20, 25, 30, 35, 40, 45 and 50).

Two assembly procedures were also tested after defining the optimal kmer length of 25. The first procedure constructed a transcriptome using pooled data from all eight libraries at the optimal kmer length. The second procedure assembled libraries independently and removed redundancy in all assemblies resulting from the different kmer lengths using CD_Hit, a clustering software designed specifically for this purpose. A significance threshold of 0.95 was used, all other parameters were set to default. The output data showed that assembling all libraries using CD_Hit did not improve the quality of the final assembly. Therefore, the first procedure (simultaneous assembly of all libraries) was chosen to generate the assembly used in later steps.

2.4. RNA-seq analysis

RNA-seq analysis of the eight libraries (four OOD-positive versus four OOD-negative) was carried out in the CLC Genomics workbench. The analysis was performed for a number of candidate assemblies, chosen based on optimal values of contig length and quality. The RNA-seq analysis (legacy) tool from the CLC Genomics workbench was used to generate these initial analyses. The following parameters were chosen; a maximum number of mismatches of 2, minimum length fraction of 0.8, minimum similarity fraction of 0.8 and a maximum number of hits for a read of 20. All other parameters were set to default. This generated read count data for all eight libraries. The differential expression of contigs in the OOD-positive libraries versus the OOD-negative libraries was then undertaken in the CLC workbench. Empirical analysis of DGE (EDGE) was then performed to yield more statically precise results, allowing for better extraction of information on differential expression.

Contigs were deemed to be significantly differentially expressed between the OOD-positive and OOD-negative libraries when the FDR adjusted EDGE test p-value was <0.05. The differentiated contigs were then subjected to verification and manual annotation as described below.

RNA-seq results were subjected to principal component analysis (PCA) depicting either all contigs or the final subset of significantly differentially expressed contigs, relying on original expression values to group the eight libraries. Heat maps were also generated to graphically represent the expression levels of different contigs in all eight libraries, by sample clustering the libraries (again based on original expression values). Distance correlation (statistical dependence between two variables) is measured using Pearson correlation, with single linkage. Both the PCA and heat map were generated on the CLC Genomics workbench.

2.5. Annotation

Differentially expressed sets of transcripts were matched to known sequences (homologs) via Blast algorithms to determine the identity and putative functions of each differentially-regulated transcript. Blastx searches were performed against the NCBI Molluscan non-redundant protein sequence database.

Blast x annotations were then matched to Gene Ontology (GO) terms, through a mapping process that matched GO protein IDs to those sequences. Specific GO terms were then extracted from the pool for each mapped sequence, using default parameters (E-value of 1.0E-06, GO Weight of 5). This filtered the terms based on reliability and specificity. Domain/motif information was also extracted for each sequence within the blastx results, predicted through InterProScan 5.0 [89]. The scan processed sequences based on the presence of valid sequence strings.

Manual validation of the 147 most differentially expressed contigs was conducted to ensure their validity. The results were verified by cross-referencing with public databases, whereby the automated annotations was either validated and accepted, or modified if there was a discrepancy. For automated annotations without a clear result, manual annotation relied on Blasting against the NCBI non-redundant database first using blastn, before moving on to blastx if there are no conclusive results from the blastn similarity search.

2.6. Validation of expression levels

To accurately quantify the expression levels of genes of interest (identified through NGS), RT-qPCR was performed. This was to ensure that the semi-quantitative differential expression fold differences obtained from the RNA-Seq analysis were reliable.

Gene specific primers were designed using OligoCalc [90] for 15 genes that were found to be most differentially expressed in the RNA-seq analysis (see Table 3). Four reference genes were also chosen. A total of 24 primer sets were generated for the 15 contigs of interest and four potential reference genes, with two of the contigs of interest and three of the reference genes having a second primer set.

Table 3: Contigs of interest and RT-qPCR primer information. For the sequence information, 'F' denotes the forward primer, whilst 'R' denotes the reverse primer. A (Y) next to the primer name denotes the primer set chosen for the final validation of relative expression step.

Contig Number	Reference Gene?	Primer Name	Sequence (5'-3')	Amplicon Size
11368	No	P11368 (Y)	F: GATTCCTCACAAAGAAGCCTG R: GAATCGACCCATAATGATCCAC	83bp
12018	No	P12018	F: CTTCTTGGTGAACTGATAATCTG R: CTCTCTAATGTCCGCTCAAAG	90bp
1456	No	P1456-1	F: GACCTTTAACCTCACTTACCAG R: CTGTATGGCTACGTCAATCTTG	118bp
		P1456-2 (Y)	F: CTGTGATAGAACGACCAACAC R: CTGATCTGACTTCGCCCTC	83bp
14728	No	P14728	F: GATGCTGCCTTGTGAGAATG R: CATCCCAGCAGTGTCATAATG	110bp
30564	No	P30564 (Y)	F: CTGGTATAAGGTTCAACACGAG R: GTGTCCCATTCAACCAAACAG	110bp
3489	No	P3489	F: GACCCAACGACAGCGTTTG R: GAGTTCTGCGGATAATTCGTTC	94bp
3750	No	P3750 (Y)	F: GACGCGTGTCACGTACAG R: GATCCTATTGGTCATCGCAATC	112bp
3889	No	P3889 (Y)	F: GACCATATAGAAGGCAACCTC R: GTGCCACTCAAGGAAACTATC	87bp
48	No	P48	F: CACGACAACCGATACACCAG R: GACAGTTCGGTCCCTATCTG	104bp

6728	No	P6728	F: CTATGATGGCAAAGACGTGTC R: CACTTCCTCTACGATATGTATC	100bp
73	No	P73-1	F: GTATCGAATGGGAATTCATTGAC R: CATGCATTCCTCTTCAAGGATG	104bp
		P73-2 (Y)	F: CACTGCCAATAATCTGGAGAG R: GATTCTCTATTGACGCGATCTAC	111bp
73887	No	P73887 (Y)	F: GTTCCAATCAAATCTGCTTCATG R: GAACATCCACTGGGGAAATTAC	81bp
78929	No	P78929 (Y)	F: GAAATTGATGATGCAGAACCTTC R: CATCGTGCCATTTCCTTCATTC	90bp
78949	No	P78949 (Y)	F: CAAGGGACATTAGACTTTCCATC R: CTTTCCAGATATGATTACGTGTG	102bp
990	No	P990	F: CACCATCAATCGGTATCGAC R: GTGGTTCTGTTACCCTGGTC	103bp
1342	Yes	P1342-1	F: GAACATAATGAGAGCCAGTCTC R: GATCTATCCTCCTCTCCATC	89bp
		P1342-2 (Y)	F: GAAGATGGCAGCCACTGTC R: CAGTCCCAGAGCTTACTCAG	98bp
20	Yes	P20-1 (Y)	F: GTTGTCACCATGCCATCCAG R: GAGGCACGTTTCAGTGAAATTC	111bp
		P20-2	F: CTCTGACCTCAATTCAAACTGTG R: CAATAAGATGGACAGCACTGAC	115bp
21224	Yes	P21224	F: GAATTACCCGATGGTCAAGTC R: CACCTGGAGCGTTCAGTC	100bp
2161	Yes	P2161-1	F: CTGTGCAAGTTGTAGCTGTC R: CTTATACTGACCATGTGTAGAG	92bp
		P2161-2 (Y)	F: GTAGTGATTTCTGCTCCATCTC R: CACATTAAGATCGTTGGTGTAC	84bp

The amplification efficiency equation was **$E = [10(^{-1/M})] - 1$**, whereby an E (amplification efficiency value) of 100% indicated a doubling of amplicons for every RT-qPCR reaction, signifying that the primers were working correctly. A primer set did not meet the selection criteria if the E value was significantly higher than 100%, which signified that there was more than a doubling of RT-qPCR products per reaction; or a low value of E which signified that the primers were not able to properly amplify the target genes. A final nine out of the 17 contigs of interest primer sets and three of the seven reference gene primer sets (primer sets marked with a '(Y)' in Table 3) were selected from those that produced a single amplicon via the melt curve analysis for validation and had a coefficient (R^2) value

higher than 0.9. Here, three reference primer sets were tested to ensure a higher probability that at least one was able to act as a reference.

cDNA of four OOD-positive and four OOD-negative libraries was used for validation purposes. Whenever possible, the libraries chosen were the same as those that underwent RNA-seq analysis. The purpose of this validation step was simply to confirm that the RNA-seq analysis results were accurate in terms of fold difference. Hence validation focused on the eight libraries analysed by NGS, rather than all 20 libraries collected, to best avoid variabilities that may otherwise be captured. As C6A (an OOD-negative library) cDNA was not available, another OOD-negative library was chosen by random to make up the four x four analysis. RT-qPCR runs were performed for each primer set against each of the eight libraries.

To select the reference gene from the three potential reference primer sets chosen, standard deviation (sd) and coefficient of variation (cv) was calculated for each. The primer set (P1342-2) with the smallest sd and cv was chosen as the reference for all further analysis. The Livak method [91] was used to calculate relative eaxpression. All expression values are normalized to the reference gene (contig number 1342), and each OOD-positive expression value was then calculated to be relative to an OOD-negative expression value of 1.

Wilcoxon rank sum tests were carried out to compare the expression levels of each target amplicon to determine whether there were statistically significant differences in expression between cDNAs from four OOD-positive and four OOD-negative libraries.

The normalized relative expression values of the significantly differentially expressed contigs as determined by the Wilcoxon rank sum test were compared to the normalized relative expression values of the same contigs obtained from the RNA-seq analysis using the correlation coefficient similar to Liu et al. [92]. This determined if the relative expression values obtained from RT-qPCR validated the values from the RNA-seq analysis.

3. Results and Discussion

The RNA-seq analysis showed substantial differences in the transcriptional profiles of OOD-positive versus OOD-negative libraries, providing evidence that OOD is a distinct disease state in pearl oysters. The data indicated the presence of a typical wound healing/immune response that is characteristic of the general stress response pathway characterized by Anderson et al. [75]. There was a lack of genes that are known in other species to be differentially expressed in response to viral infection. The general stress response pathway identified, and the lack of differentially expressed viral response genes, suggests that the causative agent of OOD is not viral.

3.1. Quality control

The number of reads obtained per library ranged from 19 to 32 million, with a total of over 192 million reads over the eight libraries. Quality checking via FastQC yielded average quality scores in the medium to high range (a quality score of 20 to 28; and 28 and above respectively) for all eight libraries.

The trimming process removed 6% of all reads, comprising those that did not meet the 30bp cut-off point post-trimming. A further 19% of reads did not have complimentary reads (either forward or reverse) after the trimming process, and hence were not included in the downstream analysis. Therefore, 75% of reads (144 million) were retained after the QC process. All these reads were of high quality (quality scores of 28 or above).

3.2. Assembly results

3.2.1. Optimal assembly kmer length

Parameters for each assembly with kmers from 20 to 50 at steps of 5 were used to determine the optimal assembly settings (Table 4). As expected, the N50 values and average lengths decreased as kmer length increased, although the variations in N50 values were not large. The relative consistency of N50 results did not significantly promote one assembly over the other. However, the highest number of contigs of 1000 bp in length, the largest contig and the highest number of resultant bp peaked for the assembly using a kmer length of 25. A higher number of larger contigs were desirable for this study, as the

focus was for *de novo* transcriptome assembly. This process prefers a larger number of better, more fully assembled contigs, as *de novo* annotation relies on the ability to generate similarity matches to non-*P. maxima* organisms. Hence, the kmer length of 25 was deemed optimal. To confirm this, a further examination of the largest contigs for each assembly was conducted. The results indicated that there was little redundancy or contigs that could have potentially been merged together.

Table 4: Results for the different assemblies for each kmer length tested (at steps of kmer length 5). All other parameters were the same in each assembly.

Parameter	**Kmer 20**	**Kmer 25**	**Kmer 30**	**Kmer 35**	**Kmer 40**	**Kmer 45**	**Kmer 50**
Number of contigs	69,579	81,236	82,900	82,521	79,696	75,711	70,257
Total number of bp	34,113,692	39,246,796	38,919,791	37,492,046	35,187,020	32,255,854	28,869,006
Number of contigs with length of 1000 bp	5,351	5,547	5,212	4,537	4,036	3,389	2,771
N75	328	329	322	314	307	298	290
N50	509	498	480	459	442	422	401
N25	1,122	990	929	858	811	755	796
Average contig length	490	483	469	454	442	426	411
Largest contig (bp)	23,447	25,381	17,607	15,473	12,535	12,989	10,490

3.2.2. Optimal assembly method

To test all assembly options that could potentially yield viable, desirable results, a number of further assembly methods were examined. These included generating assemblies for individual libraries, before assembling all eight assemblies (designated ‘separate’ in Table 5) and generating an assembly of assemblies of all seven kmer lengths tested (designated ‘multiple kmer’ in Table 5). The output of these alternative methods were compared to the original method of a ‘single’ assembly of all eight libraries, using a kmer length of 25. From the results, it was determined that a single assembly of all eight libraries was preferable to either the ‘separate’ or ‘multiple kmer’ assemblies.

Although both the N50 and average contig length were higher for the assembly of individual library assemblies, the number of large contigs (determined by higher values for number of contigs of 1000 bp in length, the largest contig and the highest number of resultant bp) still indicated that the single assembly of kmer length 25 was the preferable of the two. Examination of the resultant contigs from the assembly of individual libraries identified more redundancy due to the collapsing of sequences for each library separately. This prevented many contigs from being further matched together during the second assembly. Conversely, individual assemblies meant that there were a larger number of shorter contigs produced that could not be assembled either, and hence were removed from the assembly. This explained the lower number of contigs obtained by this method (almost a six-fold decrease), despite a much higher proportion of larger contigs.

Table 5: A comparison of different assembly methods.*

	(A) Single Assembly	(B) Separate Assemblies	(C) Multiple kmer assembly
Number of Contigs	81,236	13,573	161,283
Total number of bp	39,246,796	11,240,225	77,092,263
Number of contigs with length of 1000 bp	5,547	2,803	10,873
N50	498	962	342
Average Length	483	828	478
Largest Contig (bp)	25,381	20,201	25,381

*Method A involves a single assembly of contigs from all libraries, using a kmer length of 25. Method B involves the assembly of contigs from each assembly separately, before assembling those assemblies together, both using a kmer length of 25. Method C involves the assembly of all libraries using different kmer lengths first (20 to 50, step of 5) before assembling all assemblies together.

In contrast, the multiple kmer assembly consisted of approximately double the number of contigs, total number of base pairs, and number of contigs with length of 1000 bp compared to the single assembly. A comparison of the distribution of contigs in terms of proportions indicated that it was very similar for both the multiple kmer and single assemblies (Figure 3). This suggests that the assembly of assemblies of different kmer lengths only doubled the contigs already present, producing more redundancy.

In summary, both of the two 'assembly of assemblies' yielded contigs that were unsuitable for further analysis. Regardless of the method, the assembly of assemblies yielded unnecessary duplicates, which would pose problems downstream in the RNA-seq analysis. It was therefore concluded that a single assembly was the optimal output for RNA-seq.

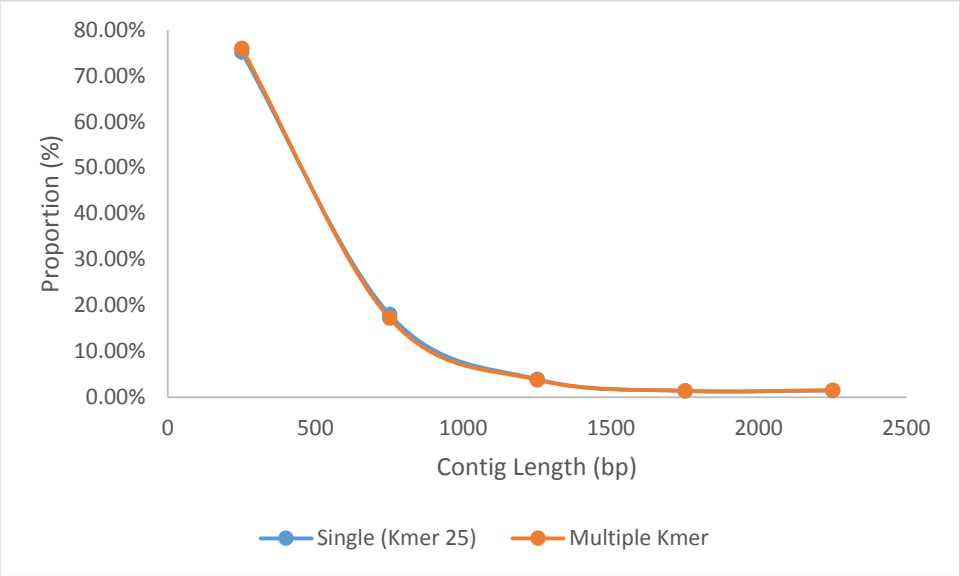


Figure 3: The contig length distribution for both the individual assembly of kmer 25, and the combined assembly of assemblies of all tested kmer lengths. The two distributions overlap over the majority of the contig length range.

