6-7-2017

Transcriptomic and Epigenetic Responses to Environmental Stress in Marine Bivalves with a Focus on Harmful Algal Blooms

Maria Victoria Suarez Ulloa
Florida International University, msuar177@fiu.edu

DOI: 10.25148/etd.FIDC001905

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

TRANSCRIPTOMIC AND EPIGENETIC RESPONSES TO ENVIRONMENTAL STRESS IN MARINE BIVALVES WITH A FOCUS ON HARMFUL ALGAL BLOOMS

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOLOGY

by

María Victoria Suárez Ulloa

2017

To: Dean Michael R. Heithaus
College of Arts, Sciences and Education

This dissertation, written by María Victoria Suárez Ulloa, and entitled Transcriptomic and Epigenetic Responses to Environmental Stress in Marine Bivalves with a Focus on Harmful Algal Blooms, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Heather Bracken-Grissom

_____
Mauricio Rodriguez-Lanetty

_____
Eric Bishop-Von Wettberg

_____
Giri Narasimhan

_____
Jose M. Eirin-Lopez, Major Professor

Date of Defense: June 7, 2017

The dissertation of María Victoria Suárez Ulloa is approved.

_____
Dean Michael R. Heithaus
College of Arts, Sciences and Education

_____
Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2017

COPYRIGHT PAGE

CHAPTER I

Part of the contents of this chapter have been published in:

Suarez-Ulloa V, Fernandez-Tajes J, Manfrin C, Gerdol M, Venier P, and Eirin-López JM (2013). Bivalve omics: state of the art and potential applications for the biomonitoring of harmful marine compounds. *Mar Drugs* 11:4370-4389. doi: 10.3390/md11114370

Suarez-Ulloa V, Gonzalez-Romero R., and Eirin-López JM (2015) Environmental epigenetics: A promising venue for developing next-generation pollution biomonitoring tools in marine invertebrates. *Mar Pollut Bull* 98:5-13. doi:10.1016/j.marpolbul.2015.06.020

CHAPTER II

Suarez-Ulloa V, Fernandez-Tajes J, Aguiar-Pulido V, Rivera-Casas C, Gonzalez-Romero R, Ausio J, Mendez J, Dorado J, and Eirin-López JM (2013). The CHROMEVALOA database: a resource for the evaluation of okadaic acid contamination in the marine environment based on the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis*. *Mar Drugs* 11:830-841. doi: 10.3390/md11030830

CHAPTER III

Suarez-Ulloa V, Fernandez-Tajes J, Aguiar-Pulido V, Prego-Faraldo MV, Florez-Barros F, Sexto-Iglesias A, Mendez J, Eirin-López JM. (2015) Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid. *PeerJ* 3:e1429 doi: 10.7717/peerj.1429

CHAPTER IV

Gonzalez-Romero R*, Suarez-Ulloa V*, Rodriguez-Casariego J, Garcia-Souto D, Diaz G, Smith A, Pasantes JJ, Rand G, Eirin-López JM (2017) Effects of Florida Red Tides on histone variant expression and DNA methylation in the Eastern oyster *Crassostrea virginica*. *Aquat Toxicol* 186:196-204. doi: 10.1016/j.aquatox.2017.03.006

*Both authors have contributed equally to this publication

DEDICATION

To my mother

ACKNOWLEDGMENTS

I would like to say thank you to all the colleagues and collaborators that have contributed in any way to the completion of this dissertation.

First and foremost, thanks to my major advisor Dr. Eirín López (Chema), for his constant guidance and support. I hope we will keep on working together in the future.

Thanks to the members of my academic committee for great advice and feedback through the years. In particular, thanks to Dr. Narasimhan (Giri) for adopting me as one more member of his own research group.

Thanks to my lab mates, most especially to Dr. González-Romero (Rodri) for all the training and the mandarins at lunch time.

Thanks also to Dr. Aguiar-Pulido (Vane), close collaborator since the early days and great guidance with the most technical aspects of this work. Let's keep it up!

To my family and friends, old and new, thank you for your support and for all the happiness through the years. I love you.

Lastly, thanks to the sources of funding that supported my work during these years, including Graduate Assistantship and Dissertation Year Fellowship from the University Graduate School at Florida International University as well as a Research Assistantship from the Spanish Government.

ABSTRACT OF THE DISSERTATION

TRANSCRIPTOMIC AND EPIGENETIC RESPONSES TO ENVIRONMENTAL STRESS IN MARINE BIVALVES WITH A FOCUS ON HARMFUL ALGAL BLOOMS

by

María Victoria Suárez Ulloa

Florida International University, 2017

Miami, Florida

Professor Jose M. Eirin-Lopez, Major Professor

Global change poses new threats for life in the oceans forcing marine organisms to respond through molecular acclimatory and adaptive strategies. Although bivalve molluscs are particularly tolerant and resilient to environmental stress, they must now face the challenge of more frequent and severe Harmful Algal Blooms (HABs) episodes. These massive outbreaks of microalgae produce toxins that accumulate in the tissues of these filter-feeder organisms, causing changes in their gene expression profiles, which in turn modify their phenotype in order to maintain homeostasis. Such modifications in gene expression are modulated by epigenetic mechanisms elicited by specific environmental stimuli, laying the foundations for long-term adaptations.

The present work aims to examine the links between environmental stress in bivalve molluscs (with especial emphasis on Harmful Algal Blooms) and specific epigenetic marks triggering responses through modifications in gene expression patterns. Overall, a better understanding of the molecular strategies underlying the conspicuous stress tolerance observed in bivalve molluscs will provide a framework for developing a

new generation of biomonitoring strategies. In addition, this strategy will represent a valuable contribution to our knowledge in acclimatization, adaptation and survival.

With that goal in mind, the present work has generated transcriptomic data using RNA-Seq and microarray technologies, facilitating the characterization and investigation of the epigenetic mechanisms used by the Mediterranean mussel *Mytilus galloprovincialis* during responses to HAB exposure. That information was made publicly available through a specialized online resource (the Chromevaloa Database, chromevaloa.com) assessing the response of chromatin-associated transcripts to Okadaic Acid.

Specific epigenetic marks have been assessed under lab-controlled exposure experiments simulating the natural development of the HAB Florida Red Tide (FRT). Results demonstrate a role for the phosphorylation of histone H2A.X and DNA methylation in the response to FRT in the Eastern oyster *Crassostrea virginica*. Lastly, the study of co-expression networks based on RNA-Seq data series from the Pacific oyster *Crassostrea gigas* reveals dynamic transcriptomic patterns that vary with time, stressor and tissue. However, consistent functional profiles support the existence of a core response to general conditions of environmental stress. Such response involves metabolic and transport processes, response to oxidative stress and protein repair or disposal, as well as the activation of immune mechanisms supporting a tightly intertwined neuroendocrine-immune regulatory system in bivalves.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER I

INTRODUCTION

Marine invertebrates constitute the largest group of macroscopic species in the sea (Ruppert et al. 2004). Among them, bivalve molluscs stand out not only for their fundamental role in the marine ecosystem, but also for their commercial value in aquaculture industry (Gosling 2003, Newell 2004). Additionally, this group of organisms displays key features supporting their application as sentinel organisms for the biomonitoring of water quality, particularly in coastal and estuarine areas, including (but not limited to): ubiquitous distribution, easy accessibility, filtering lifestyle, as well as strong resistance to a wide range of pollutants (Collin et al. 2010, Campos et al. 2012, Fernandez-Tajes et al. 2012, Luchmann et al. 2012, Zhang et al. 2012, Milan et al. 2013). Consequently, bivalves have been traditionally used for biomonitoring purposes (Wells et al. 2001, Florez-Barros et al. 2011), often following physiological or biochemical approaches. Yet, it was not until recently that omic approaches have been implemented in the study of marine bivalves, primarily due to the recent advances in sequencing technologies and the substantial reduction in the associated costs.

So far, omic studies in bivalves have been mainly focused towards the characterization of genomes, transcriptomes and proteomes, although not necessarily in this specific order. In fact, pioneer omic studies in bivalves were eminently based on transcriptomes, helping to set up the foundations for subsequent proteomic and genomic studies (Tanguy et al. 2005, Brown et al. 2006, Dondero et al. 2006, Venier et al. 2006, Zapata et al. 2009). On the other hand, the metabolomic and epigenomic characterization of bivalves still represent emerging disciplines, required to completing the necessary framework for integrative approaches.

The availability of complete reference genome sequences, along with the development of high-throughput DNA sequencing and bioinformatic methods, has ignited the interest for the epigenetic study of marine organisms, providing insights into their strategies to cope with a changing environment as well as paving the road for the development of next-generation biomarkers of marine pollution in a context of global change.

Harmful Algal Blooms: An environmental stressor of growing concern

Among the different sources of marine pollution threatening marine life, Harmful Algal Blooms (HABs) stand out due to their increasing frequency, duration and severity during the last decades (Anderson 2009). HABs are massive outbreaks of toxic microalgae (commonly dinoflagellates) with severe consequences for coastal ecosystems and human populations. HABs have an especially negative impact on filter-feeding shellfish such as mussels and oysters since these organisms are able to accumulate large amounts of toxins in their tissues, thus representing a hazard for food safety in shellfish aquaculture.

There are different types of HABs depending on the main toxin or group of toxins responsible as well as the syndrome associated to them. Among these, the most relevant include the Neurotoxic Shellfish Poisoning (NSP), the Paralytic Shellfish Poisoning (PSP), the Diarrhetic Shellfish Poisoning (DSP), the Ciguatera Shellfish Poisoning (CSP) and the Azaspirazid Shellfish Poisoning (ASP) (Van Dolah 2000). More especifically, DSP and NSP are associated with Okadaic Acid (OA) and brevetoxins respectively, and in both cases they have been linked with genotoxicity and tumor development (Murrell

and Gibson 2009, Valdiglesias et al. 2012). Diarrhetic Shellfish Poisoning (DSP) blooms represent a major threat in widespread geographic areas comprising the Atlantic coast of Europe, Chile and Japan (Reguera et al. 2014), where natural outbreaks of toxic *Dinophysis* and *Prorocentrum* microalgae produce large amounts of OA biotoxins (Sellner et al. 2003). OA is the primary cause of acute DSP intoxication in human consumers of shellfish, causing significant economic losses, most notably in the mussel aquaculture industry (the primary economic force in these areas).

Although OA is considered non-lethal, it is a well-known tumor promoter encompassing harmful effects at concentrations as low as 5 nM (Valdiglesias et al. 2011). The southeastern U.S. and its unique coastal habitats are also strongly affected by HABs, notably in the form of Florida Red Tides. These episodes are characterized by massive and colored outbreaks of the NSP-producing dinoflagellate *Karenia brevis*, responsible for the production of different toxins of the same family commonly categorized as brevetoxins or PbTx toxins (Brand and Compton 2007), causing high mortality rates in populations of marine invertebrates, fishes and marine mammals (Brand et al. 2012). Florida Red Tides also constitute a serious threat for human populations in coastal areas, as the consumption of brevetoxin-ladden shellfish affects the nervous system, and the exposure to oceanic aerosol carrying brevetoxins may cause respiratory distress and severe allergic reactions in persons with respiratory conditions (Fleming et al. 2007). Given their ability to damage DNA (Radwan and Ramsdell 2008), these biotoxins represent ideal systems to study the epigenetic mechanisms underlying the responses bivalve mollucs to environmental stress, as well as for investigating the potential application of transcriptomic signatures as biomarkers of HAB pollution in the oceans.

A *framework for the epigenetic analysis of environmental responses*

One of the most amazing features of the eukaryotic genetic material is its ability to be packed and organized within a tiny cell nucleus up to 200,000 times smaller than the cell itself. That is possible thanks to the wrapping of the DNA molecule around chromosomal proteins (among which histones are the most important, both structurally and functionally), constituting a dynamic polymer organized in fundamental nucleosome subunits known as chromatin. However, chromatin also participates in the functional classification of the information contained in the genome (Allis et al. 2007), providing a framework for the study of epigenetics, defined as the heritable changes in gene expression resulting from modifications in chromatin structure, without involving changes in the genetic information stored in the DNA sequence (Allis et al. 2007).

Various mechanisms have the potential to encode epigenetic information including DNA methylation, the replacement of canonical histones by specialized histone variants, histone post-translational modifications (PTMs), non-coding RNAs, and transcription factor regulatory networks, among others (Kouzarides 2007, Ptashne 2007, Arya et al. 2010, Talbert and Henikoff 2010, Mercer and Mattick 2013). Although different in nature, all these mechanisms are capable of triggering dynamic modifications of the chromatin structure in response to external stimuli (Talbert and Henikoff 2014). However, while some of these modifications last for a few seconds before being rapidly reverted to a basal state (e.g., acetylation of histones allowing expression of genes specifically involved in DNA repair), others may persist in the chromatin of the same specific cell for decades (e.g., DNA methylation leading to gene silencing during the differentiation of neural stem cells) (Williams et al. 2014). Furthermore, what is truly

amazing about these epigenetic marks is their ability to transcend across generations, constituting the basis for acclimatization and long-term adaptation (i.e., conserved DNA methylation imprinting in the germ line) (Gapp et al. 2014, Heard and Martienssen 2014) (Figure 1).

Overall, epigenetics constitutes the next frontier for understanding how mechanisms of temporal and spatial control of gene activity work during environmental responses (Holliday 1990). In order to do so, it is fundamental to investigate not only the links between specific epigenetic marks and the subsequent modifications in chromatin structure and gene expression, but also the environmental factors leading to these epigenetic marks in the first place (Cortessis et al. 2012). That strategy constitutes the basis for environmental epigenetic analyses (Baccarelli and Bollati 2009, Bollati and Baccarelli 2010), providing information about the mechanisms by which different environmental factors influence phenotypic variation, both within individuals and across generations (Cortessis et al. 2012, Talbert and Henikoff 2014). Therefore, environmental epigenetics opens up an innovative venue for developing fast and sensible environmental biomonitoring programs, given the dynamic and potentially reversible nature of the different types of epigenetic marks. Most importantly, it seeks a better understanding of the mechanisms that allow organisms to adapt to a changing environment, with critical implications for conservation and management purposes.

Epigenetic background in marine bivalves

So far, epigenetic studies developed in marine invertebrates have been primarily focused on two major mechanisms: DNA methylation and chromatin specialization by

histone variants (Gavery and Roberts 2010, González-Romero et al. 2012). In the former, the addition of methyl groups at gene promoters (CpG islands) constitutes a mark usually involved in silencing of gene expression. However, it is a common pattern for invertebrates to display DNA methylation predominantly within gene bodies, associated with gene expression regulation and alternative splicing (Su et al. 2011, Gavery and Roberts 2013). Genome-wide DNA methylation patterns have become recently available in the Pacific oyster *Crassostrea gigas* and the Zhikong scallop *Chlamys farreri* (Gavery and Roberts 2013, Sun et al. 2014), establishing links between DNA hypomethylation and transcription of genes potentially linked to phenotypic plasticity and adaptation (Gavery and Roberts 2010).

Similarly to the case of DNA methylation, the presence of specialized histone variants in marine invertebrates hints the existence of elaborated epigenetic mechanisms potentially involved in adaptive environmental responses (Talbert and Henikoff 2014). Yet, it is still necessary to clearly ascertain the extent of such diversification. How many different variants are there and what is their taxonomic distribution? And most importantly, how do these variants impart functional specialization to chromatin? The strategy to tackle the first question seems straightforward in those species where histone variants have been described (i.e., H2A.X, H2A.Z and macroH2A from mussels (González-Romero et al. 2012, Rivera-Casas et al. 2016a, Rivera-Casas et al. 2016b) facilitating the *in silico* characterization of homologous genes from molecular databases and their subsequent isolation in different species following PCR-based approaches.

Concomitantly, tailor-made tools based on real-time PCR techniques also enable several complementary analyses including chromosomal and physical mapping, as well

as the quantification of histone variant gene copy numbers (Eirín-López et al. 2002, Albig et al. 2003, Eirín-López et al. 2004). Overall, the availability of genomic and transcriptomic data in public repositories represents a critical advantage in order to efficiently characterize epigenetically relevant genes, transcripts and proteins in non-traditional model organisms such as bivalves.

Transcriptomic data and gene expression studies in marine bivalves

Pioneer transcriptomic studies in bivalves have been progressively complemented with differential expression analyses in response to different stressors (pollutants and pathogens) using technologies such as cDNA libraries, Suppression Subtractive Hybridization libraries (SSH) and microarrays (Venier et al. 2006, Zapata et al. 2009, Collin et al. 2010, Canesi et al. 2011, Chapman et al. 2011, de Lorgeril et al. 2011, Milan et al. 2011, Venier et al. 2011, Egas et al. 2012, Luchmann et al. 2012, Moreira et al. 2012, Philipp et al. 2012). Consequently, the development of cDNA and SSH libraries has led to a significant increase in the number of Expressed Sequence Tags (ESTs) in databases, constituting the basis for DNA microarray technology.

Microarrays have been primarily used in bivalves to study transcriptional responses to different environmental stressors (Dondero et al. 2006, Venier et al. 2006, Lockwood et al. 2010, Manfrin et al. 2010, Lockwood and Somero 2011, Venier et al. 2011). Nowadays, the combination of microarray and NGS technologies is significantly speeding up *de novo* gene discovery (Ghiselli et al. 2012), allowing transcriptomic analyses of non-model organisms (Francis et al. 2013) including bivalves (Moreira et al. 2012, Pante et al. 2012, Yue et al. 2012, Shi et al. 2013). Additionally, the RNA-Seq is

now the preferred choice (over microarray analysis) for transcriptome profiling in bivalves (Gerdol et al. 2012, Gerdol et al. 2014). RNA-Seq provides a more precise measurement of transcript levels compared with other methods, delivering unbiased and unparalleled gene expression information. However, *de novo* transcriptome assembly poses specific challenges of its own, given that the number of sequenced reads corresponding to different transcripts can vary over several orders of magnitude due to differences in expression levels (Gerdol et al. 2012). Consequently, sequencing coverage is susceptible to be heterogeneous throughout the whole transcriptome (i.e., higher coverage levels of highly expressed transcripts) requiring transcript normalization before data analysis. On the other hand, the availability of reference genome sequences (such as the case of the Pacific oyster *C. gigas*) has boosted the proliferation of transcriptomic studies in bivalves, as demonstrated by the growing number of registered Bioprojects categorized as "Transcriptome or gene expression" in the NCBI repository between 2013 and 2017 (Figure 2).

The presence of transcriptional modifications in response to fluctuations in environmental factors provides a potential tool to unveil transitory physiological adjustments, irreversible functional deficits and taxon-specific acclimatory features of the organisms. For instance, the study of transcript signatures has revealed subtle differences between the mussels *Mytilus galloprovincialis* and *M. trossulus* under salinity stress (Lockwood and Somero 2011). Additionally, transcriptomic profiles have also been analyzed in different tissues from the Eastern oyster *Crassostrea virginica* using chemical measures and DNA microarray analyses. In this case, progressive computational data testing confirmed the reliability of the DNA microarray metrics, and

provided insights on the response mechanisms to temperature and pH (Chapman et al. 2011). Concurrently, the construction of SSH libraries led to the identification of putative biomarkers involved in the response to several stress factors relevant for aquaculture, notably thermal stress (Meistertzheim et al. 2007, Lockwood et al. 2010), hypoxia (David et al. 2005) and pathogen infection (Araya et al. 2010, Morga et al. 2011).

Additionally, transcriptomic studies have helped identifying specific groups of genes involved in the acclimatization of bivalves to heterogeneous environments. This is the case of studies suggesting that genes involved in defense and the innate immune response play a pivotal role as determinants of the resistance to summer mortality in the Pacific oyster *C. gigas* (Fleury et al. 2010, Chaney and Gracey 2011). Indeed, sequencing and data mining of ESTs are essential steps for the comparative identification of molecules and related pathways of response to specific stimuli. Such task is greatly facilitated by the availability of high throughput sequencing technologies yielding unprecedented amounts of sequence data. RNA-Seq has also been used to identify developmentally regulated genes in *Crassostrea angulata* and *Meretrix meretrix* (Huan et al. 2012, Qin et al. 2012). Overall, the combination of molecular data with traditional physiological and population studies provides a new framework improving the management of livestock under different stress conditions (Coppe et al. 2012, Pante et al. 2012).

The availability of genomic and transcriptomic data supports the development of epigenetic analyses in non-model organisms, facilitating gene discovery and characterization. For instance, data mining of transcriptomic data developed during the this dissertation has recently led to the characterization of the histone variant macroH2A

for the first time in an invertebrate, the mussel *Mytilus galloprovincialis* (Rivera-Casas et al. 2016a), illustrating the relevance of genomic and transcriptomic resources for other closely related areas of research.

Specialized data resources

Most molecular information related to bivalve molluscs is stored in general repositories like the National Center for Biotechnology Information (NCBI) or the European Molecular Biology Laboratory (EMBL). Yet, a number of specialized databases have become publicly available during the last decade (see Table 1).

For instance, the Marine Genomics Project (McKillen et al. 2005) comprises ESTs and microarray data from marine organisms in a broad sense, although most recent databases aim to extend the molecular knowledge to specific groups of species. Within this context, the genome draft of the pearl oyster *Pinctada fucata* (version 1.0) has been made publicly available through a specific genome browser [OIST Marine Genomics Unit. Available online: http://marinegenomics.oist.jp/ accessed on 27 February 2013)] (Takeuchi et al. 2012). Similarly, repositories such as the Mytibase (Venier et al. 2009) represent useful resources for the transcriptomic study of the mussel *Mytilus*, providing large-scale ESTs with critical relevance for developing microarray platforms aimed to the biomonitoring of marine pollution. ESTs have been also put together for other bivalve species, including the clams *Ruditapes philippinarum* [RuphiBase (Milan et al. 2011)] and *Chamelea gallina* [ChamaleaBase (Coppe et al. 2012)], the mussel *Bathymodiolus azoricus* [DeepSeaVent database (Egas et al. 2012)], as well as the Pacific oyster *C. gigas* [GigasDatabase (Fleury et al. 2009)]. In addition, this list is completed by the

CHROMEVALOA database, constructed as part of this doctoral dissertation and described in Chapter 2. This resource provides access to the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis* in response to the toxic dinoflagellate *Prorocentrum lima* (Suarez-Ulloa et al. 2013). The list of specialized databases of ESTs and transcriptomic data from bivalves is summarized in Table 1.

Toxin biomonitoring during HABs

HABs cause highly deleterious effects on marine organisms and human populations in coastal areas. Several studies have tried to tackle the effects of marine biotoxins on bivalves from an omic perspective, notably by using the Mediterranean mussel *M. galloprovincialis* as model organism. One of these studies used a cDNA microarray to examine the effects of a 35 day-long exposure to the biotoxin okadaic acid (OA) in mussels. The obtained results identified several transcripts as potential OA stress biomarkers (Manfrin et al. 2010). Additionally, the up-regulation of several stress-related proteins involved in apoptosis, proteolysis and cytoskeleton destabilization, suggested a harmful effect of OA in mussels.

The present doctoral dissertation complemented this effort by characterizing the transcriptome of mussels exposed to sub-lethal concentrations of the OA-producing dinoflagellate *P. lima* (Suarez-Ulloa et al. 2015), as well as the identification and annotation of chromatin-associated transcripts (Suarez-Ulloa et al. 2013). As a result, this work identified a number of genes whose expression was significantly influenced by OA, revealing potential HAB biomarkers. Beyond the present dissertation, an additional study investigated the molecular mechanisms underlying responses to paralytic toxins produced

by the dinoflagellate *Alexandrium minutum* in the digestive gland of mussels. While the obtained results revealed very mild effects on gene expression, the effects of these toxins were enough to identify a few potential biomarkers of HAB contamination (Gerdol et al. 2014). Targeted gene expression analyses using quantitative PCR (qPCR), as well as histopathological and physiological methods, complement omic analyses. Accordingly, changes in the epithelium of the digestive gland from the oyster *C. gigas* were detected upon exposure to the OA-producing dinoflagellate *P. lima*, along with deregulation of stress-related genes (Romero-Geraldo et al. 2016). These results support the notion suggesting that, despite the relatively high tolerance of bivalves towards HAB exposure, the toxic effects of OA do indeed cause an impact at the cellular and molecular level.

Final remarks

The application of omic approaches based on high-throughput methods constitutes a very powerful tool for deciphering the molecular mechanisms underlying the response and acclimatization of bivalves to environmental changes. Nowadays, the analysis of gene expression profiles is helping define new metrics complementing traditional chemical and biomarker-based strategies, improving the monitoring of coastal waters. On one hand, Suppression Subtractive Hybridization (SSH) and microarray techniques have largely generated important information concerning differential expression of specific genes in response to different sources of stress in the environment. On the other hand, the high-throughput sequencing of transcriptomes (RNA-Seq) is progressively adding further depth and new details about the biology of bivalve molluscs and other sentinel organisms. The application of these technologies sets the framework

for biomarker development based on multi-gene expression profiles, improving traditional methods in terms of both sensitivity and specificity. A novel approach to this goal could be classified within the relatively recent field of research known as Environmental Epigenetics, characterizing cause-effect relationships between environmental factors and the corresponding dynamic changes in epigenetic marks leading to specific phenotypes.

Unbiased omic research and targeted approaches represent complementary strategies in modern molecular biology. However, omic research necessarily involves the analysis of massive amounts of data that require specialized computational methods. Thus, in order to answer the primary research question posed by the present dissertation (i.e., How do marine bivalves respond to environmental stress resulting from HAB pollution?) molecular biology and bioinformatics will be required in similar proportions.

In particular, this dissertation aims to provide an answer on the following specific questions:

- What genes encompass epigenetic relevance in bivalves?

(Chapter II)

- What is the transcriptomic response of bivalves exposed to HABs?

(Chapter III)

- What is the epigenetic response of bivalves exposed to HABs?

(Chapter IV)

- What is common and what is unique in the response of bivalves to different environmental stressors?

(Chapter V)

Through the use of different model organisms and different sources of environmental stress, including two different types of HAB, the present dissertation makes a relevant contribution to our current knowledge in functional genomics and epigenetics for a taxonomic group of great environmental and economic importance.

References

Albig, W., U. Warthorst, B. Drabent, E. Prats, L. Cornudella, and D. Doenecke. 2003. Mytilus edulis core histone genes are organized in two clusters devoid of linker histone genes. J. Mol. Evol. 56:597-606.

Allis, C. D., T. Jenuwein, and D. Reinberg. 2007. Epigenetics. Cold Spring Harbor Laboratory Press, New York.

Anderson, D. M. 2009. Approaches to monitoring, control and management of harmful algal blooms (HABs). Ocean Coast Manag 52:342-347.