3.3. Differential gene expression

RNA-seq analysis and subsequent DGE analysis identified 147 genes that were significantly differentially expressed between OOD-positive and OOD-negative libraries (FDR adjusted p-value < 0.05) in either direction (up-regulated and down-regulated in OOD-positive libraries versus OOD-negative libraries). The selection criteria used resembled 'double-filtering' [93], whereby large fold changes without significant p-values (as those fold changes can be caused by one or two large outliers) were ignored, as were those with low fold changes (yet with significant p-values), as these could be false signals simply due to the low variance. Through this method, both potentially undesirable types of outliers were ignored. Figure 4 depicts this selection procedure graphically, with red dots representing contigs of interest whilst blue dots represents non-differential contigs.

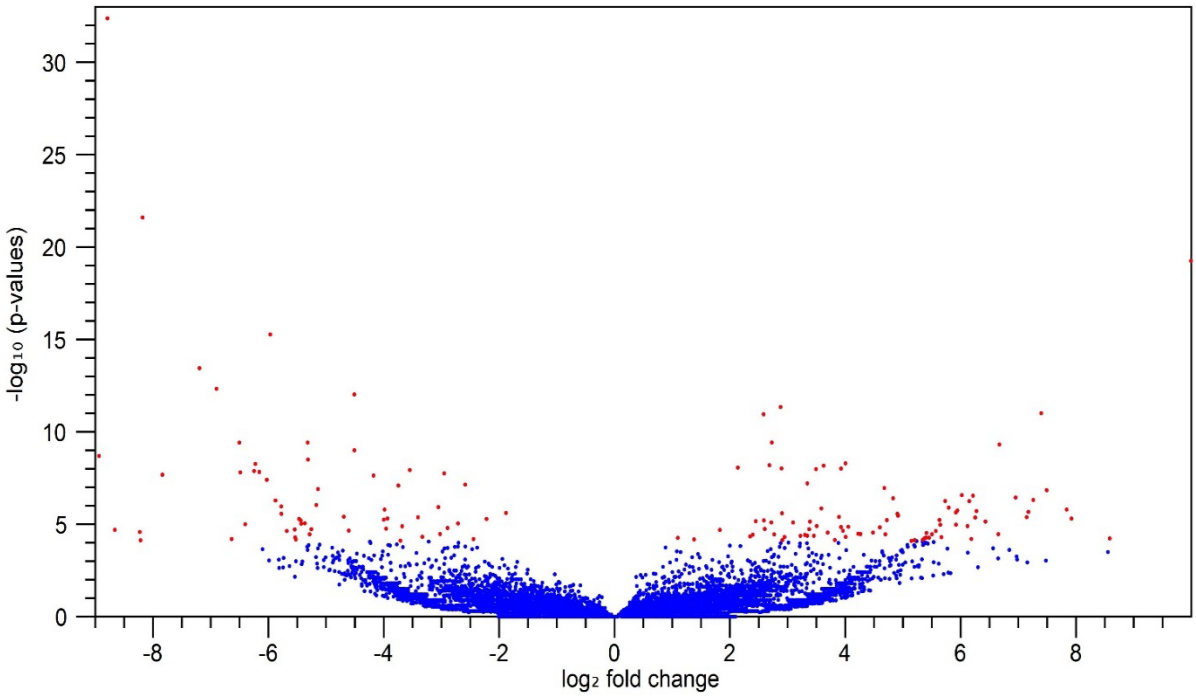
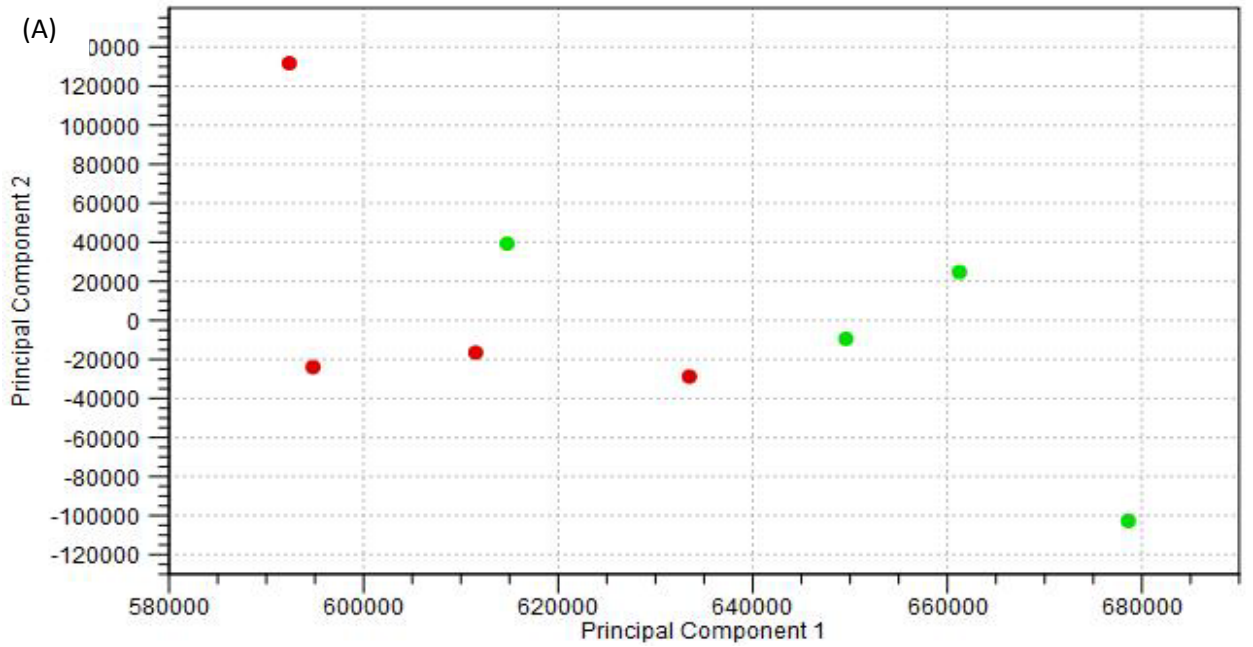


Figure 4: Volcano plot of the distribution of contigs in terms of fold change and p-value (non-FDR-adjusted) in RNA-seq comparisons of OOD-positive versus OOD-negative libraries. Red dots denote contigs that were deemed to be significantly differentially expressed. Blue dots denote no significant difference.

Principle component analysis (PCA) of the differential expression analysis of all 81,236 contigs showed a segregation of contigs from the OOD-positive libraries compared to those belonging to OOD-negative libraries (Figure 5A). This segregation is far more distinct for the subset of 147 contigs deemed to be significantly differentially expressed, as seen in the PCA of this subset (Figure 5B). It is also interesting to note that the cluster of OOD-negative libraries in the PCA of the 147 differentially expressed contigs appear to be more variable in transcriptomic responses. This suggests that OOD causes a restricted transcriptomic response in oysters, hence reducing the variability. However, it cannot be discounted that some of the OOD-negative oysters had sub-clinical infections that could lead to distinct transcriptomic responses, leading to greater variability in this group.



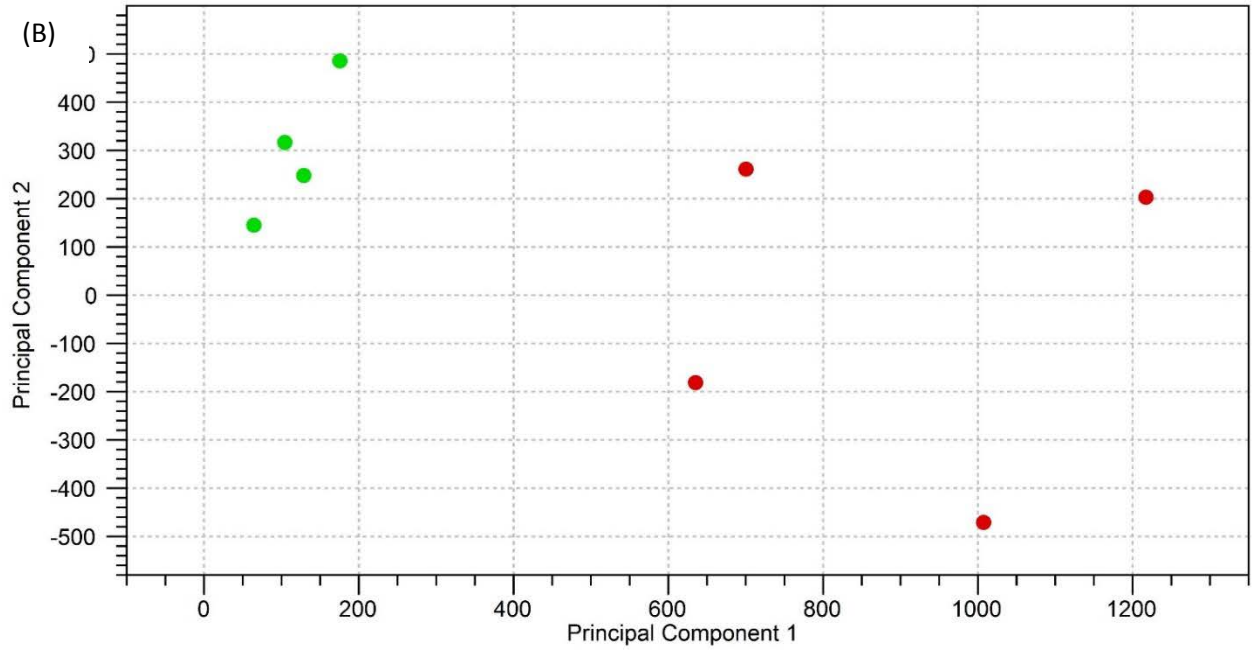


Figure 5: Principle component analysis of relative expression data for (A) the entire array of contigs obtained from the *de novo* assembly process and (B) the subset of 147 contigs deemed to be significantly differentially expressed between OOD-positive versus OOD-negative libraries. For both PCA plots, red dots signify OOD-negative libraries, whilst green dots signify OOD-positive libraries.

The validity of the 147 differentially expressed contigs was confirmed through a heatmap representation of their expression levels (Figure 6). Due to the large range of fold-differences (in either direction), not all differential contigs are visible under the resolution shown in Figure 6. However, the heat maps clearly show a marked division between OOD-positive and OOD-negative libraries.

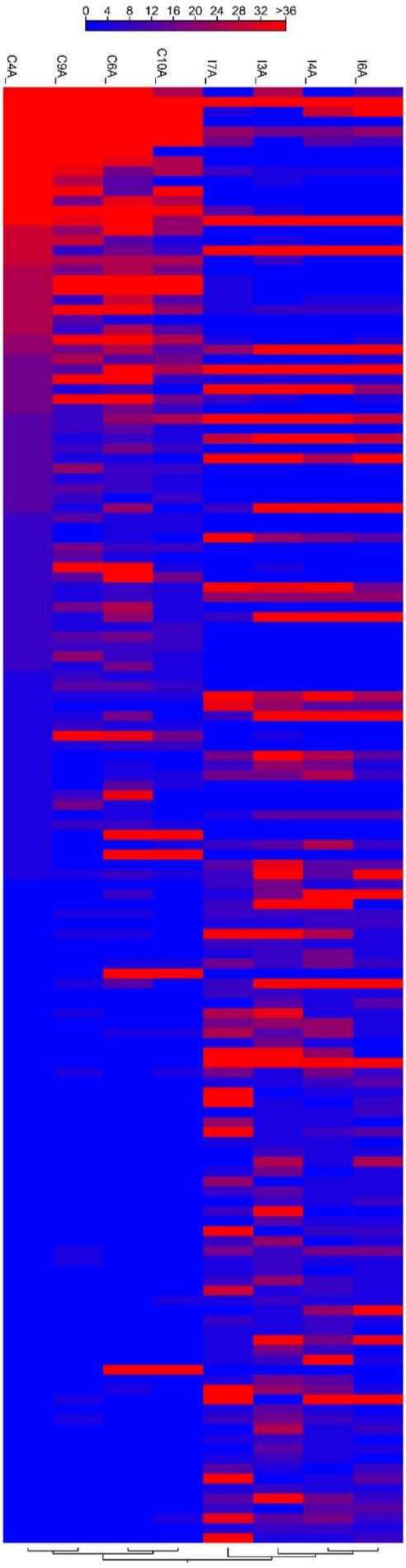


Figure 6: Heatmap showing the expression levels of the 147 contigs of interest calculated by RNA-seq analysis. Here, a scale from blue to red signifies the level of expression, with blue signifying no or undetectable expression, whilst increasing red colours signify increasingly levels of expression. Due to the level of resolution required for the generation of the heat map, not all differences between gene expression levels are represented. The four OOD-negative libraries (C4A, C9A, C6A and C10A) are on the left (numbered 1-4 in order), the four OOD-positive libraries (I7A, I3A, I4A, I6A) are on the right (numbered 5-8 in order).

In terms of fold-differences for the subset of 147 contigs, the majority of differentially expressed contigs (126) had fold-differences of less than 100 for OOD-positive versus OOD-negative libraries. Of the remaining contigs, 20 had fold-differences of between 100 and 500 (Figure 7). One contig returned a fold difference of over 1000 fold.

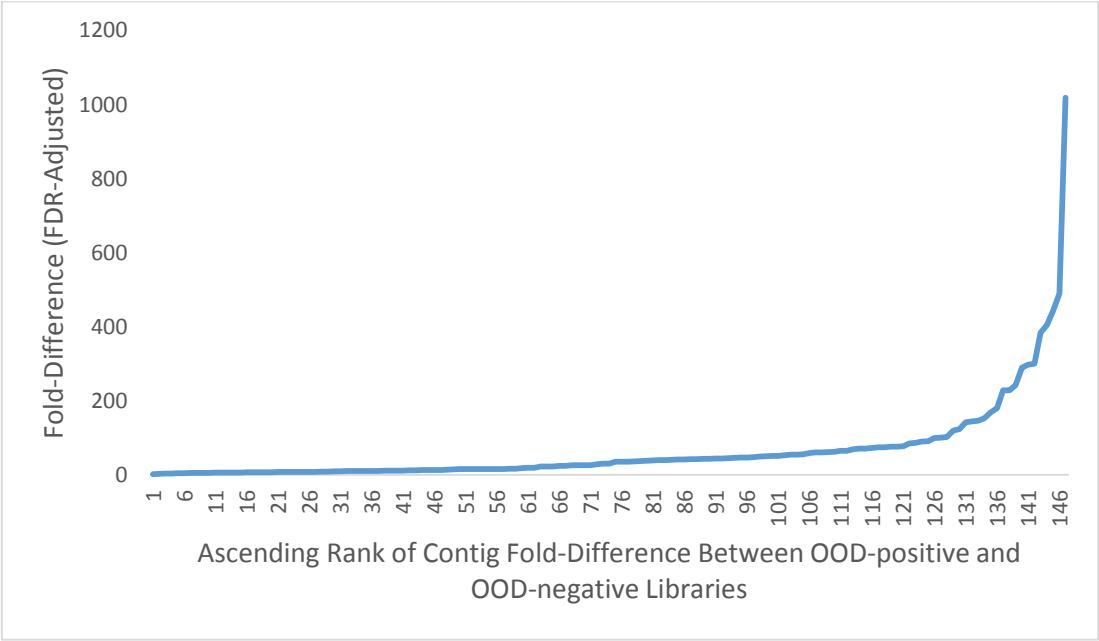


Figure 7: Fold-change (OOD-positive versus OOD-negative libraries) distribution for all 147 contigs, based on RNA-seq data, ordered from the smallest to largest fold difference.