Araya, M. T., F. Markham, D. R. Mateo, P. McKenna, G. R. Johnson, F. C. Berthe, and A. Siah. 2010. Identification and expression of immune-related genes in hemocytes of soft-shell clams, *Mya arenaria*, challenged with *Vibrio splendidus*. Fish Shellfish Immunol 29:557-564.

Arya, G., A. Maitra, and S. A. Grigoryev. 2010. A structural perspective on the where, how, why, and what of nucleosome positioning. J Biomol Struct Dyn 27:803-820.

Baccarelli, A. and V. Bollati. 2009. Epigenetics and environmental chemicals. Curr Opin Pediatr 21:243-251.

Bollati, V. and A. Baccarelli. 2010. Environmental epigenetics. Heredity 105:105-112.

Brand, L. E., L. Campbell, and E. Bresnan. 2012. *Karenia*: the biology and ecology of a toxic genus. Harmful Algae 14:156-178.

Brand, L. E. and A. Compton. 2007. Long-term increase in *Karenia brevis* abundance along the Southwest Florida Coast. Harmful Algae 6:232-252.

Brown, M., I. M. Davies, C. F. Moffat, and J. A. Craft. 2006. Application of SSH and a macroarray to investigate altered gene expression in *Mytilus edulis* in response to exposure to benzo[a]pyrene. Mar Environ Res 62 Suppl:S128-135.

Campos, A., S. Tedesco, V. Vasconcelos, and S. Cristobal. 2012. Proteomic research in bivalves: towards the identification of molecular markers of aquatic pollution. J Proteomics 75:4346-4359.

Canesi, L., A. Negri, C. Barmo, M. Banni, G. Gallo, A. Viarengo, and F. Dondero. 2011. The organophosphate Chlorpyrifos interferes with the responses to 17beta-estradiol in the digestive gland of the marine mussel *Mytilus galloprovincialis*. PLoS ONE 6:e19803.

Chaney, M. L. and A. Y. Gracey. 2011. Mass mortality in Pacific oysters is associated with a specific gene expression signature. Mol Ecol 20:2942-2954.

Chapman, R. W., A. Mancia, M. Beal, A. Veloso, C. Rathburn, A. Blair, A. F. Holland, G. W. Warr, G. Didinato, I. M. Sokolova, E. F. Wirth, E. Duffy, and D. Sanger. 2011. The transcriptomic responses of the eastern oyster, *Crassostrea virginica*, to environmental conditions. Mol Ecol 20:1431-1449.

Collin, H., A. L. Meistertzheim, E. David, D. Moraga, and I. Boutet. 2010. Response of the Pacific oyster *Crassostrea gigas*, Thunberg 1793, to pesticide exposure under experimental conditions. J Exp Biol 213:4010-4017.

Coppe, A., S. Bortoluzzi, G. Murari, I. A. Marino, L. Zane, and C. Papetti. 2012. Sequencing and characterization of striped venus transcriptome expand resources for clam fishery genetics. PLoS ONE 7:e44185.

Cortessis, V. K., D. C. Thomas, A. J. Levine, C. V. Breton, T. M. Mack, K. D. Siegmund, R. W. Haile, and P. W. Laird. 2012. Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. Hum Genet 131:1565-1589.

David, E., A. Tanguy, K. Pichavant, and D. Moraga. 2005. Response of the Pacific oyster *Crassostrea gigas* to hypoxia exposure under experimental conditions. FEBS J 272:5635-5652.

de Lorgeril, J., R. Zenagui, R. D. Rosa, D. Piquemal, and E. Bachere. 2011. Whole transcriptome profiling of successful immune response to *Vibrio* infections in the oyster *Crassostrea gigas* by digital gene expression analysis. PLoS ONE 6:e23142.

Dondero, F., L. Piacentini, F. Marsano, M. Rebelo, L. Vergani, P. Venier, and A. Viarengo. 2006. Gene transcription profiling in pollutant exposed mussels (*Mytilus* spp.) using a new low-density oligonucleotide microarray. Gene 376:24-36.

Egas, C., M. Pinheiro, P. Gomes, C. Barroso, and R. Bettencourt. 2012. The transcriptome of *Bathymodiolus azoricus* gill reveals expression of genes from endosymbionts and free-living deep-sea bacteria. Mar Drugs 10:1765-1783.

Eirín-López, J. M., A. M. González-Tizón, A. Martínez, and J. Méndez. 2002. Molecular and evolutionary analysis of mussel histone genes (*Mytilus* spp.): possible evidence of an "orphon origin" for H1 histone genes. J Mol Evol 55:272-283.

Eirín-López, J. M., M. F. Ruiz, A. M. González-Tizón, A. Martínez, L. Sánchez, and J. Méndez. 2004. Molecular evolutionary characterization of the mussel *Mytilus* histone multigene family: first record of a tandemly repeated unit of five histone genes containing an H1 subtype with "orphon" features. J Mol Evol 58:131-144.

Fernandez-Tajes, J., A. Arias-Perez, M. Fernandez-Moreno, and J. Mendez. 2012. Sharp decrease of genetic variation in two Spanish localities of razor clam *Ensis siliqua*: natural fluctuation or Prestige oil spill effects? Ecotoxicology 21:225-233.

Fleming, L. E., B. Kirkpatrick, L. C. Backer, J. A. Bean, A. Wanner, A. Reich, J. Zaias, Y. S. Cheng, R. Pierce, J. Naar, W. M. Abraham, and D. G. Baden. 2007. Aerosolized red-tide toxins (brevetoxins) and asthma. Chest 131:187-194.

Fleury, E., A. Huvet, C. Lelong, J. de Lorgeril, V. Boulo, Y. Gueguen, E. Bachere, A. Tanguy, D. Moraga, C. Fabioux, P. Lindeque, J. Shaw, R. Reinhardt, P. Prunet, G. Davey, S. Lapegue, C. Sauvage, C. Corporeau, J. Moal, F. Gavory, P. Wincker, F. Moreews, C. Klopp, M. Mathieu, P. Boudry, and P. Favrel. 2009. Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. BMC Genomics 10:341.

Fleury, E., J. Moal, V. Boulo, J. Y. Daniel, D. Mazurais, A. Henaut, C. Corporeau, P. Boudry, P. Favrel, and A. Huvet. 2010. Microarray-based identification of gonad transcripts differentially expressed between lines of Pacific oyster selected to be resistant or susceptible to summer mortality. Mar Biotechnol (NY) 12:326-339.

Florez-Barros, F., M. Prado-Alvarez, J. Mendez, and J. Fernandez-Tajes. 2011. Evaluation of genotoxicity in gills and hemolymph of clam *Ruditapes decussatus* fed with the toxic dinoflagellate *Prorocentrum lima*. J Toxicol Environ Health A 74:971-979.

Francis, W. R., L. M. Christianson, R. Kiko, M. L. Powers, N. C. Shaner, and S. H. Haddock. 2013. A comparison across non-model animals suggests an optimal sequencing depth for de novo transcriptome assembly. BMC Genomics 14:167.

Gapp, K., L. von Ziegler, R. Y. Tweedie-Cullen, and I. M. Mansuy. 2014. Early life epigenetic programming and transmission of stress-induced traits in mammals: how and when can environmental factors influence traits and their transgenerational inheritance? Bioessays 36:491-502.

Gavery, M. R. and S. B. Roberts. 2010. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). BMC Genomics 11:483.

Gavery, M. R. and S. B. Roberts. 2013. Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. PeerJ 1:e215.

Gerdol, M., G. De Moro, C. Manfrin, A. Milandri, E. Riccardi, A. Beran, P. Venier, and A. Pallavicini. 2014. RNA sequencing and de novo assembly of the digestive gland transcriptome in *Mytilus galloprovincialis* fed with toxinogenic and non-toxic strains of *Alexandrium minutum*. BMC Res Notes 7:722.

Gerdol, M., G. De Moro, C. Manfrin, P. Venier, and A. Pallavicini. 2012. Big defensins and mytimacins, new AMP families of the Mediterranean mussel *Mytilus galloprovincialis*. Dev Comp Immunol 36:390-399.

Ghiselli, F., L. Milani, P. L. Chang, D. Hedgecock, J. P. Davis, S. V. Nuzhdin, and M. Passamonti. 2012. De Novo assembly of the Manila clam *Ruditapes philippinarum* transcriptome provides new insights into expression bias, mitochondrial doubly uniparental inheritance and sex determination. Mol Biol Evol 29:771-786.

González-Romero, R., C. Rivera-Casas, L. J. Frehlick, J. Méndez, J. Ausió, and J. M. Eirín-López. 2012. Histone H2A (H2A.X and H2A.Z) variants in molluscs: molecular characterization and potential implications for chromatin dynamics. PLoS ONE 7:e30006.

Gosling, E. M. 2003. Bivalve Molluscs: Biology, Ecology and Culture. Oxford Fishing New Books, Blackwell Science, Oxford, UK.

Heard, E. and R. A. Martienssen. 2014. Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157:95-109.

Holliday, R. 1990. Mechanisms for the control of gene activity during development. Biol Rev Camb Philos Soc 65:431-471.

Huan, P., H. Wang, and B. Liu. 2012. Transcriptomic analysis of the clam *Meretrix meretrix* on different larval stages. Mar Biotechnol (NY) 14:69-78.

Kouzarides, T. 2007. Chromatin modifications and their function. Cell 128:693-705.

Lockwood, B. L., J. G. Sanders, and G. N. Somero. 2010. Transcriptomic responses to heat stress in invasive and native blue mussels (genus *Mytilus*): molecular correlates of invasive success. J Exp Biol 213:3548-3558.

Lockwood, B. L. and G. N. Somero. 2011. Transcriptomic responses to salinity stress in invasive and native blue mussels (genus *Mytilus*). Mol Ecol 20:517-529.

Luchmann, K. H., J. J. Mattos, M. N. Siebert, T. S. Dorrington, G. Toledo-Silva, P. H. Stoco, E. C. Grisard, and A. C. Bainy. 2012. Suppressive subtractive hybridization libraries prepared from the digestive gland of the oyster *Crassostrea brasiliana* exposed to a diesel fuel water-accommodated fraction. Environ Toxicol Chem 31:1249-1253.

Manfrin, C., R. Dreos, S. Battistella, A. Beran, M. Gerdol, L. Varotto, G. Lanfranchi, P. Venier, and A. Pallavicini. 2010. Mediterranean mussel gene expression profile induced by okadaic acid exposure. Environ Sci Technol 44:8276-8283.

McKillen, D. J., Y. A. Chen, C. Chen, M. J. Jenny, H. F. Trent, 3rd, J. Robalino, D. C. McLean, Jr., P. S. Gross, R. W. Chapman, G. W. Warr, and J. S. Almeida. 2005. Marine genomics: a clearing-house for genomic and transcriptomic data of marine organisms. BMC Genomics 6:34.

Meistertzheim, A. L., A. Tanguy, D. Moraga, and M. T. Thebault. 2007. Identification of differentially expressed genes of the Pacific oyster *Crassostrea gigas* exposed to prolonged thermal stress. FEBS J 274:6392-6402.

Mercer, T. R. and J. S. Mattick. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol 20:300-307.

Milan, M., A. Coppe, R. Reinhardt, L. M. Cancela, R. B. Leite, C. Saavedra, C. Ciofi, G. Chelazzi, T. Patarnello, S. Bortoluzzi, and L. Bargelloni. 2011. Transcriptome sequencing and microarray development for the Manila clam, *Ruditapes philippinarum*: genomic tools for environmental monitoring. BMC Genomics 12:234.

Milan, M., M. Pauletto, T. Patarnello, L. Bargelloni, M. G. Marin, and V. Matozzo. 2013. Gene transcription and biomarker responses in the clam *Ruditapes philippinarum* after exposure to ibuprofen. Aquat Toxicol 126:17-29.

Moreira, R., P. Balseiro, J. V. Planas, B. Fuste, S. Beltran, B. Novoa, and A. Figueras. 2012. Transcriptomics of in vitro immune-stimulated hemocytes from the Manila clam *Ruditapes philippinarum* using high-throughput sequencing. PLoS ONE 7:e35009.

Morga, B., T. Renault, N. Faury, B. Chollet, and I. Arzul. 2011. Cellular and molecular responses of haemocytes from *Ostrea edulis* during in vitro infection by the parasite *Bonamia ostreae*. Int J Parasitol 41:755-764.

Murrell, R. N. and J. E. Gibson. 2009. Brevetoxins 2, 3, 6, and 9 show variability in potency and cause significant induction of DNA damage and apoptosis in Jurkat E6-1 cells. Arch Toxicol 83:1009-1019.

Newell, R. 2004. Ecosystem influences of natural and cultivated populations of suspension-feeding bivalve molluscs: a review. J. Shellfish Res. 23:51-61.

Pante, E., A. Rohfritsch, V. Becquet, K. Belkhir, N. Bierne, and P. Garcia. 2012. SNP detection from de novo transcriptome sequencing in the bivalve *Macoma balthica*: marker development for evolutionary studies. PLoS ONE 7:e52302.

Philipp, E. E., L. Kraemer, F. Melzner, A. J. Poustka, S. Thieme, U. Findeisen, S. Schreiber, and P. Rosenstiel. 2012. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. PLoS ONE 7:e33091.

Ptashne, M. 2007. On the use of the word 'epigenetic'. Curr Biol 17:R233-236.

Qin, J., Z. Huang, J. Chen, Q. Zou, W. You, and C. Ke. 2012. Sequencing and de novo analysis of *Crassostrea angulata* (Fujian oyster) from 8 different developing phases using 454 GSFlx. PLoS ONE 7:e43653.

Radwan, F. F. and J. S. Ramsdell. 2008. Brevetoxin forms covalent DNA adducts in rat lung following intratracheal exposure. Environ Health Perspect 116:930-936.

Reguera, B., P. Riobo, F. Rodriguez, P. A. Diaz, G. Pizarro, B. Paz, J. M. Franco, and J. Blanco. 2014. Dinophysis toxins: causative organisms, distribution and fate in shellfish. Mar Drugs 12:394-461.

Rivera-Casas, C., R. Gonzalez-Romero, M. S. Cheema, J. Ausio, and J. M. Eirin-Lopez. 2016a. The characterization of macroH2A beyond vertebrates supports an ancestral origin and conserved role for histone variants in chromatin. Epigenetics 11:415-425.

Rivera-Casas, C., R. Gonzalez-Romero, A. Vizoso-Vazquez, M. S. Cheema, M. E. Cerdan, J. Mendez, J. Ausio, and J. M. Eirin-Lopez. 2016b. Characterization of mussel H2A.Z.2: a new H2A.Z variant preferentially expressed in germinal tissues from *Mytilus*. Biochem Cell Biol 94:480-490.

Romero-Geraldo, R., N. Garcia-Lagunas, and N. Y. Hernandez-Saavedra. 2016. *Crassostrea gigas* exposure to the dinoflagellate *Prorocentrum lima*: Histological

and gene expression effects on the digestive gland. Mar Environ Res 120:93-102.

Ruppert, E. E., R. S. Fox, and R. D. Barnes. 2004. Invertebrate Zoology: A Functional Evolutionary Approach. 7th edition. Cengage Learning, Stamford, CT, USA.

Sellner, K. G., G. J. Doucette, and G. J. Kirkpatrick. 2003. Harmful algal blooms: causes, impacts and detection. J Ind Microbiol Biotechnol 30:383-406.

Shi, Y., C. Yu, Z. Gu, X. Zhan, Y. Wang, and A. Wang. 2013. Characterization of the pearl oyster (*Pinctada martensii*) mantle transcriptome unravels biomineralization genes. Mar Biotechnol (NY) 15:175-187.

Su, Z., L. Han, and Z. Zhao. 2011. Conservation and divergence of DNA methylation in eukaryotes: new insights from single base-resolution DNA methylomes. Epigenetics 6:134-140.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, M. V. Prego-Faraldo, F. Florez-Barros, A. Sexto-Iglesias, J. Mendez, and J. M. Eirin-Lopez. 2015. Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid. PeerJ 3.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, C. Rivera-Casas, R. Gonzalez-Romero, J. Ausio, J. Mendez, J. Dorado, and J. M. Eirin-Lopez. 2013. The CHROMEVALOA database: a resource for the evaluation of okadaic acid contamination in the marine environment based on the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis*. Mar Drugs 11:830-841.

Sun, Y., R. Hou, X. Fu, C. Sun, S. Wang, C. Wang, N. Li, L. Zhang, and Z. Bao. 2014. Genome-wide analysis of DNA methylation in five tissues of Zhikong scallop, *Chlamys farreri*. PLoS ONE 9:e86232.

Takeuchi, T., T. Kawashima, R. Koyanagi, F. Gyoja, M. Tanaka, T. Ikuta, E. Shoguchi, M. Fujiwara, C. Shinzato, K. Hisata, M. Fujie, T. Usami, K. Nagai, K. Maeyama, K. Okamoto, H. Aoki, T. Ishikawa, T. Masaoka, A. Fujiwara, K. Endo, H. Endo, H. Nagasawa, S. Kinoshita, S. Asakawa, S. Watabe, and N. Satoh. 2012. Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. DNA Res 19:117-130.

Talbert, P. B. and S. Henikoff. 2010. Histone variants--ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol 11:264-275.

Talbert, P. B. and S. Henikoff. 2014. Environmental responses mediated by histone variants. Trends Cell Biol.

Tanguy, A., I. Boutet, J. Laroche, and D. Moraga. 2005. Molecular identification and expression study of differentially regulated genes in the Pacific oyster *Crassostrea gigas* in response to pesticide exposure. FEBS J 272:390-403.

Valdiglesias, V., J. Fernandez-Tajes, E. Pasaro, J. Mendez, and B. Laffon. 2012. Identification of differentially expressed genes in SHSY5Y cells exposed to okadaic acid by suppression subtractive hybridization. BMC Genomics 13:46.

Valdiglesias, V., B. Laffon, E. Pasaro, and J. Mendez. 2011. Okadaic acid induces morphological changes, apoptosis and cell cycle alterations in different human cell types. J Environ Monit 13:1831-1840.

Van Dolah, F. M. 2000. Marine Algal Toxins: Origins, Health Effects, and Their Increased Occurrence. Environ Health Persp 108:133-141.

Venier, P., C. De Pitta, F. Bernante, L. Varotto, B. De Nardi, G. Bovo, P. Roch, B. Novoa, A. Figueras, A. Pallavicini, and G. Lanfranchi. 2009. MytiBase: a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics 10:72.

Venier, P., C. De Pitta, A. Pallavicini, F. Marsano, L. Varotto, C. Romualdi, F. Dondero, A. Viarengo, and G. Lanfranchi. 2006. Development of mussel mRNA profiling: Can gene expression trends reveal coastal water pollution? Mutat Res 602:121-134.

Venier, P., L. Varotto, U. Rosani, C. Millino, B. Celegato, F. Bernante, G. Lanfranchi, B. Novoa, P. Roch, A. Figueras, and A. Pallavicini. 2011. Insights into the innate immunity of the Mediterranean mussel *Mytilus galloprovincialis*. BMC Genomics 12:69.

Wells, P. G., M. H. Depledge, J. N. Butler, J. J. Manock, and A. H. Knap. 2001. Rapid toxicity assessment and biomonitoring of marine contaminants - exploiting the potential of rapid biomarker assays and microscale toxicity tests. Mar. Pollut. Bull. 42:799-804.

Williams, T. D., L. Mirbahai, and J. K. Chipman. 2014. The toxicological application of transcriptomics and epigenomics in zebrafish and other teleosts. Brief Funct Genomics 13:157-171.

Yue, X., H. Wang, X. Huang, C. Wang, X. Chai, and B. Liu. 2012. Single nucleotide polymorphisms in i-type lysozyme gene and their correlation with *vibrio*-resistance and growth of clam *Meretrix meretrix* based on the selected resistance stocks. Fish Shellfish Immunol 33:559-568.

Zapata, M., A. Tanguy, E. David, D. Moraga, and C. Riquelme. 2009. Transcriptomic response of *Argopecten purpuratus* post-larvae to copper exposure under experimental conditions. Gene 442:37-46.

Zhang, G., X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que, Y. Xie, P. W. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, J. Meng, L. Yang, J. Liu, B. Wen, N. Zhang, Z. Huang, Q. Zhu, Y. Feng, A. Mount, D. Hedgecock, Z. Xu, Y. Liu, T. Domazet-Loso, Y. Du, X. Sun, S. Zhang, B. Liu, P. Cheng, X. Jiang, J. Li, D. Fan, W. Wang, W. Fu, T. Wang, B. Wang, J. Zhang, Z. Peng, Y. Li, N. Li, M. Chen, Y. He, F. Tan, X. Song, Q. Zheng, R. Huang, H. Yang, X. Du, L. Chen, M. Yang, P. M. Gaffney, S. Wang, L. Luo, Z. She, Y. Ming, W. Huang, B. Huang, Y. Zhang, T. Qu, P. Ni, G. Miao, Q. Wang, C. E. Steinberg, H. Wang, L. Qian, X. Liu, and Y. Yin. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49-54.

Figure 1: Environmental changes require swift epigenetic responses to genotoxic stress (e.g., DNA breaks) including the onset of epigenetic modifications triggering the remodeling of the chromatin fiber and modulating the access to specific genes involved in the response to DNA damage. Once the stress episode is over most of these marks will be reset, reverting chromatin structure to its basal state. Nonetheless, many of these epigenetic marks will transcend throughout generations in those cases where the environmental stress persists, securing a continuous response to genotoxicity in the cell and establishing the basis for organismal long-term adaptations.

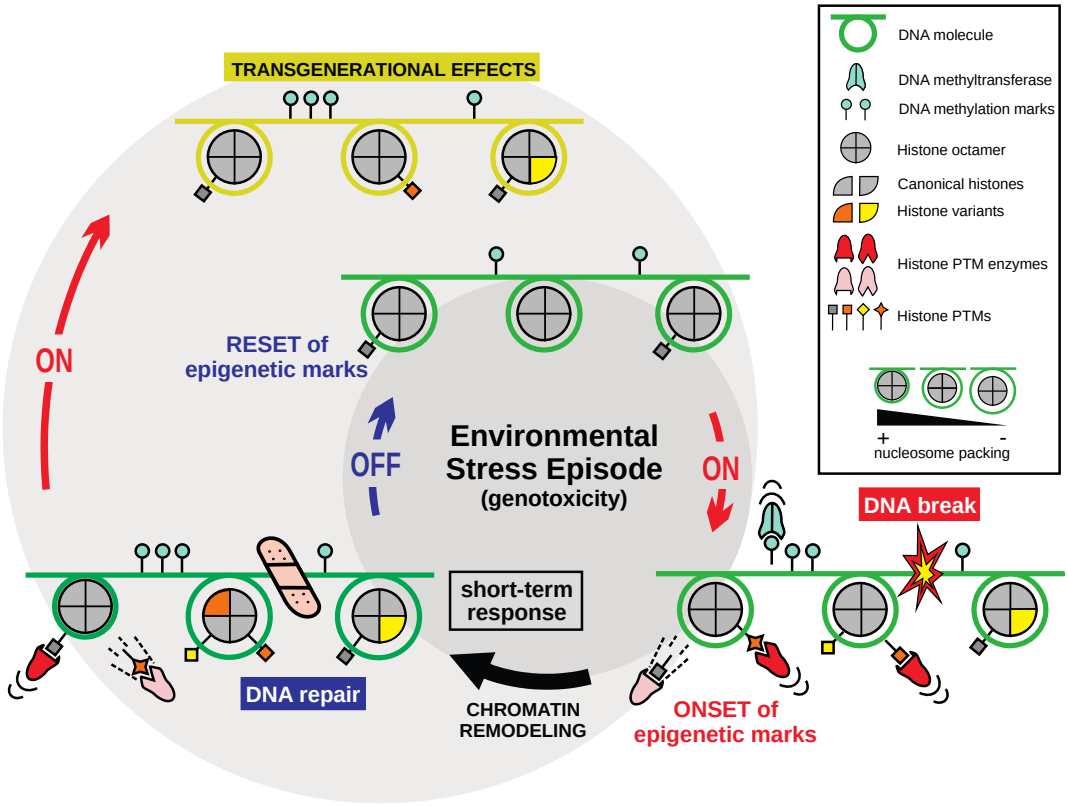


Figure 2: Increasing numbers of registered Bioprojects (2013-2017) in NCBI involving gene expression studies in bivalves classified by species.

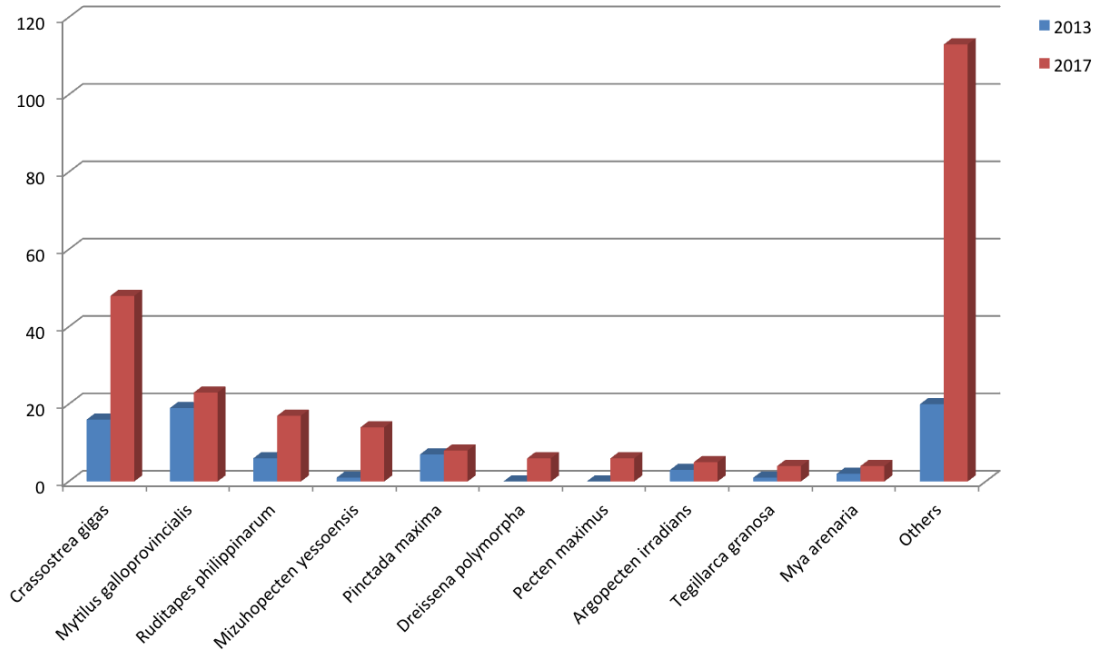


Table 1: Species-specific databases of ESTs and transcriptomic data.