3.4. Complete Transcriptome Annotation

Of the 81,236 contigs making up the whole transcriptome of the individuals tested in this study, 28,732 contigs returned significant Blastx hits (35% hit rate) in searches of the NCBI database. Of these hits, 17,796 (22%) could be assigned GO terms (Figure 8A). The subset of 147 contigs that were found to be differentially expressed between OOD-negative and OOD-positive libraries had a higher percentage of Blastx hits at 46% (80/147 contigs), although the percentage of contigs assigned GO terms was similar at 21% (31 contigs) in the subset (Figure 8B).

The three top species with which contigs shared significant homology were *C. gigas* (24,039 hits), *Lottia gigantea* (1,857 hits) and *Aplysia californica* (1,030 hits). All of these species have complete genome sequences [57, 94, 95]. Together, these three species of molluscs comprised 94% of the top hits. Of these species, *C. gigas* has the closest phylogenetic relationship to *P. maxima*, with both belonging to the family Bivalvia [96].

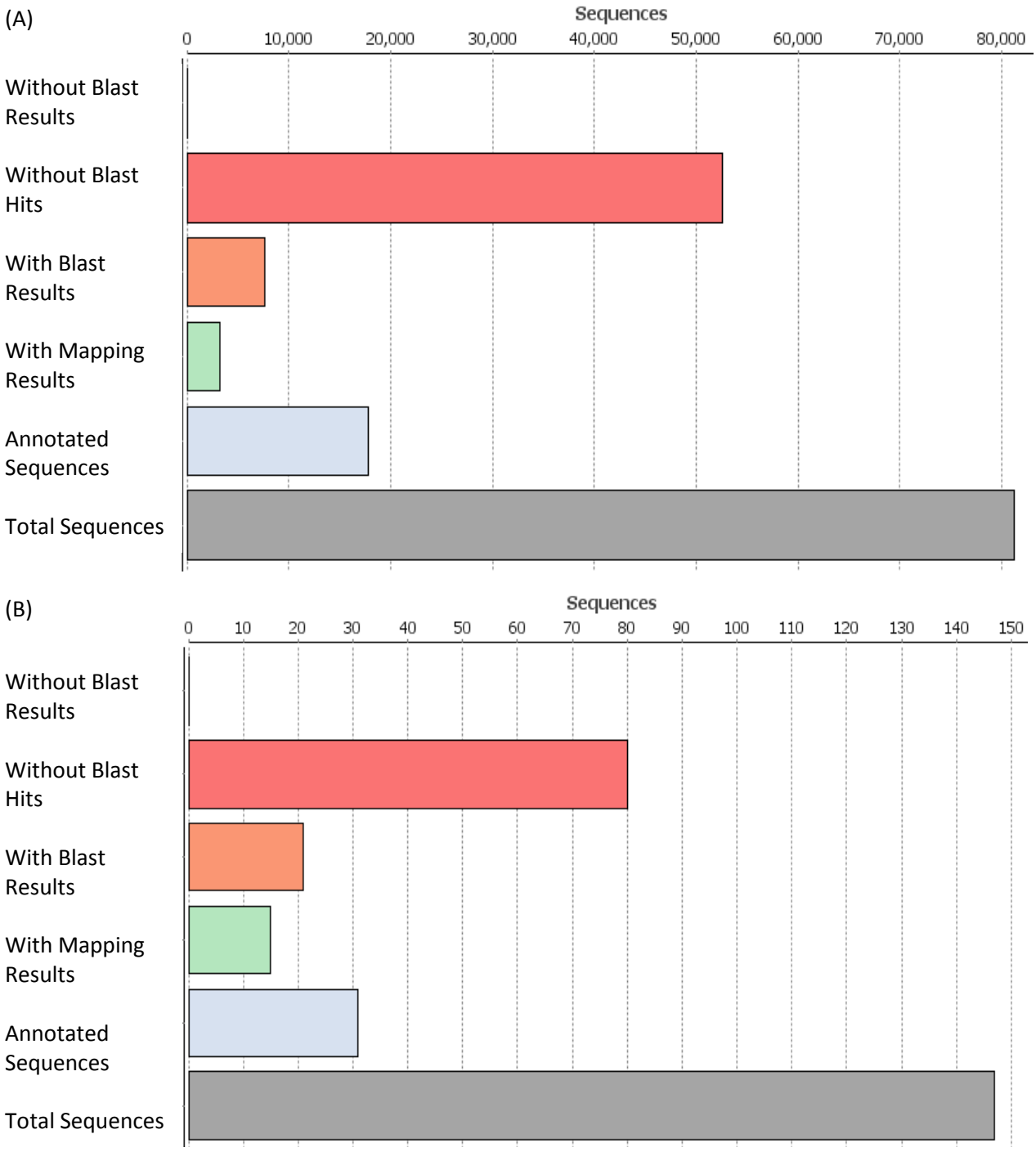


Figure 8: Blast and annotation results for (A) all 81,236 contigs that make up the transcriptome of the *P. maxima* libraries sequenced, and (B) the subset of 147 significantly differentially expressed contigs.

A breakdown of the GO term assignment reveals that of the GO terms obtained for all contigs in the whole transcriptome, 49% were biological processes, 33% were cellular components and 18% had molecular functions (Figure 9A). These are similar proportions to the subset of 147 contigs, of which 47% were biological processes, 31% were cellular components, and 22% had molecular functions (Figure 9B).

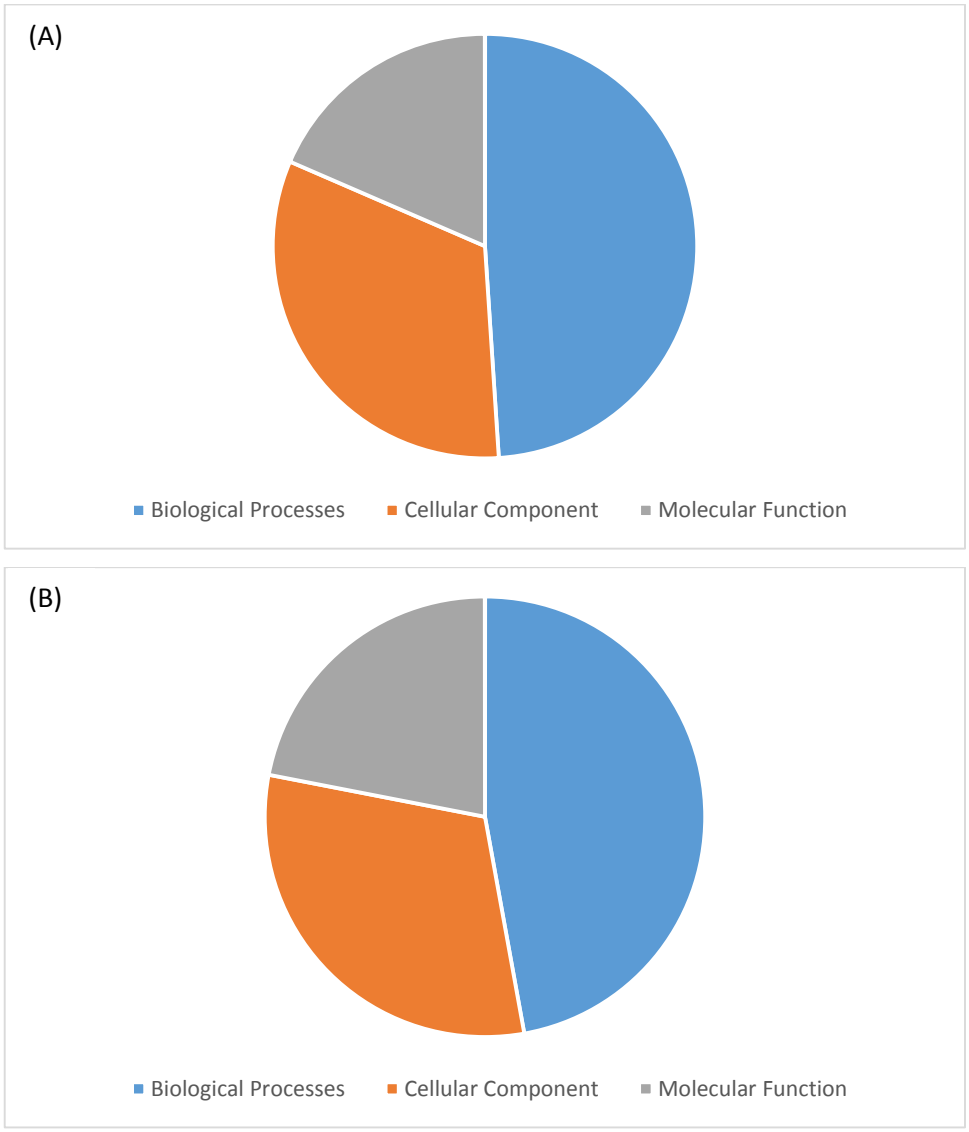
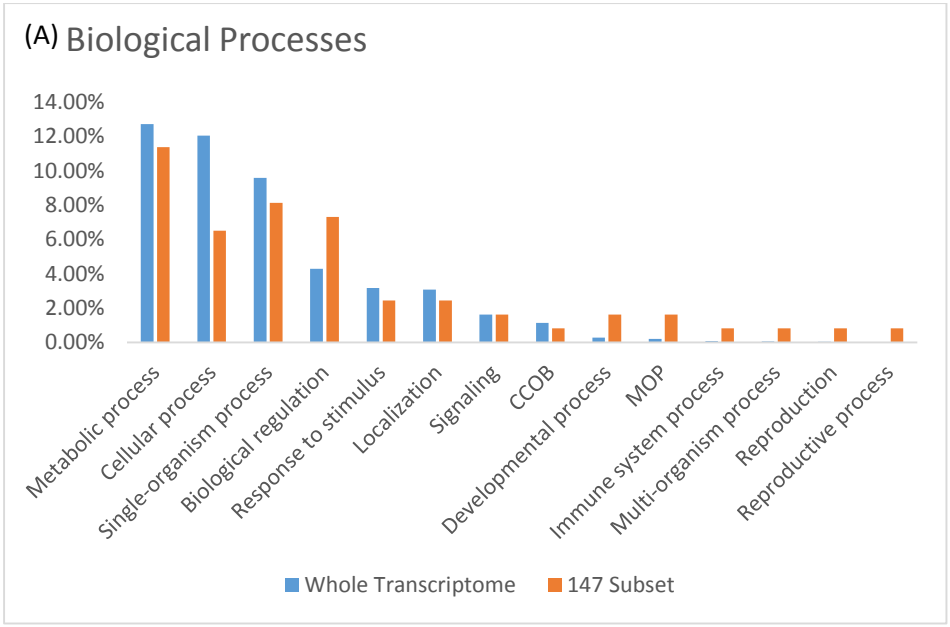


Figure 9: GO terms divided by broad ontologies for (A) the whole transcriptome, and (B) the 147 differentially expressed contigs identified through RNA-seq analysis.

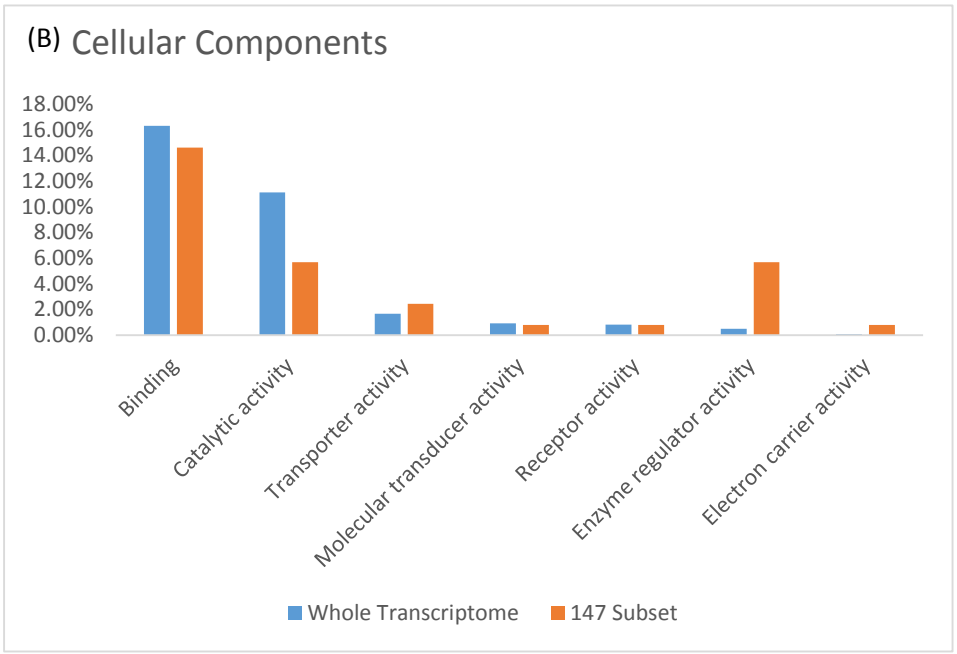
The three broad categories could be further broken down into more specific GO terms, with a total of 18 for biological functions, 14 for cellular components and ten for molecular functions. Of these, GO terms assigned to the subset of 147 significantly differentially expressed contigs only consisted of 14 biological functions (Figure 10A), seven cellular components (Figure 10B) and six molecular functions (Figure 10C).

Figure 10: Deep hierarchical division of GO terms shared by the whole transcriptome and the 147 significantly differentially expressed contigs, with a layered breakdown of the ontologies into specific pathways. The terms are divided into (A) biological processes, (B) cellular components, and (C) molecular functions. Abbreviations are as follows: CCOB = cellular component organization or biogenesis, MOP = multicellular organismal process, MC = macromolecular complex, MEL = membrane-enclosed lumen.

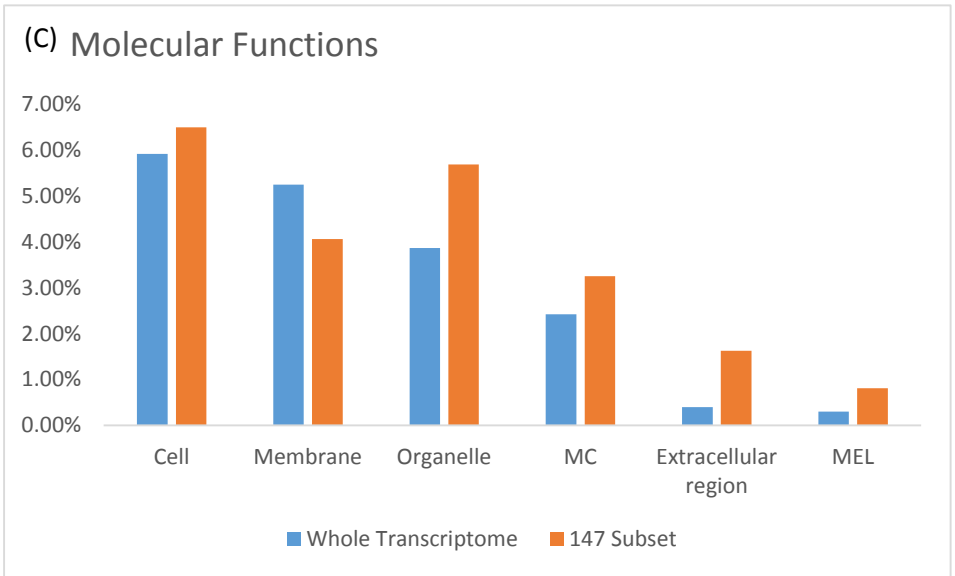
(Figure overleaf)



853



854



855

856

To determine whether for each GO term the proportion of assigned contigs were significantly different between the whole transcriptome and the subset of 147 contigs, a z-test statistical analysis was performed for each GO term. Table 6 displays all GO terms that were significantly different proportionally between the whole transcriptome and the subset of 147 contigs. The proportion of contigs that were assigned cellular process and/or catalytic activity GO terms in the subset of 147 significantly differentiated contigs was significantly less than that for the whole transcriptome, suggesting that response to OOD by *P. maxima* involved less cellular processes and catalytic activities than average. All other GO terms had a higher proportion of assigned contigs in the subset of 147 differentially expressed contigs. This suggests that there is a higher degree of involvement by genes with these GO terms in response to OOD than would be expected on average. A large proportion of these GO terms (8) were biological processes (Figure 4A), indicating that broadly, biological processes are most significantly affected during OOD.

Similar *de novo* transcriptome sequencing of both molluscs (*Crassostrea hongkongensis*, *Crassostrea angulata*, *Bathymodiolus azoricus*, and *Mizuhopecten yessoensis*) [97-100] and non-molluscs (*Apostichopus japonicus* and hybrid catfish) [101, 102] yielded between 20,056 and 144,746 (mean of 95,331) contigs, with 24-43% (mean of 33%) of the contigs having Blast hits, and 12-30% (mean of 18%) assigned GO terms. The RNA-seq analysis results of this study lies within these ranges from other similar studies, and is very similar to the mean values provided. This indicates that outputs from our transcriptome analysis are within a typical range.

Table 6: All GO terms that had a significantly different proportion of assigned contigs between the whole transcriptome and the subset of 147 contigs. There were a total of 66283 annotations for the whole transcriptome, and 123 annotations for the subset of significantly differentiated contigs. The z-test scores and corresponding p-values are shown, with all being <0.05. Comparisons that were not significant are not included in the table.

	Number of Contigs			
GO Annotation	**Whole Transcriptome**	**147 Subset**	**Z-test score**	**Corresponding p-value**
Cellular process	7985	8	-1.88744	0.02955
Biological regulation	2838	9	-1.66033	0.048424
Developmental process	191	2	-2.7537	0.002946
Multicellular organismal process	138	2	-3.42506	0.000307
Immune system process	44	1	-3.17905	0.000739
Multi-organism process	38	1	-3.4561	0.000274
Reproduction	23	1	-4.5371	2.85E-06
Reproductive process	15	1	-5.64262	8.37E-09
Catalytic activity	7378	7	-1.91725	0.027603
Enzyme regulator activity	325	7	-8.17029	1.54E-16
Electron carrier activity	52	1	-2.88214	0.001975
Extracellular region	264	2	-2.1537	0.015632

3.5. Expression validation

Six of the nine contigs of interest that underwent RT-qPCR analysis were found to be significantly differentially expressed based on the Wilcoxon rank sums test (p-value <0.05). The difference in expression levels for the reference contig (contig 1342) was not significant (p-value >0.05) under the same test, confirming that it was suitable as a reference.

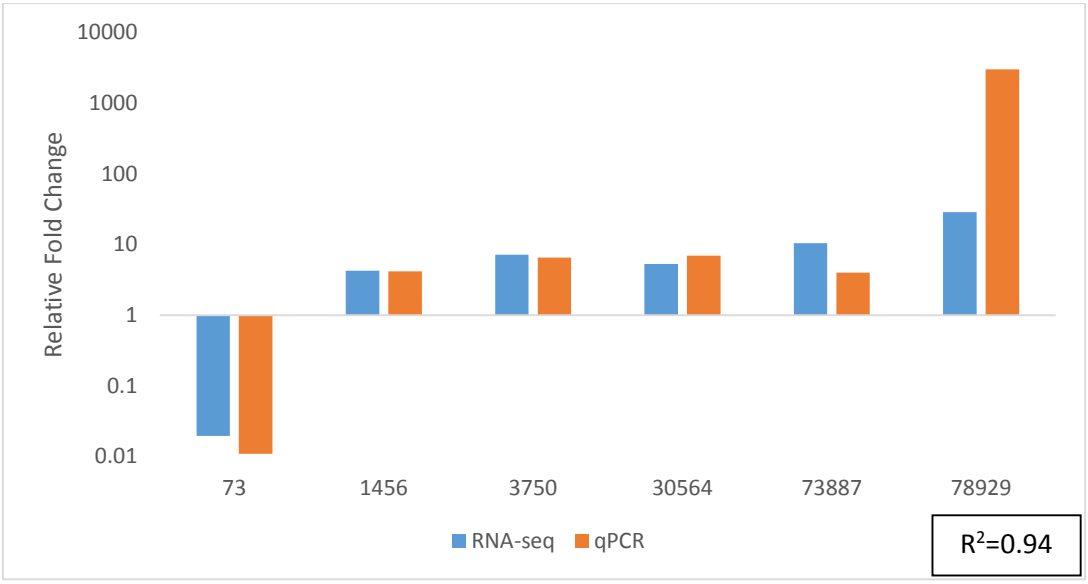


Figure 11: Relative fold differences (normalized) obtained from the RNA-seq analysis (blue) and RT-qPCR (orange) plotted against each other. The horizontal axis depicts the numerical identifiers of each of the six contigs found to be significantly differentially expressed through the RT-qPCR analysis.

The normalized relative expression values for the six contigs obtained through both the RNA-seq analysis and the RT-qPCR is shown in Figure 11, revealing a close correlation between the two analyses. The correlation coefficient (R^2 = 0.94) supports this, indicating that the RNA-seq analysis results are reliable as validated through RT-qPCR analysis.

3.6. Mollusc-specific genes that were differentially expressed

A total of 23 genes (discounting hypothetical proteins, uncharacterized, predicted or putative matches) matching entries in the molluscan database were found to be differentially expressed (Table 7). Of the 30 contigs that match these genes, 25 were up-regulated and five were down-regulated in OOD-positive relative to OOD-negative specimens. This is consistent with the finding by Anderson et al. [75] that in response to a stressor, there are generally more genes up-regulated than down-regulated.

Table 7: Automated annotated contigs matching known mollusc genes that were differentially expressed in the RNA-seq comparison of OOD-positive and OOD-negative specimens. The number of contigs with matches to known genes is shown, along with the contig ID for each match and top blast hit result information. Each gene is also assigned into a specific functional category.

Gene	Number of Contigs Matched	Contig ID	Blast Result				Functional Category
			Specific Match [Organism] / Accession Number	E-value	Identity	Fold Change (OOD-negative versus OOD-positive)	
Hydrolase	1	1790	hydrolase [*Helicobacter pullorum*] / WP_005023606.1	3.30E-35	79.45%	19.2	Unknown (bacteria)
chk1 checkpoint	1	5866	CHK1 checkpoint protein [*Echinococcus multilocularis*] / CDS35417.1	2.90E-20	76.56%	13.0	Cell cycle/apoptosis
carp-1	1	5848	clam ADP-ribosylating protein CARP-1 [*Meretrix lamarckii*] / BAF03560.1	1.60E-13	52.00%	-46.0	Immunity
Collagen alpha-2 chain	1	78949	Collagen alpha-2(VIII) chain [*Crassostrea gigas*] / EKC36664.1	4.40E-28	60.75%	61.6	Immunity/wound healing
Dispatched	1	73730	dispatched-like protein 1 [*Crassostrea gigas*] / EKC41449.1	5.80E-29	80.00%	10.4	Communication

DNA mismatch repair protein	1	13133	DNA mismatch repair protein Msh6 [*Crassostrea gigas*] / EKC25057.1	1.40E-24	76.50%	-46.6	DNA repair
Heat shock 70 kDa protein 8	1	990	Hsp70, partial [*Gallus gallus*], AAO44919.1	3.50E-48	94.30%	-7.7	Stress
Inter-alpha-trypsin inhibitor heavy chain h3	5	10635	Inter-alpha-trypsin inhibitor heavy chain H3 [*Crassostrea gigas*] / EKC29055.1	2.90E-61	63.75%	6.4	Protein regulation
		1456	Inter-alpha-trypsin inhibitor heavy chain H3 [*Crassostrea gigas*] / EKC23874.1	0.00E+00	63.60%	6.0	
		1458	Inter-alpha-trypsin inhibitor heavy chain H3 [*Crassostrea gigas*] / EKC23874.1	1.50E-29	78.00%	168.5	
		6990	Inter-alpha-trypsin inhibitor heavy chain H3 [*Crassostrea gigas*] / EKC27555.1	2.00E-61	48.20%	6.6	
		938	Inter-alpha-trypsin inhibitor heavy chain H3 [*Crassostrea gigas*] / EKC23874.1	1.10E-18	79.15%	7.4	

Inter-alpha-trypsin inhibitor heavy chain H4	2	10634	Inter-alpha-trypsin inhibitor heavy chain H4 [*Crassostrea gigas*] / EKC36390.1	3.00E-27	64%	5.3	Protein regulation
		12018	Inter-alpha-trypsin inhibitor heavy chain H4 [*Crassostrea gigas*] / EKC36390.1	3.00E-79	59.85%	5.5	
Macrophage mannose receptor	2	3750	Macrophage mannose receptor 1 [*Crassostrea gigas*] / EKC31065.1	1.90E-94	46.55%	10.1	Immunity
		7045	Macrophage mannose receptor 1, partial [*Chelonia mydas*] / EMP36986.1	5.50E-41	45.75%	7.5	
Metallothionein	1	16434	metallothionein [*Pinctada maxima*] / ACJ22893.1	2.40E-18	94.00%	50.8	Stress
Myosin heavy chain	1	73	Myosin heavy chain, striated muscle [*Crassostrea gigas*] / EKC37566.1	0.00E+00	82.15%	-36.0	Cytoskeleton

NADH dehydrogenase	1	5311	NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial [*Crassostrea gigas*] / EKC27484.1	1.40E-14	46.00%	-35.3	Metabolism
NADH dehydrogenase subunit 2	2	988	NADH dehydrogenase subunit 2 (mitochondrion) [*Pinctada maxima*] / YP_006883012.1	4.90E-53	85.00%	-442.6	Metabolism
		989	NADH dehydrogenase subunit 2 (mitochondrion) [*Pinctada maxima*] / YP_006883012.1	2.70E-54	85.00%	1018.6	
Permease	1	195	permease [*Lactobacillus crispatus*] / WP_013085634.1	2.10E-58	71.50%	29.9	Unknown (bacterial)
Prism shell protein 18	1	78932	Prism uncharacterized shell protein 18 like [*Pinctada fucata*] / BAM76253.1	7.50E-13	71.00%	383.7	Biomineralization
ALO17	1	79139	Protein ALO17 [*Crassostrea gigas*] / EKC32504.1	1.60E-20	57.67%	35.6	Protein regulation

Tar1p	1	1962	Tar1p [*Saccharomyces cerevisiae S288c*] / NP_690845.1	2.10E-34	64.35%	8.5	Unknown (fungal)
Sacsin	2	11369	Sacsin [*Crassostrea gigas*] / EKC26842.1	9.40E-26	47.25%	18.7	Stress
		77244	Sacsin [*Crassostrea gigas*] / EKC26842.1	6.30E-23	44.70%	9.3	
Short-chain collagen c4	1	14728	Short-chain collagen C4 [*Crassostrea gigas*] / EKC34668.1	1.20E-44	56.70%	5.1	Immunity/wound healing
Transforming growth factor-beta-induced protein ig-h3	1	6728	Transforming growth factor-beta-induced protein ig-h3 [*Crassostrea gigas*] / EKC23900.1	0.00E+00	51.95%	4.4	Immunity
Nacre shell protein 20	1	7760	nacre uncharacterized shell protein 20 [*Pinctada margaritifera*] / CCE46184.1	1.20E-17	95.00%	10.5	Biomineralization
von willebrand factor d and egf domain-containing protein	1	73887	von Willebrand factor D and EGF domain-containing protein [*Crassostrea gigas*] / EKC28789.1	1.70E-55	49.33%	14.8	Immunity/wound healing

937

938

Manual annotation against the entire NCBI non-redundant database identified an additional 41 differentially expressed contigs that matched genes/gene products that were not hypothetical, putative nor predicted. A large proportion of these were bacterial (approximately 66%), with seven contigs matching to spirochaetes or similar organisms. Three matched a specific spirochaete species, *Spirochaeta sp. L21-RPul-D2*. All spirochaete genes, however, were down-regulated in OOD-positive individuals, suggesting no link in the causative effect of the presence of the bacterium.

A further six contigs were found to match five molluscan genes after the manual annotation process (Table 8). These represented matches not identified by the automated process. The failure of automated searches to identify these matches was likely due to new additions to the non-redundant NCBI database that did not exist when the molluscan database used for the Blast processes was constructed. Addition of the manually annotated genes brought the total number of differentially expressed contigs with a match to molluscan genes/gene products to 48, with 33 distinct genes/gene products implicated.

Table 8: Manually annotated contigs identified as host genes that were differentially expressed in the RNA-seq comparison of OOD-positive and OOD-negative specimens. The number of contigs with matches to known genes is shown, along with the contig ID for each match and top blast hit result information. Each gene is also assigned into a specific functional category.

Gene	**Number of Contigs Matched**	**Contig ID**	**Blast Result**				**Functional Category**
			Specific Match [Organism] / Accession Number	**E-value**	**Identity**	**Fold Change (OOD-negative versus OOD-positive)**	
ATP synthase beta subunit	1	3889	ATP synthase beta subunit [*Pinctada fucata*]/ ABC86835.1	2.00E-72	98%	-7.4	Metabolism

Metalloproteinase inhibitor 3	1	4712	Metalloproteinase inhibitor 3 [*Crassostrea gigas*]/ EKC31955.1	1.00E-13	30%	11.3	Protein regulation
Sacsin	2	11368	Sacsin [*Crassostrea gigas*]/ EKC26842.1	7.00E-22	29%	-8.3	Stress
		36151	Sacsin [*Crassostrea gigas*]/ EKC26842.1	4.00E-15	35%	47.5	
Sarcoplasmic calcium-binding protein	1	4537	Sarcoplasmic calcium-binding protein [*Crassostrea gigas*]/ EKC29122.1	1.00E-27	53%	-25.9	Cytoskeleton
Uromodulin	1	30564	Uromodulin [*Crassostrea gigas*]/ EKC21727.1	8.00E-41	35%	7.5	Unknown

The characterized differentially expressed genes were placed into ten broad categories as listed in Tables 6 and 7. These were cell cycle (including apoptosis), immunity, wound healing, communication, DNA repair, stress response, protein regulation, cytoskeleton, metabolism, and biomineralization. Seven of these corresponds to intracellular processes defined by Anderson et al. [75]. Here, metabolism, cell cycle and communication related genes were less commonly identified than in Anderson et al. [75], whilst protein regulatory genes and specifically immunity related genes are found at higher rates (Figure 12). However, the methodology used to assign genes of interest to intracellular processes may be different, and could be the source of some of the discrepancy. With the three remaining broad categories fitted into the intracellular processes as described by Anderson et al. [75], the proportions could differ again.

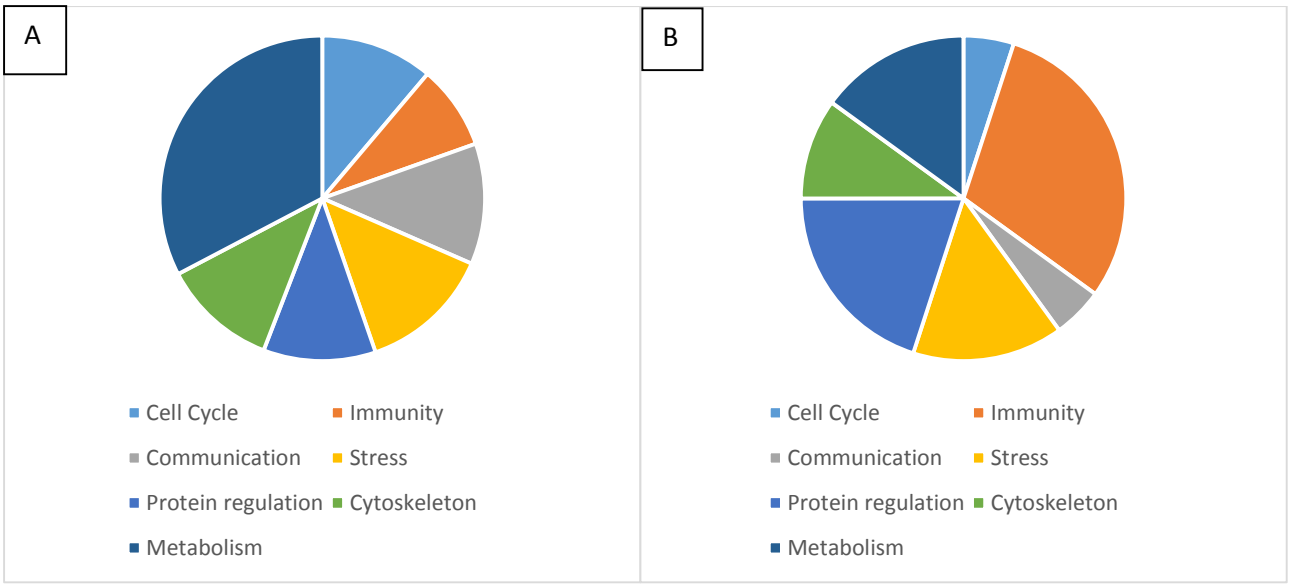


Figure 12: The proportion of genes belonging to each of the seven overlapping intracellular processes. (A) dictates the most commonly identified genes in proportion to each other by Anderson et al. [75], whilst (B) dictates the proportion of genes identified in the current study.