Database	***Organism*/# Sequences**	***Tissues***
Species-centered		
Mytibase	*M. galloprovincialis* 7112	*Digestive gland*, *gills*, *hemocytes*
GigasDatabase	*C. gigas* 29745	*Digestive gland*, *gills*, *gonad*, *hemocytes*, *mantle-edge*, *muscle*
RuphiBase	*R. philippinarum* 32606	*Mixed tissues*
ChameleaBase	*C. gallina* 36422	*Muscle*
DeepSeaVent	*B. azoricus* 35903	*Gills*
Functionally-centered		
Chromevaloa	*M. galloprovincialis* 14408	*Digestive gland*

CHAPTER II

THE CHROMEVALOA DATABASE: A RESOURCE FOR THE EVALUATION OF OKADAIC ACID CONTAMINATION IN THE MARINE ENVIRONMENT BASED ON THE CHROMATIN-ASSOCIATED TRANSCRIPTOME OF THE MUSSEL *MYTILUS GALLOPROVINCIALIS*

In this chapter, the transcriptome of the mussel *Mytilus galloprovincialis* was characterized under conditions of exposure to low concentrations of the toxic dinoflagellate *Prorocentrum Lima*, leading to the identification and classification of genes with epigenetic relevance, i.e., those coding for chromatin-associated proteins. Gene sequences and related functional information were organized and catalogued in a public repository named CHROMEVALOAdb, constituting a specialized data resource for the study of epigenetic responses of bivalves to Harmful Algal Blooms.

This work was carried out with the collaboration of Juan Fernandez Tajes[1,2], Vanessa Aguiar-Pulido[2], Ciro Rivera-Casas[2], Rodrigo Gonzalez-Romero[2,3], Juan Ausio[3], Josefina Mendez[2], Julian Dorado[2] and Jose M. Eirin-Lopez[4], under affiliations to the University of Oxford (U.K.) [1], University of A Coruña (Spain)[2], University of Victoria (Canada)[3], and Florida International University (U.S.A.) [4].

Abstract

Okadaic Acid (OA) constitutes the main active principle in Diarrhetic Shellfish Poisoning (DSP) toxins produced during Harmful Algal Blooms (HABs), constituting a serious threat for human consumers of edible shellfish. Furthermore, OA conveys critical deleterious effects for marine organisms as well due to its genotoxic potential. Many efforts have been dedicated to OA biomonitoring during the last three decades. However, it is only now with the current availability of detailed molecular information on DNA organization and the mechanisms involved in the maintenance of genome integrity that a new arena starts opening up for the study of OA contamination. In the present work we

address the links between OA genotoxicity and chromatin by combining Next Generation Sequencing (NGS) technologies and bioinformatics. To this end, we introduce CHROMEVALOAdb (http://chromevaloa.udc.es), a public database containing the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis* (a sentinel model organism) in response to OA exposure. This resource constitutes a leap forward for the development of chromatin-based biomarkers, paving the road towards the generation of powerful and sensitive tests for the detection and evaluation of the genotoxic effects of OA in coastal areas.

Introduction

Massive algal proliferations are among the most important sources of contamination in the sea. These episodes may arise as a consequence of either natural or anthropogenic causes, leading to large accumulations of algae in the marine environment (Cardozo et al. 2007). Quite often, massive algal proliferations include blooms of toxin-producing organisms known as Harmful Algal Blooms (HABs), producing high concentrations of potentially harmful biotoxins that are accumulated throughout the food chain. Among HAB biotoxins, Diarrhetic Shellfish Poisoning (DSP) toxins are especially predominant across European coasts, causing alterations in the gastrointestinal system of human consumers of contaminated shellfish (Aune and Yndestad 1993, James et al. 2010). The main active principle in DSPs is Okadaic Acid (OA) (Vale 2010), which is synthesized by dinoflagellates of the genera *Dinophysis* and *Prorocentrum* (Yasumoto et al. 1980). OA has genotoxic potential, constituting a tumor promoter and apoptosis inducer able to cause DNA oxidative damage (Suganuma et al. 1988, Leira et al. 2001).

Particularly, DNA Double Strand Breaks (DSBs) stand out for their severity among the genotoxic effects exerted by OA and require the activation of prompt repair mechanisms in order to avoid serious damage in the cell (Florez-Barros et al. 2011, Valdiglesias et al. 2011).

During the last 30 years, fisheries and aquaculture-based industries have experienced important economic losses. This is visibly due to the dramatic increase in the diversity of toxic algal species and the toxins they produce (Van Dolah and Ramsdell 1992), which constitute a serious threat for human consumers (Cardozo et al. 2007). Consequently, a very important effort has been devoted to OA biomonitoring in estuarine areas by using sentinel organisms, most notably bivalve molluscs (Wells et al. 2001, Florez-Barros et al. 2011). These studies have progressively transitioned from traditional biomonitoring methods (based on physicochemical and physiological parameters) to more sensitive molecular probes (Marcaillou-Le Baut et al. 1994, Manfrin et al. 2010, Ledreux et al. 2012, Sassolas et al. 2012). Given the role of chromosomal proteins in the modulation of chromatin structure and DNA metabolism (including DNA repair) (Dinant et al. 2008), the study of chromatin-associated biomarkers constitutes a powerful and sensitive approach for the evaluation of genotoxicity. The usefulness of chromatin-based genotoxicity tests has already been demonstrated in mammals, where histone H2A.X phosphorylation has been used to assess the extent of DNA repair following exposure of cells to DNA-damaging agents (Albino et al. 2009, Dickey et al. 2009, Watters et al. 2009). Yet, this approach is largely unexplored in those organisms where chromatin information is scarce, including bivalve molluscs (González-Romero et al. 2012a). Furthermore, the lack of knowledge regarding gene and protein sequences in these

organisms constitutes a very important barrier for the analysis of high-throughput -omic data, especially as it pertains to data assembly and annotation of highly divergent and/or lineage-specialized genes (Eirín-López et al. 2004, Eirín-López et al. 2006, González-Romero et al. 2012a, González-Romero et al. 2012b). Even though the genome sequence of the Pacific oyster *Crassostrea gigas* has been recently published (Zhang et al. 2012), the amount of information available for marine bivalves remains scarce compared to other model organisms in spite of their environmental value.

In the present work, we specifically address the links between OA genotoxicity and potential chromatin-associated biomarkers by combining Next Generation Sequencing (NGS) technologies and bioinformatics. To this end, we introduce CHROMEVALOAdb (chromevaloa.udc.es), a database containing the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis* in response to OA exposure. The information provided in this database includes fully traceable raw ESTs assembled into consensus sequences and classified into unigenes linked to Gene Ontology (GO) information (function, process and subcellular compartment) as well as to expression information in response to OA. CHROMEVALOAdb allows for the manual browsing and keyword-based search of chromatin-associated contigs. In addition, the whole OA-specific transcriptome can be accessed using built in BLAST and CLUSTAL W tools. Overall, the present work constitutes a leap forward in the study of the genotoxic effect exerted by OA in these organisms, paving the road towards the development of chromatin-based tests for detecting and evaluating the genotoxic effect of OA in the marine environment.

Results and discussion

Sequencing and annotation of OA-specific ESTs in M. galloprovincialis

Mussels (*M. galloprovincialis*) sampled in the Galician coast (northwest Spain), obtained from an area with a low impact of dinoflagellates blooms, were experimentally exposed to OA in the laboratory (Figure 1) using a set of conditions that were previously proven to cause significant genotoxic damage (200 cells/L of the OA-producing dinoflagellate *Prorocentrum lima*, 1 day exposure) (Carvalho Pinto-Silva et al. 2005, Florez-Barros et al. 2011). The accumulation of OA in digestive gland tissue was subsequently confirmed by HPLC-MS quantification (Table 1).

Raw normalized libraries constructed from mussel specimens exposed and non-exposed to OA were sequenced using pyrosequencing technology at 40x depth, producing 493,440 and 491,109 raw reads for the OA-exposed (NORM_MGC) and the control (NORM_MGT) libraries, respectively. These data allowed the assembly of 16,395 consensus sequences in the case of the control library and 24,624 consensus sequences from the OA-exposed library, with average length values of 712 and 644 bp, respectively. Approximately 44% of the assembled sequences (17,952) were annotated by using BLAST (blastx) homology searches against non-redundant (nr) protein databases, including 7335 contigs in the control library and 10,617 contigs in the OA-exposed library (38% and 45%, respectively), setting an expectation (*e*) value of $1 \times 10{-}6$ or better (Table 2).

Novel chromatin-associated transcripts in CHROMEVALOAdb

Chromatin-associated transcripts were identified from the assembled OA-specific transcriptome from *M. galloprovincialis* by following two complementary strategies (see Experimental Section for details). On one hand, a list of keywords identifying chromatin-associated components was used to screen annotated transcripts regarding sequence description and related gene ontology terms (Supplementary Figures S1 and S2). On the other hand, BLAST homology comparisons were performed against specialized chromatin databases. The combination of both strategies resulted in the identification of 14,480 chromatin-associated contigs in control and OA-exposed libraries among which 1124 were identified as chromatin-associated unigenes (Table 2). The analysis of gene expression profiles (Supplementary Figure S3) allowed us to define groups of statistically significant unigenes upregulated and downregulated in the presence of OA (a total number of 1254) among which 90 were identified as chromatin-associated (Table 2). This information, along with gene ontology and expression profile data, constitutes the core of CHROMEVALOAdb.

The ontological analysis of the biological processes on which the identified chromatin-associated unigenes could be potentially involved revealed that cellular and metabolic processes are most significantly deregulated in response to OA (Figure 2). Furthermore, a significant deregulation of genes involved in chromatin remodeling (inhibited) and transmembrane transport (overexpressed) was identified through global ontological analyses based on the whole OA-specific transcriptome (Fisher's exact test approach using topGO R-bioconductor package, Supplementary Figure S4). Even though additional experimental studies will be needed to decipher the functional role of

chromatin-associated unigenes in response to OA, these results may be indicative of the activation in protective detoxifying mechanisms in mussels after one day of exposure to OA, once DNA has been repaired.

Comparisons between OA-specific EST information from CHROMEVALOAdb and *Mytilus* ESTs information from the MytiBase EST knowledge database (Venier et al. 2009) revealed that approximately 25% of the chromatin-associated sequences contained in CHROMEVALOAdb are redundant with MytiBase sequences. This extends also to the case of the complete OA-specific transcriptome, with a 30% of the ESTs being redundant with MytiBase sequences considering no identity cutoff value. In other words, approximately 75% of the ESTs contained in CHROMEVALOAdb constitute previously unknown transcripts in the mussel *M. galloprovincialis*, establishing a very important contribution not only for the study of OA chromatin-associated biomarkers, but also for the characterization of the mussel genome.

Availability, management and application of data stored in CHROMEVALOAdb

Management of data quality constitutes a basic requirement of NGS projects that is often overlooked, resulting in the loss of important information for fine sequence curation and identification of DNA polymorphisms, among other quantitative analyses. The structure of CHROMEVALOAdb strengthens this aspect by providing full access to raw reads used to assemble the consensus sequences annotated in the database. This feature facilitates the alignment of quality-filtered raw sequences, establishing links with specific expression patterns in response to OA. Furthermore, the availability of the full dataset of contigs allows users to retrieve anonymous sequences by using the BLAST

tool interface and communicate new chromatin-related findings through a standardized feedback form, contributing to the curation of the information in CHROMEVALOAdb. Processed data, on the other hand, is also downloadable as flat text files containing information that can be filtered by keywords (Figure 3).

The information contained in CHROMEVALOAdb serves a dual purpose. First, it helps identify previously unknown chromatin-associated factors in the mussel *M. galloprovincialis*, specially histone variants and chromatin remodeling factors (Figure 4A,B). This aim is motivated by the role of chromatin-associated proteins in the maintenance of genome integrity, most notably in the case of DNA DSB repair (González-Romero et al. 2012a, González-Romero et al. 2012b). Within this context, the generation of new molecular data and its organization in CHROMEVALOAdb helps increase the knowledge about chromatin in molluscs, setting up a framework for studying its role in DNA repair. The second purpose of CHROMEVALOAdb is to establish cause-effect relationships between OA exposure and specific expression patterns of chromatin-associated factors involved in the maintenance of genome integrity. Given the critical roles played by these factors during DNA DSB repair, the screening of their expression patterns will help identify potentially sensitive biomarkers of OA genotoxic effect. To this end, CHROMEVALOAdb provides differential expression information for chromatin-associated unigenes, using an intuitive and practical format based on arrows (up-regulated and down-regulated transcripts, Figure 4C). The combination of the newly characterized DNA sequences together with their associated expression information in response to OA paves the road towards the development of chromatin-based tests for detecting and evaluating the genotoxic effect of OA in the marine environment.