There were also a number of genes with unknown functions, two of which are bacterial in origin. Contig 1790 matches to a hydrolase, which is a large class of enzymes catalysing the hydrolysis of a chemical bond, such as those of mammalian hepoxilin and soluble epoxide [103]. As the match was to a bacterial (*Helicobacter pullorum*) hydrolase, the specific function of this gene in *P. maxima* cannot be identified. The second unknown bacterial match (contig 195) is a permease, which are membrane transport proteins of molecules such as L-tryptophan [104] oxyanions like molybdite [105]. Similarly to the hydrolase, the function in *P. maxima* is highly uncertain, as the match is to a bacteria (*Lactobacillus crispatus*). Tar1p (contig 1962) is another non-mollusc gene, which in *Saccharomyces cerevisiae* plays a role in respiratory regulation [106, 107], and has been shown to also play a similar role in at least one other yeast species (Kluyveromyces lactis) [108]. However, the role of Tar1p has yet to be identified in non-yeast species, and so it is unclear what role Tar1p plays in *P. maxima*. Uromodulin (contig 30564) was the final gene with an uncharacterized molluscan function.

The remaining genes could be grouped into two broad categories of responses – a general stress response, involving changes in communication, stress response, protein regulation, cytoskeleton, and metabolism; and a concurrent wound healing response, involving changes in the cell cycle, immunity, wound healing, DNA repair and biomineralization.

3.6.1. General stress response

Genes/gene products involved with the cytoskeleton are those that controls the structure of cells, and maintenance of such structure. Both sarcoplasmic calcium-binding protein and myosin heavy chain are cytoskeletal proteins. In molluscs, sarcoplasmic calcium-binding proteins are regulatory proteins in the muscle [109], whilst the myosin heavy chain is involved in muscle contraction [110]. In this case, both are down-regulated, which is consistent with the symptoms of OOD: loss of epithelial cells and mantle retraction. Myosin heavy chains are also known to fuse with anaplastic lymphoma kinase (ALK), which can also be fused to ALO17 [111, 112]. Contig 79139 was matched to ALO17, and was up-regulated in response to OOD. Whether there is a relationship between ALO17 and myosin heavy chains in this case must be further investigated. Von willebrand factor D and EGF domain-containing protein could also play a role in relation to the symptomology of OOD. The EGF (epidermal growth factor) links it to the outer layers of tissues and cells, whilst the Von willebrand factor D plays a role in blood clotting in humans [113]. However, its role in molluscs is unknown.

There were a number of gene products involved in general stress responses which are activated ubiquitously in response to varied stressors [75]. These generic stress response genes include molecular chaperones (heat shock proteins). Molecular chaperones are commonly accompanied by co-chaperones that form complexes together, such as that of the heat shock protein 70 (hsp70) DnaK and its co-chaperone, Dna-J [114]. In the current study, both a member of the hsp70 family (heat shock 70 kDa protein 8, HSPA8) and an hsp70 co-chaperone (sacsin) were found to be differentially expressed. HSPA8 antagonizes apoptosis [115-117], and therefore its reduction among OOD-positive individuals could indicate that apoptosis is uninhibited during OOD. Four contigs are found to match the hsp70 co-chaperone, sacsin, with three showing up-regulation in OOD-positive libraries and one being down-regulated. All four matches have the same accession number (EKC26842.1), suggesting that all four contigs could potentially be one single read. However, the fact that they are expressed at different levels from one another suggests that these could be different sacsin isoforms, and the one down-regulated sacsin isoform (contig 11368) could be the co-chaperone to the hsp70 found in this study.

Metallothionein also plays a role in general stress responses, specifically as a potential defence against metal toxicity [118] by trafficking metals (mainly zinc) and sequestering them [119]. Increased concentrations of vanadium [120] or cadmium [121] for example, has been reported to result in increased expression of metallothionein by mussels (*Mytilus sp.*). Metallothionein has therefore been investigated as a potential biomarker of metal toxicity. However it is known that there are other factors that can trigger expression of metallothioneins as well, including generalized stress [122].

In humans and other mammals, inter-alpha-trypsin inhibitor heavy chains inhibits proteases and binds to chemicals such as hyaluronan [123], and is also up-regulated in response to infections [124, 125]. Two forms (chains h3 and h4) were identified in the current study, with all matched contigs up-regulated in response to OOD. The number of matches found in this study and the ubiquitous role inter-alpha-trypsin inhibitor heavy chains supports the notion of a general response to infection in OOD-impacted oysters.

Genes involved in metabolism control the production of energy, and the catabolism/anabolism of chemicals within an organism. A major proportion of energy production takes place in the mitochondria. NADH dehydrogenase is located within the inner mitochondrial membrane, and is responsible for catalysing the dehydrogenation of NADH into NAD^+. The behaviour of NADH dehydrogenases are enigmatic in terms of their potential role in OOD. Whilst one NADH dehydrogenase subunit 2 (contig 988) had a 442.63 negative fold-change (down-regulation) in OOD-positive individuals, the other (contig 999) saw a 1018.64 positive fold-change (up-regulation). This was the biggest fold-change amongst the significantly differentially expressed contigs. Whilst it is known that NADH dehydrogenase subunits can show different patterns of expression based on different time points during responses to stress [64, 65], there is high sequence similarity between the two contigs, and actual expression values for the two contigs suggested that they were allelic and unique to the collection sites. Whilst the average expression value of contig 988 for OOD-positive oysters was zero, there was a much higher expression value for OOD-negative oysters. This was the reverse for contig 999, suggesting that contig 988 was exclusive to OOD-negative oysters, whilst contig 989 was exclusive to OOD-positive oysters. Also in terms of metabolism, ATP synthase beta subunit (contig 3889) was found to be down-regulated [126, 127].

Finally, biomineralization, especially in pearl oysters, may be in response to contamination or other abiotic and biotic factors [72, 73]. Nacre and prisms are two types of calcium carbonate deposits, which are controlled by two very different pathways [71], although they can form a highly desirable composite. Both nacre and prism related proteins were up-regulated in response to OOD, suggesting a more orchestrated stress response.

3.6.2. Immune response and wound healing

In general, genes involved in immune responses and wound healing act to protect the organism from harm and when not possible, repair damage done to the organism. A total of ten genes were identified in these categories during the current study.

There were a number of genes found to play a role in responding to infections. Metalloproteinase inhibitor 3 was found to be up-regulated in response to OOD. Metalloproteinase inhibitors are inhibitors of matrix metalloproteinase, with roles in the degradation and remodelling of tissues [128]. In molluscs, the gene is found to play a role in wound healing and defense mechanisms, specifically in response to shell damage and bacterial infections [129, 130]. Although uncovered by Zhang et al. for *C. gigas* [57], no specific function was described for uromodulin. In mammals such as humans, uromodulin is secreted by kidneys and seems to play a role in immune responses against urinary infections [131, 132]. It is unclear if uromodulin plays a similar role in *P. maxima*. Macrophage mannose receptor-1 is involved in inflammation and wound healing, specifically in times of infection [133]. It acts as a receptor for pathogen associated molecular patterns. Its differential expression in OOD-positive oysters may suggest that there is an infectious agent involved in OOD.

Chk1 is a regulator of the cell cycle and associated DNA damage in mammals such as humans [134, 135] and mice [136, 137], amongst other animals such as frogs [138], and even fungi [139, 140]. This suggests that the role of Chk1 is ubiquitous, including in *P. maxima*. Similarly, carp-1 is suggested to play a role in the cell cycle, specifically up-regulation of the gene induced apoptosis [141, 142]. It (contig 5848) was down-regulated in OOD-positive oysters, and therefore suggests a suppression of apoptosis. DNA mismatch repair proteins are a large class of proteins, including proteins such as the heterodimeric MutS homolog (MSH) [143]. Here, Msh6 (contig 13133) was found to be

down-regulated, which has been shown elsewhere to be the direct effect of the presence of a stressor [144, 145]. It has been found that Msh6 induces apoptosis [146, 147], and therefore the suppression of Msh6 in OOD-positive oysters may inhibit apoptosis.

Transforming growth factor-beta-induced protein ig-h3 (TGFBI), represented by contig 6728, plays a role in inducing collagen expression [148], of which two were identified to be differentially expressed: collagen alpha-2 chain (contig 78949) and short-chain collagen c4 (contig 14728). Collagen is known to play a role in immune responses by binding to pathogens, and antagonizing apoptosis by binding to apoptosis-inducing cells [149]. TGFBI is also known to play other roles, such as in inhibiting cell attachment [150]. The up-regulation of all three of these genes in OOD-positive oysters strongly suggests an immune-response, and once again exhibits a clear indication that apoptosis is antagonized in response to OOD.

Dispatched (contig 73730) is a protein that plays a role in membranal molecular trafficking, which along with a hedgehog gene is involved in cellular growth and development [151-153]. However, without the differential expression of the hedgehog gene in this case, it is unsure that dispatched plays this same role in response to OOD.

3.7. OOD response profile and comparison to the viral infection response profile

As determined in section *1.6.*, the viral infection response profile is very specific. Table 1 lists five proteins/protein families that are differentially expressed in response to a viral infection, which are distinct from those involved in response to other stressors. These proteins have been definitively implicated in response to viral infections in both oysters [79, 80, 154] and non-oysters [81, 82].

A total of 27 genes/gene products make up the OOD response profile. However, no gene/gene product determined through the current study matches with any of the protein/protein families that makes up the typical viral infection response profile. This excludes the viral stressor as a potential cause of OOD.

On the other hand, a large proportion of the OOD response profile matches to other genes/gene products identified in Table 1 that are associated with stress, wound healing and immune responses. NADH dehydrogenase was commonly found to be differentially

expressed, along with molecular chaperones (heat shock proteins and their co-chaperones). Many proteins that play a role in apoptosis inhibition were also identified. All of these changes are characteristic of responses to non-viral infections. Comparison of the broad intracellular processes identified in this study found that there was no significant difference (p-value=0.282, >0.05) between our data and the meta-analysis of Anderson et al. [75]. This indicated that there was a concerted, but generalized response to stress in OOD-affected oysters. This is in agreement with the observation that a large portion of the OOD response profile (21 genes out of 27) plays a role in stress responses (13 genes) or immune responses/wound healing (8 genes).

4. Conclusion

The transcriptomic analysis of the host response of the pearl oyster *P. maxima* to OOD has validated the method as a viable investigative tool for the future. Through RNA-seq analysis, a comparison of expression levels of OOD-positive and OOD-negative specimens was accomplished, and a set of differentially expressed genes obtained and collated into a response profile of 28 genes. A comparison of this response profile to the typical viral infection response profile of molluscs yielded no match. In contrast, there were clear indications that an alternative stressor, yielding a more 'general' response to stress, is involved in OOD. In particular, this generalised stress response involved the differential expression of genes involved in wound healing and immune responses.

This study has therefore been able to narrow the focus of further research, which will be able to concentrate on non-viral stressors. With more stressor response profiles built in the future, studies such as this will become more precise and powerful. Even now, this study has already established the feasibility of generating the data necessary for comparison.

5. References

1. Carraschi SP, Barbuio R, Ikefuti CV, Florencio T, da Cruz C, Ranzani-Paiva MJT. Effectiveness of therapeutic agents in disease treatment in *Piaractus mesopotamicus*. Aquaculture. 2014 431:124-8.
2. Pikitch EK, Doukakis P, Lauck L, Chakrabarty P, Erickson DL. Status, trends and management of sturgeon and paddlefish fisheries. Fish and Fisheries. 2005 6:233-65.
3. Fletcher W, Friedman K, Weir V, McCrea J, Clark R. Pearl Oyster Fishery. Western Australia: Department of Fisheries; 2006.
4. Shumway SE. Shellfish Aquaculture and the Environment: Wiley-Blackwell; 2011.
5. Bostock J, McAndrew B, Richards R, Jauncey K, Telfer T, Lorenzen K, et al. Aquaculture: global status and trends. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010 365:2897-912.
6. Flegel TW, Alday-Sanz V. The crisis in Asian shrimp aquaculture: current status and future needs. Journal of Applied Ichthyology. 1998 14:269-73.
7. Jones JB, Creeper J. Diseases of pearl oysters and other molluscs: A Western Australian perspective. Journal of Shellfish Research. 2006 25:233-8.
8. Petton B, Pernet F, Robert R, Boudry P. Temperature influence on pathogen transmission and subsequent mortalities in juvenile Pacific oysters *Crassostrea gigas*. Aquaculture Environment Interactions. 2013 3:257-73.
9. Taylor JJ, Rose RA, Southgate PC. Effects of stocking density on the growth and survival of juvenile silver-lip pearl oysters (*Pinctada maxima*, Jameson) in suspended and bottom culture. Journal of Shellfish Research. 1997 16:569-72.
10. Kuchel RP, O'Connor WA, Raftos DA. Environmental stress and disease in pearl oysters, focusing on the Akoya pearl oyster (*Pinctada fucata* Gould 1850). Reviews in Aquaculture. 2011 3:138-54.
11. Pass DA, Dybdahl R, Mannion MM. Investigations into the causes of mortality of the pearl oyster, *Pinctada maxima* (Jameson), in Western Australia. Aquaculture. 1987 65:149-69.
12. Lacoste A, Jalabert F, Malham S, Cueff A, Gelebart F, Cordevant C, et al. A *Vibrio splendidus* strain is associated with summer mortality of juvenile oysters *Crassostrea gigas* in the Bay of Morlaix (North Brittany, France). Diseases of Aquatic Organisms. 2001 46:139-45.
13. Perkins FO. Shell disease in the gold lip pearl oyster, *Pinctada maxima* and the Eastern oyster, *Crassostrea virginica*. Aquatic Living Resources. 1996 9:159-68.
14. Lockwood BL, Sanders JG, Somero GN. Transcriptomic responses to heat-stress reveal the molecular basis for the success of invasive mussels. Integrative and Comparative Biology. 2010 50:E103-E.
15. Humphrey JD, Barton MA. An unsuccessful attempt to experimentally transmit "oyster oedema disease syndrome" to pearl oyster spat. . Technote (Darwin) 2009 No 125. 2009:7.
16. Jones JB, Crockford M, Creeper J, Stephens F. Histopathology of oedema in pearl oysters *Pinctada maxima*. Diseases of Aquatic Organisms. 2010 91:67-73.
17. Evans JJ, Klesius PH, Shoemaker CA. Therapeutic and prophylactic immunization against *Streptococcus iniae* infection in hybrid striped bass (*Morone chrysops* x *Morone saxatilis*). Aquaculture Research. 2006 37:742-50.
18. Matsubara D, Tanaka M, Soumyou Y, Hirakawa K, Doi RJ, Nakai T. Therapeutic effects of antimicrobial compounds against bacillary necrosis of larval pacific oyster. Fish Pathology. 2002 37:183-8.
19. Yamashita N, Yasojima M, Nakada N, Miyajima K, Komori K, Suzuki Y, et al. Effects of antibacterial agents, levofloxacin and clarithromycin, on aquatic organisms. Water Science and Technology. 2006 53:65-72.
20. Cabello FC, Godfrey HP, Tomova A, Ivanova L, Dolz H, Millanao A, et al. Antimicrobial use in aquaculture re-examined: its relevance to antimicrobial resistance and to animal and human health. Environmental Microbiology. 2013 15:1917-42.
21. Heuer OE, Kruse H, Grave K, Collignon P, Karunasagar I, Angulo FJ. Human Health Consequences of Use of Antimicrobial Agents in Aquaculture. Clinical Infectious Diseases. 2009 49:1248-53.
22. Tamminen M, Karkman A, Lohmus A, Muziasari WI, Takasu H, Wada S, et al. Tetracycline resistance genes persist at aquaculture farms in the absence of selection pressure. Environmental Science & Technology. 2011 45:386-91.