Methods

Synthesis of ESTs libraries and transcriptome assembly

Mussel specimens (*M. galloprovincialis*) were sampled in Valcobo beach, Galicia (northwest coast of Spain, 43°19'02.71"N 8°21'56.35"W) and immediately transported to the laboratory thereafter where they were maintained under controlled light/temperature conditions and fed with a standard mixture of the microalgae *Isochrysis galbana* and *Tetraselmis suecica* (Figure 1). Individuals were subsequently divided into a control group and a group exposed to OA that was additionally fed with a culture of the DSP-producing microalgae *Prorocentrum lima* (200 cells/L for 24 hours). Extraction of mRNA was subsequently performed from pooled digestive gland tissue (hepatopancreas) from five individuals in each group. The choice of this tissue as mRNA source is motivated by its ability to accumulate the biotoxin in large amounts and its detoxifying role in mussel metabolism (Svensson 2003).

cDNA libraries were synthesized using the SMARTerTM PCR cDNA synthesis kit (Clontech, Mountain View, CA) with an extra purification step using GeneJET^{TM} PCR Purification Kit (Thermo Scientific, Waltham, MA), and normalization was performed following the protocol of the Trimer cDNA Normalization Kit (Evrogen, Moscow, Russia). Libraries were sequenced using Roche-454 FLX+ Titanium pyrosequencing, obtaining forward (exposed) and reverse (control) datasets. Reads from both libraries were pre-processed (quality filtering and contaminantion removal) by combining the CD-HIT-454 (Niu et al. 2010) and the BLAST+ software (Altschul et al. 1990) implemented in the SeqtrimNext pipeline (Falgueras et al. 2010), as well as the Cutadapt v1.0 software (Martin 2011). Sequence assembly was carried out using MIRA

v.3.4.0 sequence assembler (Chevreux et al. 1999). The sequences described in this work are available at the Sequence Read Archive (SRA) database under the accession number SRA056210.

Database contents, accessibility and tool implementation

The relational structure of CHROMEVALOAdb was developed using MySQL, allowing full traceability of raw ESTs from consensus sequences of individual genes. Contigs are classified into unigenes to eliminate redundancy based on BLAST analysis parameters (same top blast hit, mean similarity larger than 80% and an e-value below 1e-10), resulting in a total number of 2,131 unigenes among which 1,124 were identified as chromatin-related (Table 2). The descriptions of the unigenes are linked to their corresponding contigs and to ontology annotations. All the information stored in CHROMEVALOAdb is freely available for browsing and downloading without login or registering requirements. The information gathered by CHROMEVALOAdb is managed through Perl-written Common Gateway Interfaces (CGIs) that communicate with the Relational Database Management System (RDBMS) MySQL using Perl's database interface (DBI) module. Server-side tools for sequence alignment, data visualization and result formatting/retrieval are administered by built in HTML web interfaces. BLAST results are formatted and interactively presented in HTML format including graphics, using Bioperl packages. Multiple sequence alignments are generated using CLUSTAL W (Thompson et al. 1994) and displayed with an embedded applet of the alignment editor Jalview (Clamp et al. 2004, Waterhouse et al. 2009).

Local data is linked to reference public databases such as NCBI repositories for extended homolog sequence descriptions and AmiGO (Carbon et al. 2009) for gene ontology term definitions.

Gene annotation and expression analysis

The functional annotation of the consensus read assemblies was carried out using the Blast2GO suite (Gotz et al. 2008), combining Gene Ontology (GO), InterProScan (IPS) protein domain information (Zdobnov and Apweiler 2001) and annotation enrichment using ANNEX (Myhre et al. 2006). Additionally, full-length transcripts were subsequently identified using the Full-Lengther tool (Lara et al. 2007). Identification of chromatin-associated transcripts was subsequently implemented following two complementary strategies. First, a keyword-based routine was defined to identify chromatin-associated transcripts among sequence descriptions and related ontology terms (Supplementary Material S2). Secondly, BLAST (blastn and blastx) homology searches were performed against the Histone Database (Marino-Ramirez et al. 2006), as well as against ChromDB (Gendler et al. 2008) and CREMOFAC (Shipra et al. 2006) databases, setting an e-value threshold of 1e-10. Functionally annotated and classified sequences, along with relevant metadata, are organized and stored in CHROMEVALOAdb.

The biological processes on which the identified chromatin-associated unigenes could be potentially involved were studied by performing ontological analyses based on GO terms (Supplementary Figure S3). Expression profiles in response to OA were further studied by comparing control and OA-exposed libraries, using the edgeR package from R-Bioconductor (Robinson et al. 2010) with the False Discovery Rate (FDR) threshold

set to 0.1 (Supplementary Figure S4). Read count for each assembled sequence was performed using SQL-based queries on the raw data contained in CHROMEVALOAdb. This approach allowed us to define groups of statistically significant unigenes upregulated and downregulated in the presence of OA.

Conclusions

CHROMEVALOAdb provides a powerful resource to investigate the molecular basis underlying the genotoxic effect of OA in mussels and for understanding the chromatin-associated mechanisms that counteract the harmful effect of this toxin in these organisms (i.e., mechanisms involved in DNA repair). Furthermore, it allows the establishment of cause-effect relationships between OA and the differential expression of chromatin-associated factors involved in DNA DSB repair, helping to identify potential sensitive biomarkers for the development of chromatin-based OA genotoxicity tests. The implementation of these tests in natural populations has critical implications for the evaluation of DNA damage in commercially relevant organisms, the optimization of their harvesting and the elaboration of additional tests designed to evaluate the safety of their consumption and potential implications for consumer's health. The design of CHROMEVALOAdb sets the basis for the future integration of model-based and semi-automated curation systems. In addition, the characterization of additional transcriptomes (i.e., at different stages of the genotoxic stress and in different tissues), together with data integration and workflow automation for interactome network development, constitute future objectives for the improvement of the database. Altogether, these approaches will help increase the knowledge of the chromatin-associated mechanisms involved in the

response to the genotoxic effect of OA, by using Knowledge Discovery in Databases (KDD) techniques.

References

Albino, A. P., E. D. Jorgensen, P. Rainey, G. Gillman, T. J. Clark, H. Zhao, F. Traganos, and Z. Darzynkiewicz. 2009. gamma-H2AX: a potential DNA damage response biomarker for assessing toxicological risk of tobacco products. Mutat. Res. 678:43-52.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Aune, T. and M. Yndestad. 1993. Diarrhetic shellfish poisoning. Pages 87-104 *in* I. R. Falconer, editor. Algal toxins in seafood and drinking water. Academic Press, London.

Carbon, S., A. Ireland, C. J. Mungall, S. Shu, B. Marshall, and S. Lewis. 2009. AmiGO: online access to ontology and annotation data. Bioinformatics 25:288-289.

Cardozo, K. H., T. Guaratini, M. P. Barros, V. R. Falcao, A. P. Tonon, N. P. Lopes, S. Campos, M. A. Torres, A. O. Souza, P. Colepicolo, and E. Pinto. 2007. Metabolites from algae with economical impact. Comp Biochem Physiol C Toxicol Pharmacol 146:60-78.

Carvalho Pinto-Silva, C. R., E. E. Creppy, and W. G. Matias. 2005. Micronucleus test in mussels *Perna perna* fed with the toxic dinoflagellate *Prorocentrum lima*. Arch Toxicol 79:422-426.

Chevreux, B., T. Wetter, and S. Suhai. 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. Pages 45-56 *in* Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB).

Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton. 2004. The Jalview Java alignment editor. Bioinformatics 20:426-427.

Dickey, J. S., C. E. Redon, A. J. Nakamura, B. J. Baird, O. A. Sedelnikova, and W. Bonner. 2009. H2AX: Functional roles and potential applications. Chromosoma 118:683-695.

Dinant, C., A. B. Houtsmuller, and W. Vermeulen. 2008. Chromatin structure and DNA damage repair. Epigenetics Chromatin 1:9.

Eirín-López, J. M., J. D. Lewis, L. Howe, and J. Ausió. 2006. Common phylogenetic origin of protamine-like (PL) proteins and histone H1: evidence from bivalve PL genes. Mol Biol Evol 23:1304-1317.

Eirín-López, J. M., M. F. Ruiz, A. M. González-Tizón, A. Martínez, L. Sánchez, and J. Méndez. 2004. Molecular evolutionary characterization of the mussel *Mytilus* histone multigene family: first record of a tandemly repeated unit of five histone genes containing an H1 subtype with "orphon" features. J Mol Evol 58:131-144.

Falgueras, J., A. J. Lara, N. Fernandez-Pozo, F. R. Canton, G. Perez-Trabado, and M. G. Claros. 2010. SeqTrim: a high-throughput pipeline for pre-processing any type of sequence read. BMC Bioinformatics 11:38.

Florez-Barros, F., M. Prado-Alvarez, J. Mendez, and J. Fernandez-Tajes. 2011. Evaluation of genotoxicity in gills and hemolymph of clam *Ruditapes decussatus* fed with the toxic dinoflagellate *Prorocentrum lima*. J Toxicol Environ Health A 74:971-979.

Gendler, K., T. Paulsen, and C. Napoli. 2008. ChromDB: the chromatin database. Nucleic Acids Res 36:D298-302.

González-Romero, R., C. Rivera-Casas, J. Fernandez-Tajes, J. Ausio, J. Méndez, and J. M. Eirín-López. 2012a. Chromatin specialization in bivalve molluscs: A leap forward for the evaluation of okadaic acid genotoxicity in the marine environment. Comp Biochem Physiol C Toxicol Pharmacol 155:175-181.

González-Romero, R., C. Rivera-Casas, L. J. Frehlick, J. Méndez, J. Ausió, and J. M. Eirín-López. 2012b. Histone H2A (H2A.X and H2A.Z) variants in molluscs: molecular characterization and potential implications for chromatin dynamics. PLoS ONE 7:e30006.

Gotz, S., J. M. Garcia-Gomez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talon, J. Dopazo, and A. Conesa. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420-3435.

James, K. J., B. Carey, J. O'Halloran, F. N. van Pelt, and Z. Skrabakova. 2010. Shellfish toxicity: human health implications of marine algal toxins. Epidemiol Infect 138:927-940.

Lara, A. J., G. Perez-Trabado, D. P. Villalobos, S. Diaz-Moreno, F. R. Canton, and M. G. Claros. 2007. A Web Tool to Discover Full-Length Sequences: Full-Lengther.*in* E. Corchado, J. M. Corchado, and A. Abraham, editors. Innovations in Hibrid Intelligent Systems. Springer, Berlin.

Ledreux, A., A. L. Serandour, B. Morin, S. Derick, R. Lanceleur, S. Hamlaoui, C. Furger, R. Bire, S. Krys, V. Fessard, M. Troussellier, and C. Bernard. 2012. Collaborative study for the detection of toxic compounds in shellfish extracts using cell-based assays. Part II: application to shellfish extracts spiked with lipophilic marine toxins. Anal Bioanal Chem 403:1995-2007.

Leira, F., C. Alvarez, J. M. Vieites, M. R. Vieytes, and L. M. Botana. 2001. Study of cytoskeletal changes induced by okadaic acid in BE(2)-M17 cells by means of a quantitative fluorimetric microplate assay. Toxicol In Vitro 15:277-282.

Manfrin, C., R. Dreos, S. Battistella, A. Beran, M. Gerdol, L. Varotto, G. Lanfranchi, P. Venier, and A. Pallavicini. 2010. Mediterranean mussel gene expression profile induced by okadaic acid exposure. Environ Sci Technol 44:8276-8283.

Marcaillou-Le Baut, C., Z. Amzil, J. P. Vernoux, Y. F. Pouchus, M. Bohec, and J. F. Simon. 1994. Studies on the detection of okadaic acid in mussels: preliminary comparison of bioassays. Nat Toxins 2:312-317.

Marino-Ramirez, L., B. Hsu, A. D. Baxevanis, and D. Landsman. 2006. The Histone Database: a comprehensive resource for histones and histone fold-containing proteins. Proteins 62:838-842.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10-12.

Myhre, S., H. Tveit, T. Mollestad, and A. Laegreid. 2006. Additional gene ontology structure for improved biological reasoning. Bioinformatics 22:2020-2027.

Niu, B., L. Fu, S. Sun, and W. Li. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. BMC Bioinformatics 11:187.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Sassolas, A., A. Hayat, G. Catanante, and J.-L. Marty. 2012. Detection of the marine toxin okadaic acid: assessing seafood safety. Talanta in press.

Shipra, A., K. Chetan, and M. R. Rao. 2006. CREMOFAC--a database of chromatin remodeling factors. Bioinformatics 22:2940-2944.

Suganuma, M., H. Fujiki, H. Suguri, S. Yoshizawa, M. Hirota, M. Nakayasu, M. Ojika, K. Wakamatsu, K. Yamada, and T. Sugimura. 1988. Okadaic acid: an additional non-phorbol-12-tetradecanoate-13-acetate-type tumor promoter. Proc. Natl. Acad. Sci. U S A 85:1768-1771.

Svensson, S. 2003. Effects, dynamics and management of okadaic acid in blue mussels, *Mytilus edulis*. Doctoral Dissertation. Göteborg University, Strömstad.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. Nucl. Acids Res. 22:4673-4680.

Valdiglesias, V., B. Laffon, E. Pasaro, and J. Mendez. 2011. Evaluation of okadaic acid-induced genotoxicity in human cells using the micronucleus test and gammaH2AX analysis. J Toxicol Environ Health A 74:980-992.

Vale, P. 2010. Profiles of fatty acids and 7-O-acyl okadaic acid esters in bivalves: can bacteria be involved in acyl esterification of okadaic acid? Comp Biochem Physiol C Toxicol Pharmacol 151:18-24.

Van Dolah, F. M. and J. S. Ramsdell. 1992. Okadaic acid inhibits a protein phosphatase activity involved in formation of the mitotic spindle of GH4 rat pituitary cells. J. Cell Physiol. 151:190-198.

Venier, P., C. De Pitta, F. Bernante, L. Varotto, B. De Nardi, G. Bovo, P. Roch, B. Novoa, A. Figueras, A. Pallavicini, and G. Lanfranchi. 2009. MytiBase: a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. BMC Genomics 10:72.

Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp, and G. J. Barton. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189-1191.

Watters, G. P., D. J. Smart, J. S. Harvey, and C. A. Austin. 2009. H2AX phosphorylation as a genotoxicity endpoint. Mutat. Res. 679:50-58.