23. Naciri-Graven Y, Martin AG, Baud JP, Renault T, Gerard A. Selecting the flat oyster *Ostrea edulis* (L) for survival when infected with the parasite *Bonamia ostreae*. Journal of Experimental Marine Biology and Ecology. 1998 224:91-107.
24. Renault T. Controlling viral diseases in aquaculture: new developments. New Technologies in Aquaculture: Improving Production Efficiency, Quality and Environmental Management. 2009:244-66.
25. Lynch SA, Flannery G, Hugh-Jones T, Hugh-Jones D, Culloty SC. Thirty-year history of Irish (Rossmore) *Ostrea edulis* selectively bred for disease resistance to *Bonamia ostreae*. Diseases of Aquatic Organisms. 2014 110:113-21.
26. Evans F, Matson S, Brake J, Langdon C. The effects of inbreeding on performance traits of adult Pacific oysters (*Crassostrea gigas*). Aquaculture. 2004 230:89-98.
27. Jouaux A, Lafont M, Blin JL, Houssin M, Mathieu M, Lelong C. Physiological change under OsHV-1 contamination in Pacific oyster *Crassostrea gigas* through massive mortality events on fields. Bmc Genomics. 2013 14:14.
28. Leite RB, Milan M, Coppe A, Bortoluzzi S, dos Anjos A, Reinhardt R, et al. mRNA-Seq and microarray development for the Grooved carpet shell clam, *Ruditapes decussatus*: a functional approach to unravel host -parasite interaction. Bmc Genomics. 2013 14:17.
29. Wang SL, Peatman E, Liu H, Bushek D, Ford SE, Kucuktas H, et al. Microarray analysis of gene expression in eastern oyster (*Crassostrea virginica*) reveals a novel combination of antimicrobial and oxidative stress host responses after dermo (*Perkinsus marinus*) challenge. Fish & Shellfish Immunology. 2010 29:921-9.
30. Milan M, Ferraresso S, Ciofi C, Chelazzi G, Carrer C, Ferrari G, et al. Exploring the effects of seasonality and chemical pollution on the hepatopancreas transcriptome of the Manila clam. Molecular Ecology. 2013 22:2157-72.
31. Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. Nature Reviews Microbiology. 2005 3:281-94.
32. Chaney ML, Gracey AY. Mass mortality in Pacific oysters is associated with a specific gene expression signature. Molecular Ecology. 2011 20:2942-54.
33. Muramoto Y, Shoemaker JE, Le MQ, Itoh Y, Tamura D, Sakai-Tagawa Y, et al. Disease severity is associated with differential gene expression at the early and late phases of infection in nonhuman primates infected with different H5N1 highly pathogenic Avian Influenza Viruses. Journal of Virology. 2014 88:8981-97.
34. Yang YH, Li MJ, Chen XJ, Wang PF, Wang FQ, Lin WX, et al. De novo characterization of the *Rehmannia glutinosa* leaf transcriptome and analysis of gene expression associated with replanting disease. Molecular Breeding. 2014 34:905-15.
35. Bolon-Canedo V, Sanchez-Marono N, Alonso-Betanzos A, Benitez JM, Herrera F. A review of microarray datasets and applied feature selection methods. Information Sciences. 2014 282:111-35.
36. Karunatilaka KS, Rueda D. Post-transcriptional modifications modulate conformational dynamics in human U2-U6 snRNA complex. Rna-a Publication of the Rna Society. 2014 20:16-23.
37. Briolat V, Jouneau L, Carvalho R, Palha N, Langevin C, Herbomel P, et al. Contrasted innate responses to two viruses in zebrafish: insights into the ancestral repertoire of vertebrate IFN- stimulated genes. Journal of Immunology. 2014 192:4328-41.
38. Gaska M, Kusmider M, Solich J, Faron-Gorecka A, Krawczyk MJ, Kulakowski K, et al. Analysis of region-specific changes in gene expression upon treatment with citalopram and desipramine reveals temporal dynamics in response to antidepressant drugs at the transcriptome level. Psychopharmacology. 2012 223:281-97.
39. McNamara LE, Dalby MJ, Tsimbouri MP. The use of microarrays and fluorescence *in situ* hybridization for the study of mechanotransduction from topography. In: Piel M, Thery M, editors. Micropatterning in Cell Biology, Pt A. San Diego: Elsevier Academic Press Inc; 2014, p. 293-309.
40. Westermann AJ, Gorski SA, Vogel J. Dual RNA-seq of pathogen and host. Nature Reviews Microbiology. 2012 10:618-30.
41. Okoniewski M, Miller C. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. BMC Bioinformatics. 2006 7:276.

42. Royce TE, Rozowsky JS, Gerstein MB. Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification. Nucleic Acids Research. 2007 35:e99.
43. Harbers M, Carninci P. Tag-based approaches for transcriptome research and genome annotation. Nature Methods. 2005 2:495-502.
44. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science. 1995 270:484-7.
45. Kodzius R, Kojima M, Nishiyori H, Nakamura M, Fukuda S, Tagami M, et al. CAGE: cap analysis of gene expression. Nature Methods. 2006 3:211-22.
46. Nakamura M, Carninci P. Cap analysis gene expression: CAGE. Tanpakushitsu Kakusan Koso. 2004 49:2688-93.
47. Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, et al. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proceedings of the National Academy of Sciences. 2003 100:15776-81.
48. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, et al. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotech. 2000 18:630-4.
49. Peiffer J, Kaushik S, Sakai H, Arteaga-Vazquez M, Sanchez-Leon N, Ghazal H, et al. A spatial dissection of the *Arabidopsis* floral transcriptome by MPSS. BMC Plant Biology. 2008 8:43.
50. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics. 2009 10:57-63.
51. Chen HH, Zha JM, Liang XF, Bu JH, Wang M, Wang ZJ. Sequencing and *de novo* assembly of the Asian clam (*Corbicula fluminea*) transcriptome using the Illumina GAIIx method. Plos One. 2013 8:12.
52. Ghiselli F, Milani L, Chang PL, Hedgecock D, Davis JP, Nuzhdin SV, et al. *De novo* assembly of the Manila clam *Ruditapes philippinarum* transcriptome provides new insights into expression bias, mitochondrial doubly uniparental inheritance and sex determination. Molecular Biology and Evolution. 2012 29:771-86.
53. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, et al. Global identification of human transcribed sequences with genome tiling arrays. Science. 2004 306:2242-6.
54. David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, et al. A high-resolution map of transcription in the yeast genome. Proceedings of the National Academy of Sciences. 2006 103:5320-5.
55. Hegedus Z, Zakrzewska A, Agoston VC, Ordas A, Racz P, Mink M, et al. Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. Molecular Immunology. 2009 46:2918-30.
56. Ordas A, Hegedus Z, Henkel CV, Stockhammer OW, Butler D, Jansen HJ, et al. Deep sequencing of the innate immune transcriptomic response of zebrafish embryos to *Salmonella* infection. Fish & Shellfish Immunology. 2011 31:716-24.
57. Zhang GF, Fang XD, Guo XM, Li L, Luo RB, Xu F, et al. The oyster genome reveals stress adaptation and complexity of shell formation. Nature. 2012 490:49-54.
58. Fu X, Sun Y, Wang J, Xing Q, Zou J, Li R, et al. Sequencing-based gene network analysis provides a core set of gene resource for understanding thermal adaptation in Zhikong scallop *Chlamys farreri*. Molecular Ecology Resources. 2014 14:184-98.
59. McDowell IC, Nikapitiya C, Aguiar D, Lane CE, Istrail S, Gomez-Chiarri M. Transcriptome of American oysters, *Crassostrea virginica*, in response to bacterial challenge: insights into potential mechanisms of disease resistance. Plos One. 2014 9:14.
60. Tomanek L. Environmental proteomics of the mussel *Mytilus*: implications for tolerance to stress and change in limits of biogeographic ranges in response to climate change. Integrative and Comparative Biology. 2012 52:648-64.
61. Li Q, Zhao X, Kong L, Yu H. Transcriptomic response to stress in marine bivalves. ISJ-Invertebr Surviv J. 2013 10:84-93.
62. Meng XL, Liu M, Jiang KY, Wang BJ, Tian X, Sun SJ, et al. *De novo* characterization of Japanese Scallop *Mizuhopecten yessoensis* transcriptome and analysis of its gene expression following Cadmium exposure. Plos One. 2013 8:10.
63. Meng XL, Tian X, Liu M, Nie GX, Jiang KY, Wang BJ, et al. The transcriptomic response to copper exposure by the gill tissue of Japanese scallops (*Mizuhopecten yessoensis*) using deep-sequencing technology. Fish & Shellfish Immunology. 2014 38:287-93.

64. Silva-Aciares F, Zapata M, Tournois J, Moraga D, Riquelme C. Identification of genes expressed in juvenile *Haliotis rufescens* in response to different copper concentrations in the north of Chile under controlled conditions. Marine Pollution Bulletin. 2011 62:2671-80.
65. Zapata M, Tanguy A, David E, Moraga D, Riquelme C. Transcriptomic response of *Argopecten purpuratus* post-larvae to copper exposure under experimental conditions. Gene. 2009 442:37-46.
66. Huning A, Melzner F, Thomsen J, Gutowska MA, Kramer L, Frickenhaus S, et al. Impacts of seawater acidification on mantle gene expression patterns of the Baltic Sea Blue Mussel: implications for shell formation and energy metabolism. Marine Biology. 2013 160:1845-61.
67. Negri A, Oliveri C, Sforzini S, Mignione F, Viarengo A, Banni M. Transcriptional response of the mussel *Mytilus galloprovincialis* (Lam.) following exposure to heat stress and copper. Plos One. 2013 8:11.
68. Zhao XL, Yu H, Kong LF, Li Q. Transcriptomic Responses to Salinity Stress in the Pacific Oyster *Crassostrea gigas*. Plos One. 2012 7:8.
69. Segarra A, Mauduit F, Faury N, Trancart S, Degremont L, Tourbiez D, et al. Dual transcriptomics of virus-host interactions: comparing two Pacific oyster families presenting contrasted susceptibility to ostreid herpesvirus 1. Bmc Genomics. 2014 15:13.
70. Mohamed B, Hajer A, Susanna S, Caterina O, Flavio M, Hamadi B, et al. Transcriptomic responses to heat stress and nickel in the mussel *Mytilus galloprovincialis*. Aquatic Toxicology. 2014 148:104-12.
71. Marie B, Joubert C, Tayale A, Zanella-Cleon I, Belliard C, Piquemal D, et al. Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell. Proceedings of the National Academy of Sciences of the United States of America. 2012 109:20986-91.
72. McGinty EL, Zenger KR, Jones DB, Jerry DR. Transcriptome analysis of biomineralisation-related genes within the pearl sac: Host and donor oyster contribution. Marine Genomics. 2012 5:27-33.
73. Zhao XX, Wang QH, Jiao Y, Huang RL, Deng YW, Wang H, et al. Identification of Genes Potentially Related to Biomineralization and Immunity by Transcriptome Analysis of Pearl Sac in Pearl Oyster *Pinctada martensii*. Marine Biotechnology. 2012 14:730-9.
74. Duckett CS, Nava VE, Gedrich RW, Clem RJ, VanDongen JL, Gilfillan MC, et al. A conserved family of cellular genes related to the baculovirus iap gene and encoding apoptosis inhibitors. Embo Journal. 1996 15:2685-94.
75. Anderson K, Taylor DA, Thompson EL, Melwani AR, Nair SV, Raftos DA. Meta-analysis of studies using suppression subtractive hybridization and microarrays to investigate the effects of environmental stress on gene transcription in oysters. Plos One. 2015 10:15.
76. Gargano LM, Moser JM, Speck SH. Role for MyD88 signaling in murine gammaherpesvirus 68 latency. Journal of Virology. 2008 82:3853-63.
77. Hallen LC, Burki Y, Ebeling M, Broger C, Siegrist F, Oroszlan-Szovik K, et al. Antiproliferative activity of the human IFN-alpha- inducible protein IFI44. Journal of Interferon and Cytokine Research. 2007 27:675-80.
78. Uchiyama R, Chassaing B, Zhang BY, Gewirtz AT. MyD88-mediated TLR signaling protects against acute rotavirus infection while inflammasome cytokines direct Ab response. Innate Immunity. 2015 21:416-28.
79. Green TJ, Montagnani C. Poly I:C induces a protective antiviral immune response in the Pacific oyster (*Crassostrea gigas*) against subsequent challenge with Ostreid herpesvirus (OsHV-1 mu var). Fish & Shellfish Immunology. 2013 35:382-8.
80. Green TJ, Barnes AC. Inhibitor of REL/NF-B-K is regulated in Sydney rock oysters in response to specific double-stranded RNA and *Vibrio alginolyticus*, but the major immune anti-oxidants EcSOD and Prx6 are non-inducible. Fish & Shellfish Immunology. 2009 27:260-5.
81. He NH, Qin QW, Xu X. Differential profile of genes expressed in hemocytes of White Spot Syndrome Virus-resistant shrimp (*Penaeus japonicus*) by combining suppression subtractive hybridization and differential hybridization. Antiviral Research. 2005 66:39-45.
82. Robalino J, Browdy CL, Prior S, Metz A, Parnell P, Gross P, et al. Induction of antiviral immunity by double-stranded RNA in a marine invertebrate. Journal of Virology. 2004 78:10442-8.
83. Rosa RD, Barracco MA. Shrimp interferon is rather a portion of the mitochondrial F0-ATP synthase than a true alpha-interferon. Molecular Immunology. 2008 45:3490-3.

84. Rosa RD, de Lorgeril J, Tailliez P, Bruno R, Piquemal D, Bachere E. A hemocyte gene expression signature correlated with predictive capacity of oysters to survive *Vibrio* infections. Bmc Genomics. 2012 13:12.
85. Queiroga FR, Marques-Santos LF, Hegaret H, Soudant P, Farias ND, Schlindwein AD, et al. Immunological responses of the mangrove oysters *Crassostrea gasar* naturally infected by *Perkinsus sp.* in the Mamanguape Estuary, Paraiba state (Northeastern, Brazil). Fish & Shellfish Immunology. 2013 35:319-27.
86. Roberts SB, Sunila I, Wikfors GH. Immune response and mechanical stress susceptibility in diseased oysters, *Crassostrea virginica*. Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology. 2012 182:41-8.
87. Rio DC, Ares M, Jr., Hannon GJ, Nilsen TW. Purification of RNA using TRIzol (TRI reagent). Cold Spring Harbor protocols. 2010 2010:pdb.prot5439.
88. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 30:2114-20.
89. Jones P, Binns D, Chang HY, Fraser M, Li WZ, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014 30:1236-40.
90. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Research. 2007 35:W43-W6.
91. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods. 2001 25:402-8.
92. Liu SQ, Rich JO, Anderson A. Antibacterial activity of a cell wall hydrolase from *Lactobacillus paracasei* NRRL B-50314 produced by recombinant *Bacillus megaterium*. Journal of Industrial Microbiology & Biotechnology. 2015 42:229-35.
93. Li W. Volcano plots in analyzing differential expressions with mRNA microarrays. Journal of bioinformatics and computational biology. 2012 10.
94. Knudsen B, Kohn AB, Nahir B, McFadden CS, Moroz LL. Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: Conservation of the gene order in Euthyneura. Molecular Phylogenetics and Evolution. 2006 38:459-69.
95. Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, et al. Insights into bilaterian evolution from three spiralian genomes. Nature. 2013 493:526-31.
96. Temkin I. Molecular phylogeny of pearl oysters and their relatives (Mollusca, Bivalvia, Pterioidea). BMC Evolutionary Biology. 2010 10:342.
97. Bettencourt R, Pinheiro M, Egas C, Gomes P, Afonso M, Shank T, et al. High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*. Bmc Genomics. 2010 11:16.
98. Qin J, Huang ZX, Chen J, Zou Q, You WW, Ke CH. Sequencing and *de novo* Analysis of *Crassostrea angulata* (Fujian Oyster) from 8 Different developing phases using 454 GSFlx. Plos One. 2012 7:8.
99. Wang S, Hou R, Bao ZM, Du HX, He Y, Su HL, et al. Transcriptome sequencing of Zhikong Scallop (*Chlamys farreri*) and comparative transcriptomic Yesso Scallop (*Patinopecten yessoensis*). Plos One. 2013 8:10.
100. Zhao XL, Yu H, Kong LF, Liu SK, Li Q. Comparative transcriptome analysis of two oysters, *Crassostrea gigas* and *Crassostrea hongkongensis* provides insights into adaptation to hypo-osmotic conditions. Plos One. 2014 9:11.
101. Liu SK, Wang XL, Sun FY, Zhang JR, Feng JB, Liu H, et al. RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation in response to heat stress in catfish. Physiological Genomics. 2013 45:462-76.
102. Zhou ZC, Dong Y, Sun HJ, Yang AF, Chen Z, Gao S, et al. Transcriptome sequencing of sea cucumber (*Apostichopus japonicus*) and the identification of gene-associated markers. Molecular Ecology Resources. 2014 14:127-38.
103. Cronin A, Decker M, Arand M. Mammalian soluble epoxide hydrolase is identical to liver hepoxilin hydrolase. Journal of Lipid Research. 2011 52:712-9.
104. Rosenfeld H, Feigelson P. Product induction in *Pseudomonas acidovorans* of a permease system which transports L-tryptophan. Journal of bacteriology. 1969 97:705-14.

105. Gisin J, Muller A, Pfander Y, Leimkuhler S, Narberhaus F, Masepohl B. A *Rhodobacter capsulatus* member of a universal permease family imports molybdate and other oxyanions. Journal of Bacteriology. 2010 192:5943-52.
106. Coelho PSR, Bryan AC, Kumar A, Shadel GS, Snyder M. A novel mitochondrial protein, Tar1p, is encoded on the antisense strand of the nuclear 25S rDNA. Genes & Development. 2002 16:2755-60.
107. Bonawitz ND, Chatenay-Lapointe M, Wearn CM, Shadel GS. Expression of the rDNA-encoded mitochondrial protein Tar1p is stringently controlled and responds differentially to mitochondrial respiratory demand and dysfunction. Current Genetics. 2008 54:83-94.
108. Galopier A, Hermann-Le Denmat S. Mitochondria of the yeasts *Saccharomyces cerevisiae* and *Kluyveromyces lactis* contain nuclear rDNA-encoded proteins. Plos One. 2011 6:11.
109. Hermann A, Cox JA. Sarcoplasmic calcium-binding protein. Comp Biochem Physiol B Biochem Mol Biol. 1995 111:337-45.
110. SzentGyorgyi AG. Regulation of contraction by calcium binding myosins. Biophysical Chemistry. 1996 59:357-63.
111. Cools J, Wlodarska I, Somers R, Mentens N, Pedeutour F, Maes B, et al. Identification of novel fusion partners of ALK, the anaplastic lymphoma kinase, in anaplastic large-cell lymphoma and inflammatory myofibroblastic tumor. Genes Chromosomes & Cancer. 2002 34:354-62.
112. Moritake H, Shimonodan H, Marutsuka K, Kamimura S, Kojima H, Nunoi H. C-MYC rearrangement may induce an aggressive phenotype in anaplastic lymphoma kinase positive anaplastic large cell lymphoma: Identification of a novel fusion gene ALO17/C-MYC. American Journal of Hematology. 2011 86:75-8.
113. Wagner DD. Cell biology of von Willebrand factor. Annual Review of Cell Biology. 1990 6:217-42.
114. Suh WC, Lu CZ, Gross CA. Structural features required for the interaction of the Hsp70 molecular chaperone DnaK with its cochaperone DnaJ. Journal of Biological Chemistry. 1999 274:30534-9.
115. Beere HM, Wolf BB, Cain K, Mosser DD, Mahboubi A, Kuwana T, et al. Heat-shock protein 70 inhibits apoptosis by preventing recruitment of procaspase-9 to the Apaf-1 apoptosome. Nature Cell Biology. 2000 2:469.
116. Beere HM, Green DR. Stress management – heat shock protein-70 and the regulation of apoptosis. Trends in Cell Biology. 2001 11:6-10.
117. Ravagnan L, Gurbuxani S, Susin SA, Maisse C, Daugas E, Zamzami N, et al. Heat-shock protein 70 antagonizes apoptosis-inducing factor. Nature Cell Biology. 2001 3:839.
118. Davis SR, Cousins RJ. Metallothionein expression in animals: a physiological perspective on function. Journal of Nutrition. 2000 130:1085-8.
119. Kikuchi S, Inohara M, Okamura I, Oshima Y, Takeuchi T, Miura T, et al. Purification and characterization of a metallothionein-like, zinc-binding protein of scallops, *Patinopecten yessoensis*. Bioscience Biotechnology and Biochemistry. 1992 56:1434-8.
120. Amiard JC, Journel R, Bacheley H. Influence of field and experimental exposure of mussels (*Mytilus sp.*) to nickel and vanadium on metallothionein concentration. Comparative Biochemistry and Physiology C-Toxicology & Pharmacology. 2008 147:378-85.
121. Bebianno MJ, Langston WJ. Cadmium induction of metallothionein synthesis in *Mytilus galloprovincialis*. Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology. 1992 103:79-85.
122. Cosson RP. Bivalve metallothionein as a biomarker of aquatic ecosystem pollution by trace metals: limits and perspectives. Cellular and Molecular Biology. 2000 46:295-309.
123. Odum L, Andersen CY, Jessen TE. Characterization of the coupling activity for the binding of inter-alpha-trypsin inhibitor to hyaluronan in human and bovine follicular fluid. Reproduction. 2002 124:249-57.
124. Liony C, Sesbue R, Manchon ND, Bercoff E, Delzant G, Martin JP, et al. Inter-alpha-trypsin inhibitor and its derivatives in inflammatory syndromes. Presse Medicale. 1991 20:203-6.
125. Pineiro M, Andres M, Iturralde M, Carmona S, Hirvonen J, Pyorala S, et al. ITIH4 (inter-alpha-trypsin inhibitor heavy chain 4) is a new acute-phase protein isolated from cattle during experimental infection. Infection and Immunity. 2004 72:3777-82.
126. Chi SL, Pizzo SV. Cell surface F1Fo ATP synthase: a new paradigm? Annals of Medicine. 2006 38:429-38.

127. Thomas D, Bron P, Weimann T, Dautant A, Giraud MF, Paumard P, et al. Supramolecular organization of the yeast F1Fo-ATP synthase. Biology of the Cell. 2008 100:591-601.
128. Woessner JF. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. The FASEB Journal. 1991 5:2145-54.
129. Montagnani C, Le Roux F, Berthe F, Escoubas JM. Cg-TIMP, an inducible tissue inhibitor of metalloproteinase from the Pacific oyster *Crassostrea gigas* with a potential role in wound healing and defense mechanisms. Febs Letters. 2001 500:64-70.
130. Wang Q, Bao YB, Huo LH, Gu HL, Lin ZH. A novel tissue inhibitor of metalloproteinase in blood clam *Tegillarca granosa*: Molecular cloning, tissue distribution and expression analysis. Fish & Shellfish Immunology. 2012 33:645-51.
131. Iorember FM, Vehaskari VM. Uromodulin: old friend with new roles in health and disease. Pediatric Nephrology. 2014 29:1151-8.
132. Prajczer S, Heidenreich U, Pfaller W, Kotanko P, Lhotta K, Jennings P. Evidence for a role of uromodulin in chronic kidney disease progression. Nephrology Dialysis Transplantation. 2010 25:1896-903.
133. Lee SJ, Evers S, Roeder D, Parlow AF, Risteli J, Risteli L, et al. Mannose receptor-mediated regulation of serum glycoprotein homeostasis. Science. 2002 295:1898-901.
134. Latif C, den Elzen NR, O'Connell MJ. DNA damage checkpoint maintenance through sustained Chk1 activity. Journal of Cell Science. 2004 117:3489-98.
135. Pabla N, Bhatt K, Dong Z. Checkpoint kinase 1 (Chk1)-short is a splice variant and endogenous inhibitor of Chk1 that regulates cell cycle and DNA damage checkpoints. Proceedings of the National Academy of Sciences of the United States of America. 2012 109:197-202.
136. Fishler T, Li YY, Wang RH, Kim HS, Sengupta K, Vassilopoulos A, et al. Genetic instability and mammary tumor formation in mice carrying mammary-specific disruption of Chk1 and p53. Oncogene. 2010 29:4007-17.
137. Takai H, Tominaga K, Motoyama N, Minamishima YA, Nagahama H, Tsukiyama T, et al. Aberrant cell cycle checkpoint function and early embryonic death in Chk1(-/-) mice. Genes & Development. 2000 14:1439-47.
138. Nakajo N, Oe T, Uto K, Sagata N. Involvement of Chk1 kinase in prophase I arrest of *Xenopus oocytes*. Developmental Biology. 1999 207:432-44.
139. den Elzen NR, O'Connell MJ. Recovery from DNA damage checkpoint arrest by PP1-mediated inhibition of Chk1. Embo Journal. 2004 23:908-18.
140. Paparatto D, Fletcher D, Piwowar K, Baldino K, Morel C, Dunaway S. The *Schizosaccharomyces pombe* checkpoint kinases Chk1 and Cds1 are important for cell survival in response to cisplatin. Plos One. 2009 4:8.
141. Jiang Y, Puliyappadamba VT, Zhang L, Wu W, Wali A, Yaffe MB, et al. A novel mechanism of cell growth regulation by Cell Cycle and Apoptosis Regulatory Protein (CARP)-1. J Mol Signal. 2010 5:7.
142. Nakano T, Matsushima-Hibiya Y, Yamamoto M, Enomoto S, Matsumoto Y, Totsuka Y, et al. Purification and molecular cloning of a DNA ADP-ribosylating protein, CARP-1, from the edible clam *Meretrix lamarckii*. Proceedings of the National Academy of Sciences of the United States of America. 2006 103:13652-7.
143. Bowers J, Tran PT, Joshi A, Liskay RM, Alani E. MSH-MLH complexes formed at a DNA mismatch are disrupted by the PCNA sliding clamp. Journal of Molecular Biology. 2001 306:957-68.
144. Hargreaves VV, Shell SS, Mazur DJ, Hess MT, Kolodner RD. Interaction between the Msh2 and Msh6 Nucleotide-binding Sites in the *Saccharomyces cerevisiae* Msh2-Msh6 Complex. Journal of Biological Chemistry. 2010 285:9301-10.
145. Hsu T, Tsai HT, Huang KM, Luan MC, Hsieh CR. Sublethal levels of cadmium down-regulate the gene expression of DNA mismatch recognition protein MutS homolog 6 (MSH6) in zebrafish (*Danio rerio*) embryos. Chemosphere. 2010 81:748-54.
146. Narine KAD, Keuling AM, Gombos R, Tron VA, Andrew SE, Young LC. Defining the DNA mismatch repair-dependent apoptotic pathway in primary cells: Evidence for p53-independence and involvement of centrosomal caspase 2. DNA Repair. 2010 9:161-8.

147. Young LC, Peters AC, Maeda T, Edelmann W, Kucherlapati R, Andrew SE, et al. DNA mismatch repair protein msh6 is required for optimal levels of ultraviolet-B-induced apoptosis in primary mouse fibroblasts. Journal of Investigative Dermatology. 2003 121:876-80.
148. Varga J, Jimenez SA, Rosenbloom J. Transforming growth-factor-beta induces concerted and persistent elevation in levels of type-I and type-III collagens and fibronectin messenger-RNA and corresponding peptide expression by cultured fibroblasts. Clinical Research. 1987 35:A802-A.
149. Fraser DA, Tenner AJ. Directing an appropriate immune response: The role of defense collagens and other soluble pattern recognition molecules. Current Drug Targets. 2008 9:113-22.
150. Skonier J, Bennett K, Rothwell V, Kosowski S, Plowman G, Wallace P, et al. beta ig-h3: a transforming growth factor-beta-responsive gene encoding a secreted protein that inhibits cell attachment in vitro and suppresses the growth of CHO cells in nude mice. DNA Cell Biol. 1994 13:571-84.
151. Callejo A, Bilioni A, Mollica E, Gorfinkiel N, Andres G, Ibanez C, et al. Dispatched mediates Hedgehog basolateral release to form the long-range morphogenetic gradient in the *Drosophila* wing disk epithelium. Proceedings of the National Academy of Sciences of the United States of America. 2011 108:12591-8.
152. Kawakami T, Kawcak T, Li YJ, Zhang WH, Hu YM, Chuang PT. Mouse dispatched mutants fall to distribute hedgehog proteins and are defective in hedgehog signaling. Development. 2002 129:5753-65.
153. Nakano Y, Kim HR, Kawakami A, Roy S, Schier AF, Ingham PW. Inactivation of dispatched 1 by the chameleon mutation disrupts hedgehog signalling in the zebrafish embryo. Developmental Biology. 2004 269:381-92.
154. Anderson K, Elizur A. Hepatic reference gene selection in adult and juvenile female Atlantic salmon at normal and elevated temperatures. BMC Research Notes. 2012 5:21.

6. Appendix

6.1. Journal format instructions

The chosen journal is **Fish and Shellfish Immunology**, with instructions printed at http://www.elsevier.com/journals/fish-and-shellfish-immunology/1050-4648/guide-for-authors.

Your Paper Your Way

We now differentiate between the requirements for new and revised submissions. You may choose to submit your manuscript as a single Word or PDF file to be used in the refereeing process. Only when your paper is at the revision stage, will you be requested to put your paper in to a 'correct format' for acceptance and provide the items required for the publication of your article.
To find out more, please visit the Preparation section below.

Introduction

Types of Papers

Research papers

Review articles

Short Communications: Where the subject matter does not warrant a full paper, a Short Communication is more likely to be acceptable for publication. Communications should be restricted in length to no more than the equivalent of two printed pages of Fish and Shellfish Immunology, that total length to include references and normally no more than one illustration and one short table. No abstract will be required, nor should a communication necessarily have subheadings or be subdivided as are full papers, but an introductory sentence or sentences must make its purport clear. In other respects submitted manuscripts should comply with the instructions given above. A Short Communication may be concerned with any subject within the scope of Fish and Shellfish Immunology but should be confined to a single point or issue of progress, and to such as an unusual occurrence, an interesting observation or a topical and timely finding.

Short Sequence Reports:Where the reporting of a novel fish or shellfish sequence with limited functional analysis does not warrant a full paper, a Short Sequence Report is more likely to be acceptable for publication. Communications should be restricted in length as for Short Communications (see above). No abstract or subheadings are required but an introductory sentence or sentences must make it clear why the sequence is novel and the approach used to obtain the gene. In other respects the submitted manuscripts should comply with the instructions given above, and be within the subject scope of the journal. Accepted reports will undergo rapid publication.

Submitting Short Sequence Reports: For submitting Short Sequence Reports, please choose "Short Communication" as the article type and in the "Sections/Categories", please choose "Short Sequence Reports" from the drop down list.

Before You Begin

Ethics in publishing

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Human and animal rights

If the work involves the use of animal or human subjects, the author should ensure that the work described has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans http://www.wma.net/en/30publications/10policies/b3/index.html; EU Directive 2010/63/EU for animal experiments http://ec.europa.eu/environment/chemicals/lab_animals/legislation_en.htm; Uniform Requirements for manuscripts submitted to Biomedical journals http://www.icmje.org. Authors should include a statement in the manuscript that informed consent was obtained for experimentation with human subjects. The privacy rights of human subjects must always be observed.

Conflict of interest

All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work. See also http://www.elsevier.com/conflictsofinterest. Further information and an example of a Conflict of Interest form can be found at: http://help.elsevier.com/app/answers/detail/a_id/286/p/7923.

Submission declaration

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/sharingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere including electronically in the same form, in English or in any other language, without the written consent of the copyright-holder.

Authorship

All authors should have made substantial contributions to all of the following: (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data, (2) drafting the article or revising it critically for important intellectual content, (3) final approval of the version to be submitted.

Changes to authorship

This policy concerns the addition, deletion, or rearrangement of author names in the authorship of accepted manuscripts:
Before the accepted manuscript is published in an online issue: Requests to add or remove an author, or to rearrange the author names, must be sent to the Journal Manager from the corresponding author of the accepted manuscript and must include: (a) the reason the name should be added or removed, or the author names rearranged and (b) written confirmation (e-mail, fax, letter) from all authors that they agree with the addition, removal or rearrangement. In the case of

addition or removal of authors, this includes confirmation from the author being added or removed. Requests that are not sent by the corresponding author will be forwarded by the Journal Manager to the corresponding author, who must follow the procedure as described above. Note that: (1) Journal Managers will inform the Journal Editors of any such requests and (2) publication of the accepted manuscript in an online issue is suspended until authorship has been agreed.
After the accepted manuscript is published in an online issue: Any requests to add, delete, or rearrange author names in an article published in an online issue will follow the same policies as noted above and result in a corrigendum.