Wells, P. G., M. H. Depledge, J. N. Butler, J. J. Manock, and A. H. Knap. 2001. Rapid toxicity assessment and biomonitoring of marine contaminants--exploiting the potential of rapid biomarker assays and microscale toxicity tests. Mar. Pollut. Bull. 42:799-804.

Yasumoto, T., Y. Oshima, W. Sugawara, Y. Fukuyo, H. Oguri, T. Igarashi, and N. Fujita. 1980. Identification of *Dinophysis fortii* as the causative organism of diarrhetic shellfish poisoning. Bulletin of the Japanese Society of Scientific Fisheries 46:1405-1411.

Zdobnov, E. M. and R. Apweiler. 2001. InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847-848.

Zhang, G., X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que, Y. Xie, P. W. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, J. Meng, L. Yang, J. Liu, B. Wen, N. Zhang, Z. Huang, Q. Zhu, Y. Feng, A. Mount, D. Hedgecock, Z. Xu, Y. Liu, T. Domazet-Loso, Y. Du, X. Sun, S. Zhang, B. Liu, P. Cheng, X. Jiang, J. Li, D. Fan, W. Wang, W. Fu, T. Wang, B. Wang, J. Zhang, Z. Peng, Y. Li, N. Li, M. Chen, Y. He, F. Tan, X. Song, Q. Zheng, R. Huang, H. Yang, X. Du, L. Chen, M. Yang, P. M. Gaffney, S. Wang, L. Luo, Z. She, Y. Ming, W. Huang, B. Huang, Y. Zhang, T. Qu, P. Ni, G. Miao, Q. Wang, C. E. Steinberg, H. Wang, L. Qian, X. Liu, and Y. Yin. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49-54.

Figure 1. Experimental settings for the exposure of mussels to Okadaic Acid (OA), specifying the environmental conditions for treated (additionally fed with OA-producing microalgae *P. lima*) and control groups of mussel individuals.

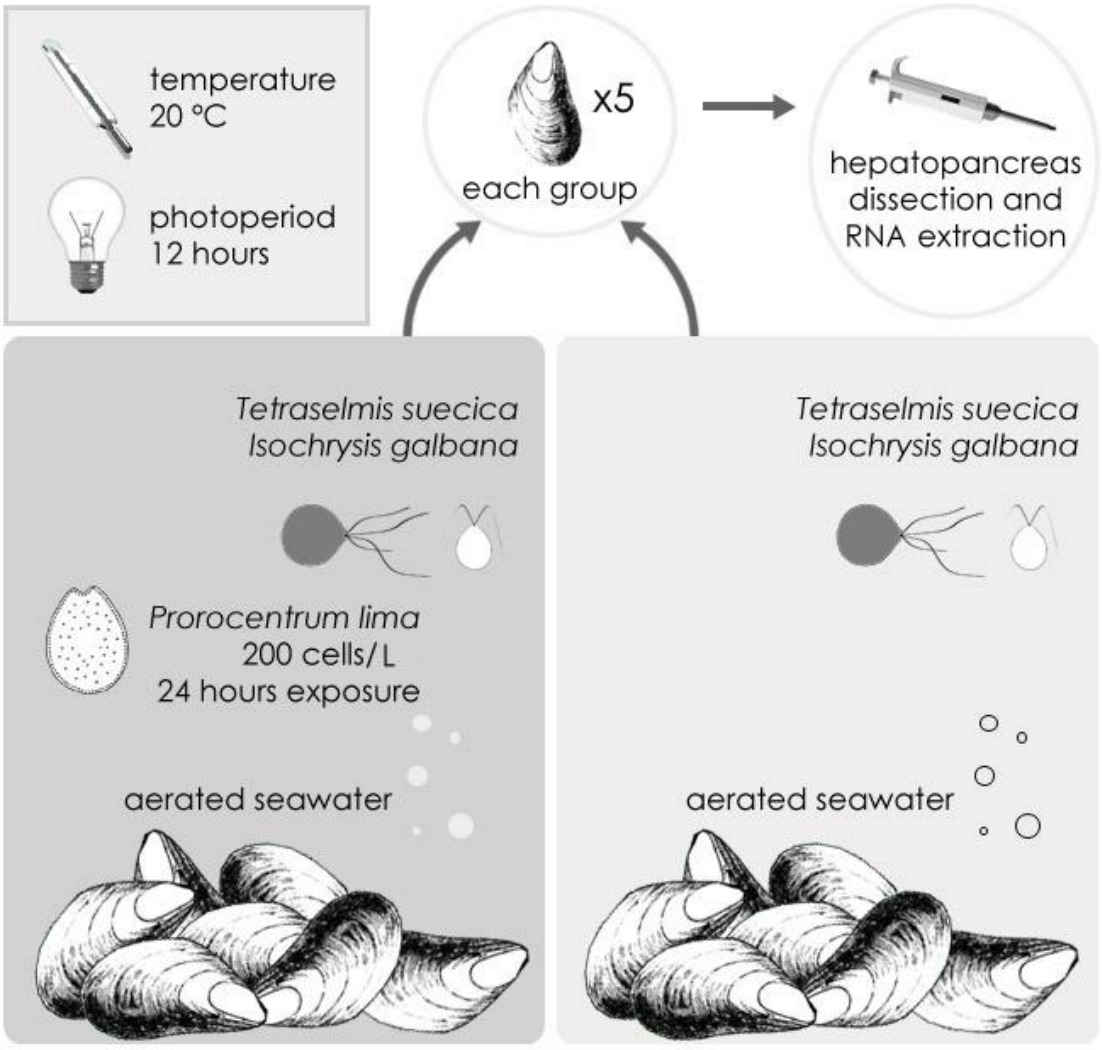


Figure 2. Biological processes on which chromatin-associated unigenes could be potentially involved during the response to OA.

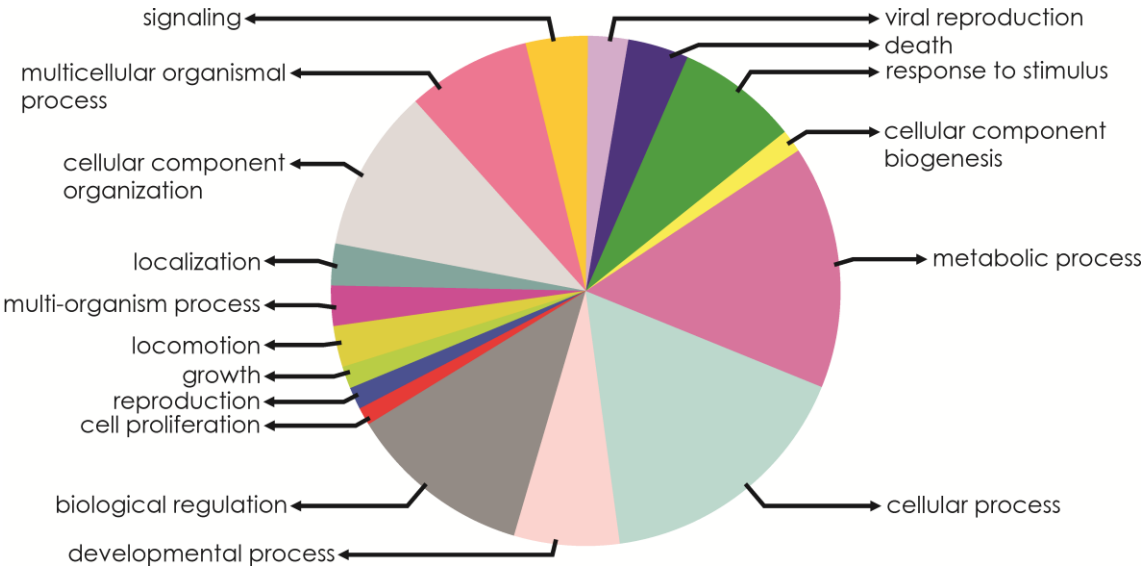


Figure 3. Diagram showing the pipeline of data management in CHROMEVALOAdb. Starting from files containing the fully annotated transcript libraries, the selection of chromatin-associated sequences is carried out through semantic and homology search approaches. Sequences and annotations are organized in the relational structures of the database and made available through web interface, including data retrieval and feedback utilities.

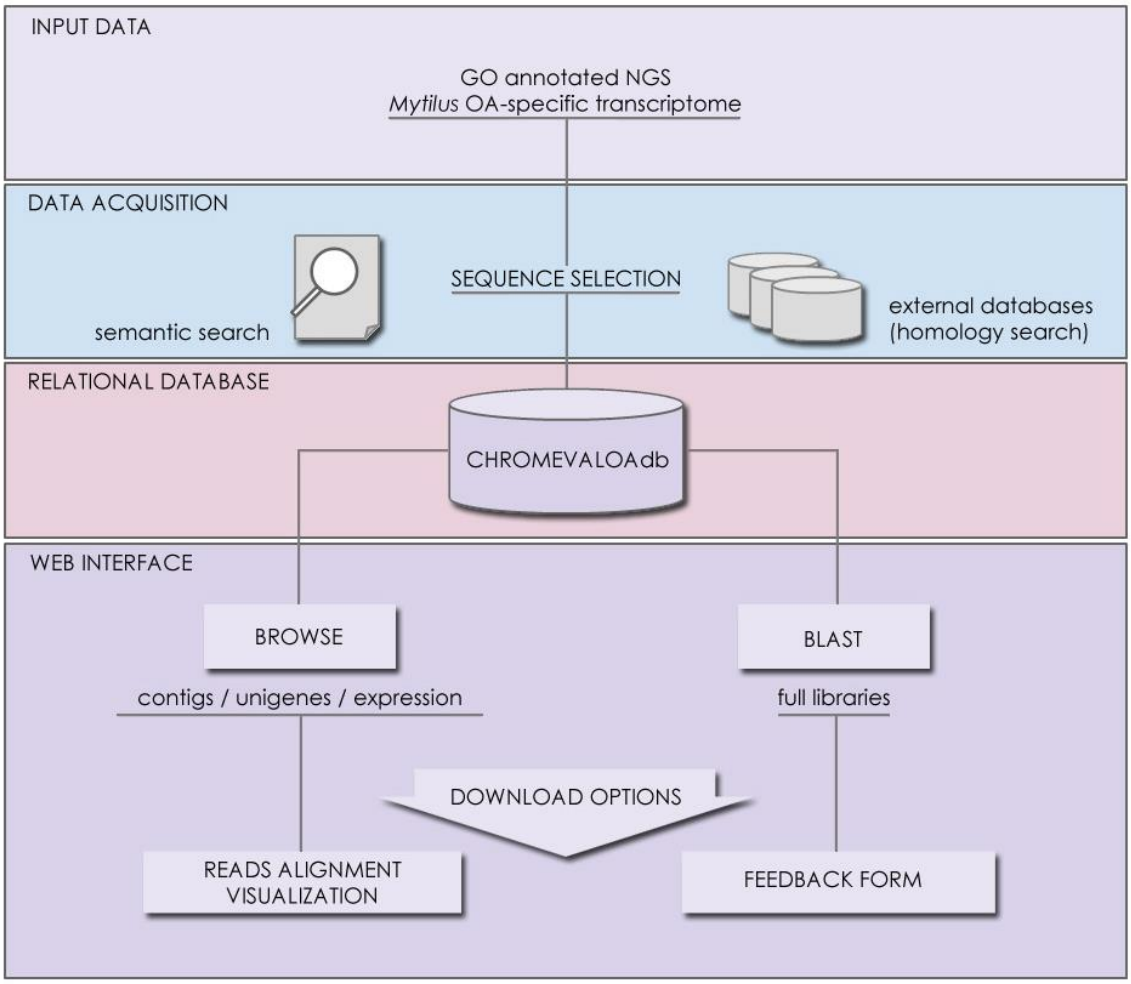


Figure 4. Chromatin-associated sequence query and results. CHROMEVALOAdb provides access to a search engine allowing users to find transcripts differentially expressed in response to OA. **A**) Searches can be performed on the basis of sequence homology (BLAST) or keywords. **B**) Results from individual unigenes provide gene ontology information as well as details on the contigs included in a given unigene. **C**) Differential expression information (upregulated and downregulated transcripts) for the chromatin-associated unigenes is presented through an intuitive format using arrow icons.