Article transfer service

This journal is part of our Article Transfer Service. This means that if the Editor feels your article is more suitable in one of our other participating journals, then you may be asked to consider transferring the article to one of those. If you agree, your article will be transferred automatically on your behalf with no need to reformat. Please note that your article will be reviewed again by the new journal. More information about this can be found here:
http://www.elsevier.com/authors/article-transfer-service.

Copyright

Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (for more information on this and copyright, see http://www.elsevier.com/copyright). An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.

Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations (please consult http://www.elsevier.com/permissions). If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases: please consult http://www.elsevier.com/permissions.

For open access articles: Upon acceptance of an article, authors will be asked to complete an 'Exclusive License Agreement' (for more information see http://www.elsevier.com/OAauthoragreement). Permitted third party reuse of open access articles is determined by the author's choice of user license (see http://www.elsevier.com/openaccesslicenses).

Author rights
As an author you (or your employer or institution) have certain rights to reuse your work. For more information see http://www.elsevier.com/copyright.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated.

Funding body agreements and policies

Elsevier has established a number of agreements with funding bodies which allow authors to comply with their funder's open access policies. Some authors may also be reimbursed for associated publication fees. To learn more about existing agreements please visit http://www.elsevier.com/fundingbodies.

Open access

This journal offers authors a choice in publishing their research:

Open access
• Articles are freely available to both subscribers and the wider public with permitted reuse
• An open access publication fee is payable by authors or on their behalf e.g. by their research funder or institution
Subscription
• Articles are made available to subscribers as well as developing countries and patient groups through our universal access programs (http://www.elsevier.com/access).
• No open access publication fee payable by authors.

Regardless of how you choose to publish your article, the journal will apply the same peer review criteria and acceptance standards.

For open access articles, permitted third party (re)use is defined by the following Creative Commons user licenses:

Creative Commons Attribution (CC BY)
Lets others distribute and copy the article, create extracts, abstracts, and other revised versions, adaptations or derivative works of or from an article (such as a translation), include in a collective work (such as an anthology), text or data mine the article, even for commercial purposes, as long as they credit the author(s), do not represent the author as endorsing their adaptation of the article, and do not modify the article in such a way as to damage the author's honor or reputation.

Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)
For non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

The open access publication fee for this journal is **USD 3000**, excluding taxes. Learn more about Elsevier's pricing policy: http://www.elsevier.com/openaccesspricing.

Language (usage and editing services)

Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who feel their English language manuscript may require editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English may wish to use the English Language Editing service available from Elsevier's WebShop (http://webshop.elsevier.com/languageediting/) or visit our customer support site (http://support.elsevier.com) for more information.

Submission

Our online submission system guides you stepwise through the process of entering your article details and uploading your files. The system converts your article files to a single PDF file used in

the peer-review process. Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the Editor's decision and requests for revision, is sent by e-mail.

For submission of your article to Fish and Shellfish Immunology please go to http://ees.elsevier.com/fsim.

Submit your article
Please submit your article via http://ess.elsevier.com/yfsim

Referees

Please submit, with the manuscript, the names, addresses and e-mail addresses of 5 potential referees. Note that the editor retains the sole right to decide whether or not the suggested reviewers are used.

Preparation

NEW SUBMISSIONS

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. The system automatically converts your files to a single PDF file, which is used in the peer-review process.
As part of the Your Paper Your Way service, you may choose to submit your manuscript as a single file to be used in the refereeing process. This can be a PDF file or a Word document, in any format or lay-out that can be used by referees to evaluate your manuscript. It should contain high enough quality figures for refereeing. If you prefer to do so, you may still provide all or some of the source files at the initial submission. Please note that individual figure files larger than 10 MB must be uploaded separately.

References

There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/book title, chapter title/article title, year of publication, volume number/book chapter and the pagination must be present. Use of DOI is highly encouraged. The reference style used by the journal will be applied to the accepted article by Elsevier at the proof stage. Note that missing data will be highlighted at proof stage for the author to correct.

Formatting requirements

There are no strict formatting requirements but all manuscripts must contain the essential elements needed to convey your manuscript, for example Abstract, Keywords, Introduction, Materials and Methods, Results, Conclusions, Artwork and Tables with Captions.
If your article includes any Videos and/or other Supplementary material, this should be included in your initial submission for peer review purposes.
Divide the article into clearly defined sections.

Please ensure your paper has consecutive line numbering - this is an essential peer review requirement.

Figures and tables embedded in text
Please ensure the figures and the tables included in the single file are placed next to the relevant text in the manuscript, rather than at the bottom or the top of the file.

REVISED SUBMISSIONS

Use of word processing software
Regardless of the file format of the original submission, at revision you must provide us with an editable file of the entire article. Keep the layout of the text as simple as possible. Most formatting codes will be removed and replaced on processing the article. The electronic text should be prepared in a way very similar to that of conventional manuscripts (see also the Guide to Publishing with Elsevier: http://www.elsevier.com/guidepublication). See also the section on Electronic artwork.
To avoid unnecessary errors you are strongly advised to use the 'spell-check' and 'grammar-check' functions of your word processor.

Subdivision - numbered sections
Divide your article into clearly defined and numbered sections. Subsections should be numbered 1.1 (then 1.1.1, 1.1.2, ...), 1.2, etc. (the abstract is not included in section numbering). Use this numbering also for internal cross-referencing: do not just refer to 'the text'. Any subsection may be given a brief heading. Each heading should appear on its own separate line.

Introduction
State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Material and methods
Provide sufficient detail to allow the work to be reproduced. Methods already published should be indicated by a reference: only relevant modifications should be described.

Results
Results should be clear and concise.

Discussion
This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Conclusions
The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

Essential title page information

• **Title.** Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.
• **Author names and affiliations.** Please clearly indicate the given name(s) and family name(s) of each author and check that all names are accurately spelled. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower-case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.

• **Corresponding author.** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. **Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author.**
• **Present/permanent address.** If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Abstract

A concise and factual abstract is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

Graphical abstract

Although a graphical abstract is optional, its use is encouraged as it draws more attention to the online article. The graphical abstract should summarize the contents of the article in a concise, pictorial form designed to capture the attention of a wide readership. Graphical abstracts should be submitted as a separate file in the online submission system. Image size: Please provide an image with a minimum of 531 × 1328 pixels (h × w) or proportionally more. The image should be readable at a size of 5 × 13 cm using a regular screen resolution of 96 dpi. Preferred file types: TIFF, EPS, PDF or MS Office files. See http://www.elsevier.com/graphicalabstracts for examples. Authors can make use of Elsevier's Illustration and Enhancement service to ensure the best presentation of their images and in accordance with all technical requirements: Illustration Service.

Highlights

Highlights are mandatory for this journal. They consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate editable file in the online submission system. Please use 'Highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). See http://www.elsevier.com/highlights for examples.

Immediately after the abstract, provide 5-10 keywords, avoiding general and plural terms and multiple concepts (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible.

Abbreviations

Define abbreviations that are not standard in this field in a footnote to be placed on the first page of the article. Such abbreviations that are unavoidable in the abstract must be defined at their first mention there, as well as in the footnote. Ensure consistency of abbreviations throughout the article.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

Nomenclature and units

Follow internationally accepted rules and conventions: use the international system of units (SI). If other quantities are mentioned, give their equivalent in SI. You are urged to consult IUB: Biochemical Nomenclature and Related Documents: http://www.chem.qmw.ac.uk/iubmb/ for further information.

Footnotes

Footnotes should be used sparingly. Number them consecutively throughout the article. Many word processors build footnotes into the text, and this feature may be used. Should this not be the case, indicate the position of footnotes in the text and present the footnotes themselves separately at the end of the article.

Artwork

Electronic artwork
General points
• Make sure you use uniform lettering and sizing of your original artwork.
• Preferred fonts: Arial (or Helvetica), Times New Roman (or Times), Symbol, Courier.
• Number the illustrations according to their sequence in the text.
• Use a logical naming convention for your artwork files.
• Indicate per figure if it is a single, 1.5 or 2-column fitting image.
• For Word submissions only, you may still provide figures and their captions, and tables within a single file at the revision stage.
• Please note that individual figure files larger than 10 MB must be provided in separate source files.
A detailed guide on electronic artwork is available on our website:
http://www.elsevier.com/artworkinstructions.
You are urged to visit this site; some excerpts from the detailed information are given here.
Formats
Regardless of the application used, when your electronic artwork is finalized, please 'save as' or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below):
EPS (or PDF): Vector drawings. Embed the font or save the text as 'graphics'.
TIFF (or JPG): Color or grayscale photographs (halftones): always use a minimum of 300 dpi.
TIFF (or JPG): Bitmapped line drawings: use a minimum of 1000 dpi.
TIFF (or JPG): Combinations bitmapped line/half-tone (color or grayscale): a minimum of 500 dpi is required.
Please do not:
• Supply files that are optimized for screen use (e.g., GIF, BMP, PICT, WPG); the resolution is too low.
• Supply files that are too low in resolution.
• Submit graphics that are disproportionately large for the content.

Color artwork
Please make sure that artwork files are in an acceptable format (TIFF (or JPEG), EPS (or PDF), or MS Office files) and with the correct resolution. If, together with your accepted article, you submit usable color figures then Elsevier will ensure, at no additional charge, that these figures will appear in color online (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in color in the printed version. **For color reproduction in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article**. Please

indicate your preference for color: in print or online only. For further information on the preparation of electronic artwork, please see http://www.elsevier.com/artworkinstructions.
Please note: Because of technical complications that can arise by converting color figures to 'gray scale' (for the printed version should you not opt for color in print) please submit in addition usable black and white versions of all the color illustrations.

Figure captions
Ensure that each illustration has a caption. A caption should comprise a brief title (**not** on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

Tables

Please submit tables as editable text and not as images. Tables can be placed either next to the relevant text in the article, or on separate page(s) at the end. Number tables consecutively in accordance with their appearance in the text and place any table notes below the table body. Be sparing in the use of tables and ensure that the data presented in them do not duplicate results described elsewhere in the article. Please avoid using vertical rules.

References

Citation in text
Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either 'Unpublished results' or 'Personal communication'. Citation of a reference as 'in press' implies that the item has been accepted for publication.

Web references
As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

References in a special issue
Please ensure that the words 'this issue' are added to any references in the list (and any citations in the text) to other articles in the same Special Issue.

Reference management software
Most Elsevier journals have a standard template available in key reference management packages. This covers packages using the Citation Style Language, such as Mendeley (http://www.mendeley.com/features/reference-manager) and also others like EndNote (http://www.endnote.com/support/enstyles.asp) and Reference Manager (http://refman.com/support/rmstyles.asp). Using plug-ins to word processing packages which are available from the above sites, authors only need to select the appropriate journal template when preparing their article and the list of references and citations to these will be formatted according to the journal style as described in this Guide. The process of including templates in these packages is constantly ongoing. If the journal you are looking for does not have a template available yet, please see the list of sample references and citations provided in this Guide to help you format these according to the journal style.

If you manage your research with Mendeley Desktop, you can easily install the reference style for this journal by clicking the link below:
http://open.mendeley.com/use-citation-style/fish-and-shellfish-immunology
When preparing your manuscript, you will then be able to select this style using the Mendeley plug-ins for Microsoft Word or LibreOffice. For more information about the Citation Style Language, visit http://citationstyles.org.

Reference formatting
There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/book title, chapter title/article title, year of publication, volume number/book chapter and the pagination must be present. Use of DOI is highly encouraged. The reference style used by the journal will be applied to the accepted article by Elsevier at the proof stage. Note that missing data will be highlighted at proof stage for the author to correct. If you do wish to format the references yourself they should be arranged according to the following examples:

Reference style
Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given.
List: Number the references (numbers in square brackets) in the list in the order in which they appear in the text.
Examples:
Reference to a journal publication:
[1] Van der Geer J, Hanraads JAJ, Lupton RA. The art of writing a scientific article. J Sci Commun 2010;163:51–9.
Reference to a book:
[2] Strunk Jr W, White EB. The elements of style. 4th ed. New York: Longman; 2000.
Reference to a chapter in an edited book:
[3] Mettam GR, Adams LB. How to prepare an electronic version of your article. In: Jones BS, Smith RZ, editors. Introduction to the electronic age, New York: E-Publishing Inc; 2009, p. 281–304.
Note shortened form for last page number. e.g., 51–9, and that for more than 6 authors the first 6 should be listed followed by 'et al.' For further details you are referred to 'Uniform Requirements for Manuscripts submitted to Biomedical Journals' (J Am Med Assoc 1997;277:927–34) (see also http://www.nlm.nih.gov/bsd/uniform_requirements.html).

Video data

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include links to these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the files in one of our recommended file formats with a preferred maximum size of 150 MB. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages at http://www.elsevier.com/artworkinstructions. Note: since video and animation

cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

AudioSlides

The journal encourages authors to create an AudioSlides presentation with their published article. AudioSlides are brief, webinar-style presentations that are shown next to the online article on ScienceDirect. This gives authors the opportunity to summarize their research in their own words and to help readers understand what the paper is about. More information and examples are available at http://www.elsevier.com/audioslides. Authors of this journal will automatically receive an invitation e-mail to create an AudioSlides presentation after acceptance of their paper.

Supplementary material

Elsevier accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. In order to ensure that your submitted material is directly usable, please provide the data in one of our recommended file formats. Authors should submit the material in electronic format together with the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at http://www.elsevier.com/artworkinstructions.

Database linking

Elsevier encourages authors to connect articles with external databases, giving readers access to relevant databases that help to build a better understanding of the described research. Please refer to relevant database identifiers using the following format in your article: Database: xxxx (e.g., TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN). See http://www.elsevier.com/databaselinking for more information and a full list of supported databases.

Submission checklist

The following list will be useful during the final checking of an article prior to sending it to the journal for review. Please consult this Guide for Authors for further details of any item.
Ensure that the following items are present:
One author has been designated as the corresponding author with contact details:
• E-mail address
• Full postal address
All necessary files have been uploaded, and contain:
• Keywords
• All figure captions
• All tables (including title, description, footnotes)
Further considerations
• Manuscript has been 'spell-checked' and 'grammar-checked'
• All references mentioned in the Reference list are cited in the text, and vice versa
• Permission has been obtained for use of copyrighted material from other sources (including the Internet)
Printed version of figures (if applicable) in color or black-and-white
• Indicate clearly whether or not color or black-and-white in print is required.
• For reproduction in black-and-white, please supply black-and-white versions of the figures for

printing purposes.
For any further information please visit our customer support site at http://support.elsevier.com.

After Acceptance

Use of the Digital Object Identifier

The Digital Object Identifier (DOI) may be used to cite and link to electronic documents. The DOI consists of a unique alpha-numeric character string which is assigned to a document by the publisher upon the initial electronic publication. The assigned DOI never changes. Therefore, it is an ideal medium for citing a document, particularly 'Articles in press' because they have not yet received their full bibliographic information. Example of a correctly given DOI (in URL format; here an article in the journal Physics Letters B):
http://dx.doi.org/10.1016/j.physletb.2010.09.059
When you use a DOI to create links to documents on the web, the DOIs are guaranteed never to change.

Online proof correction

Corresponding authors will receive an e-mail with a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to MS Word: in addition to editing text, you can also comment on figures/tables and answer questions from the Copy Editor. Web-based proofing provides a faster and less error-prone process by allowing you to directly type your corrections, eliminating the potential introduction of errors.
If preferred, you can still choose to annotate and upload your edits on the PDF version. All instructions for proofing will be given in the e-mail we send to authors, including alternative methods to the online version and PDF.
We will do everything possible to get your article published quickly and accurately. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the article as accepted for publication will only be considered at this stage with permission from the Editor. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely your responsibility.

Offprints

The corresponding author, at no cost, will be provided with a personalized link providing 50 days free access to the final published version of the article on ScienceDirect. This link can also be used for sharing via email and social networks. For an extra charge, paper offprints can be ordered via the offprint order form which is sent once the article is accepted for publication. Both corresponding and co-authors may order offprints at any time via Elsevier's WebShop (http://webshop.elsevier.com/myarticleservices/offprints). Authors requiring printed copies of multiple articles may use Elsevier WebShop's 'Create Your Own Book' service to collate multiple articles within a single cover (http://webshop.elsevier.com/myarticleservices/booklets).

Author Inquiries

You can track your submitted article at http://www.elsevier.com/track-submission. You can track your accepted article at http://www.elsevier.com/trackarticle. You are also welcome to contact Customer Support via http://support.elsevier.com.