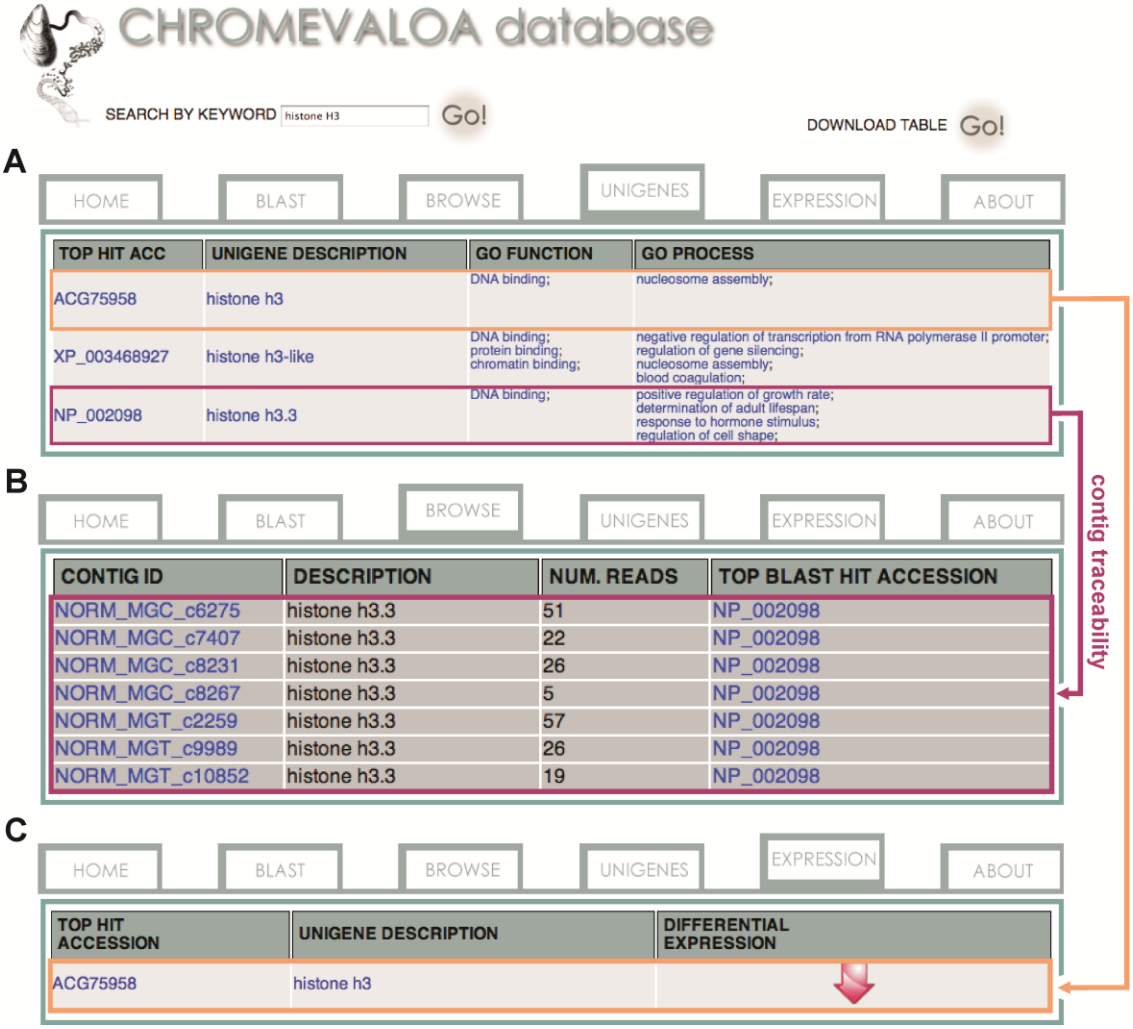


Table 1. HPLC-MS quantification of OA in digestive gland tissue.

Experimental conditions	OA-content (ng/g)
Control	Below detection limit (~0)
OA-exposed	18.27

Table 2. Amount of data in each step of the data processing pipeline.

Library	Reads	Contigs	Annotated Sequences
NORM_MGC (control)	493,440	16,395	7335
NORM_MGT (OA-exposed)	401,109	24,624	10,617
	Contigs	**Unigenes**	**Differentially Expressed**
TOTAL	41,019	2131	1254
CHROMATIN-ASSOCIATED	14,480	1124	90

CHAPTER III

UNBIASED HIGH-THROUGHPUT CHARACTERIZATION OF MUSSEL TRANSCRIPTOMIC RESPONSES TO SUBLETHAL CONCENTRATIONS OF THE BIOTOXIN OKADAIC ACID

In the current chapter, the transcriptomic characterization of the mussel Mytilus is expanded including different techniques of cDNA library preparation. Subsequently, differential expression and functional analyses are carried our using a tailor-made oligonucleotide microarray, providing a comprehensive description of the response of mussels exposed to Okadaic Acid-producers *Prorocentrum lima*.

This work was carried out in collaboration with Juan Fernandez-Tajes[1], Vanessa Aguiar-Pulido[4], M. Veronica Prego-Faraldo[2], Fernanda Florez-Barros[3], Alexia Sexto-Iglesias[2], Josefina Mendez[2] and Jose M. Eirin-Lopez[4], under respective affiliations to the University of Oxford (U.K.)[1], University of A Coruña (Spain)[2], University College London (U.K.)[3], and Florida International University (U.S.A.)[4].

Abstract

Background: Diarrhetic Shellfish Poisoning (DSP) Harmful Algal Blooms (HABs) represent a major threat for human consumers of shellfish. The biotoxin Okadaic Acid (OA), a well-known phosphatase inhibitor and tumor promoter, is the primary cause of acute DSP intoxications. Although several studies have described the molecular effects of high OA concentrations on sentinel organisms (e.g., bivalve molluscs), the effect of prolonged exposures to low (sublethal) OA concentrations is still unknown. In order to fill this gap, this work combines Next-Generation sequencing and custom-made microarray technologies to develop an unbiased characterization of the transcriptomic response of mussels during early stages of a DSP bloom.

Methods: Mussel specimens were exposed to a HAB episode simulating an early stage DSP bloom (200 cells/L of the dinoflagellate *Prorocentrum lima* for 24 hours). The

unbiased characterization of the transcriptomic responses triggered by OA was carried out using two complementary methods of cDNA library preparation: normalized and Suppression Subtractive Hybridization (SSH). Libraries were sequenced and read datasets were mapped to Gene Ontology and KEGG databases. A custom-made oligonucleotide microarray was developed based on these data, completing the expression analysis of digestive gland and gill tissues.

Results: Our findings show that exposure to sublethal concentrations of OA is enough to induce gene expression modifications in the mussel *Mytilus*. Transcriptomic analyses revealed an increase in proteasomal activity, molecular transport, cell cycle regulation, energy production and immune activity in mussels. Oppositely, a number of transcripts hypothesized to be responsive to OA (notably the Serine/Threonine phosphatases PP1 and PP2A) failed to show substantial modifications. Both digestive gland and gill tissues responded similarly to OA, although expression modifications were more dramatic in the former, supporting the choice of this tissue for future biomonitoring studies.

Discussion: Exposure to OA concentrations within legal limits for safe consumption of shellfish is enough to disrupt important cellular processes in mussels, eliciting sharp transcriptional changes as a result. By combining the study of cDNA libraries and a custom-made OA-specific microarray, our work provides a comprehensive characterization of the OA-specific transcriptome, improving the accuracy of the analysis of expresion profiles compared to single-replicated RNA-Seq methods.

The combination of our data with related studies helps understanding the molecular mechanisms underlying molecular responses to DSP episodes in marine organisms, providing useful information to develop a new generation of tools for the monitoring of OA pollution.

Introduction

Harmful Algal Blooms (HABs) constitute an environmental phenomenon encompassing critical relevance due to their increasing frequency and impact in coastal areas (Anderson 2009). Diarrhetic Shellfish Poisoning (DSP) blooms represent a major threat in widespread geographic areas comprising the Atlantic coast of Europe, Chile and Japan (Reguera et al. 2014), where natural outbreaks of toxic *Dinophysis* and *Prorocentrum* microalgae produce large amounts of DinophysisToXins (DTXs) and Okadaic Acid (OA) biotoxins (Sellner et al. 2003). OA is the primary cause of acute DSP intoxication of human consumers of shellfish, causing strong economic losses for the aquaculture industry. This biotoxin constitutes a well-known phosphatase inhibitor encompassing tumorigenic and apoptotic effects, even at low concentrations (Prego-Faraldo et al. 2015). Indeed, OA is capable of inducing genotoxic and cytotoxic damage, representing a hazard under chronic exposure conditions (Prego-Faraldo et al. 2013, Valdiglesias et al. 2013).

Given the noted risks of OA for human health and marine ecosystems, DSP events represent one of the most important threats for the shellfish aquaculture industry. Consequently, important efforts have been dedicated to develop rapid and sensible DSP biomonitoring methods, most notably using bivalve molluscs (e.g., mussels, oysters,

clams, etc.) as sentinel organisms (Manfrin et al. 2010, Fernandez-Tajes et al. 2011, McNabb et al. 2012, Romero-Geraldo et al. 2014, Huang et al. 2015). The choice of these organisms is supported by their wide distribution, sessile and filter-feeding lifestyles as well as their ability to accumulate high amounts of biotoxins, while displaying a particularly strong resilience to their harmful effects (Svensson et al. 2003, Prado-Alvarez et al. 2012, Prado-Alvarez et al. 2013). During the last decade, the increasing availability of genomic resources in bivalves has improved classical biomonitoring approaches (e.g., quantification of biotoxin content in mollusc tissues), notably by developing molecular high-throughput studies evaluating omic (transcriptomic and proteomic) responses to HAB stress and their potential biomarker application (Manfrin et al. 2010, Suarez-Ulloa et al. 2013a, Gerdol et al. 2014, Huang et al. 2015). Nonetheless, while this approach has proven to be a promising venue for pollution biomonitoring (Campos et al. 2012, Suarez-Ulloa et al. 2013b), additional efforts are still required to clarify the cause-effect relationship between the environmental stressor and changes in the pattern of expresion, and thus transform the extraordinary amount of molecular data resulting from omic experiments into a practical tool for marine pollution biomonitoring.

Mussels start accumulating OA in their tissues during early stages of DSP blooms, however, their commercialization is still allowed by the applicable legislation as long as the concentration of this biotoxin does not exceed the legal threshold of 160 µg OA equivalents/kg shellfish meat (European Union legislation). Nonetheless, it has been demonstrated that exposure to low OA concentrations for short periods of time is enough to produce genotoxic and cytotoxic effects *in vitro* (Prego-Faraldo et al. 2015). The present work aims to provide a better understanding of the molecular mechanisms

underlying the environmental responses of bivalve molluscs to sublethal concentrations of OA. For this purpose, Next-Generation sequencing and custom-made microarray technologies were combined to develop an unbiased characterization of the transcriptomic response of bivalve molluscs (mussels) to OA during early stages of a DSP bloom. These analyses build on previous studies (including our own) focused on specific subsets of genes [i.e., chromatin structure/function (Suarez-Ulloa et al. 2013a); oxidative stress, cell cycle regulation and immune response (Romero-Geraldo et al. 2014, Romero-Geraldo and Hernandez-Saavedra 2014)], as well as on the application of microarray technology to study the OA-specific transcriptome (Manfrin et al. 2010). Our results expand the scope, dimension and methodological approaches of these studies, improving the description of the cellular processes involved in the mussel response to OA toxicity. In doing so, this study generates omic information encompassing relevance for developing molecular signatures of marine pollution during DSP blooms. Such approach to pollution monitoring offers advantages versus a quantitative analytical method (i.e., LC-MS) since it has the potential of selectively identifying stressors of very different nature while assessing the magnitude of the toxic effects for the organisms as well as for the benthic community. In addition, it provides further insights into the molecular strategies underlying the extraordinary resilience of bivalve molluscs to environmental stress.

Methods

Specimen collection and experimental simulation of DSP HABs

Mussel specimens [*Mytilus galloprovincialis* (Lam.)] were collected in Valcobo beach, Galicia, NW Spain (43°19′02.71″N 8°21′56.35″W) in an area free of OA pollution during the resting period of the reproductive cycle (March) (Banni et al. 2011). Sampled individuals, adults between 10-15 cm, were randomly divided into two experimental groups; exposed (exposed group) and non-exposed (control group) to the OA-producing dinoflagellate *Prorocentrum lima.* Both groups were kept in aerated seawater tanks and fed continuously with a suspension of the microalgae *Tetraselmis suecica* and *Isochrysis galbana*. After acclimation (one week), the exposed group was fed with 200 cells/L of a *Prorocentrum lima* culture (exponential phase) for 24 hours. Specimens were dissected immediately after exposure, collecting samples from digestive gland and gill tissues. Each experimental sample consisted of tissue obtained from 5 individuals per group, dissected and pooled for RNA extraction.

RNA extraction and construction of cDNA libraries

The OA content in exposed and control samples was quantified using high-resolution mass spectrometry (Domenech et al. 2014). Total RNA was extracted from digestive gland and gill tissues using TRIzol® (Thermo Scientific, Waltham, MA) following the manufacturer's instructions. RNA concentration and quality check was measured using a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA) and a Bioanalyzer (Agilent Technologies, Santa Clara, CA). cDNA library construction and pyrosequencing were performed using digestive gland samples, based on the larger

absorption and accumulation of OA in this tissue. cDNA libraries were obtained from digestive gland tissue using the SMARTer™ PCR cDNA synthesis kit (Clontech, Mountain View, CA) and purified with GeneJET™ PCR Purification Kit (Thermo Scientific, Waltham, MA) according to the manufacturer's instructions.

The construction of normalized cDNA libraries (norm), for both exposed (mgt) and control (mgc) samples was carried out using the Trimer cDNA Normalization Kit (Evrogen, Moscow, Russia) following manufacturer's protocol. This method enhances the detection of rare (lower concentration) transcripts by decreasing the prevalence of highly abundant transcripts (Bogdanova et al. 2011). The Suppression Subtractive Hybridization (SSH) libraries were constructed using the PCR-Select™ cDNA subtraction kit (Clontech, Mountain View, CA), following manufacturer's instructions. Accordingly, two types of SSH libraries were produced: forward (fwd) and reverse (rev), representing upregulated and downregulated transcripts, respectively. This method was used to optimize the isolation of differentially expressed transcripts by removing commonly abundant cDNAs (Diatchenko et al. 1996).

Library sequencing and characterization

Normalized (exposed and control) and SSH (forward and reverse) libraries were sequenced by means of Roche-454 FLX+ Titanium pyrosequencing (Roche Diagnostics, Indianapolis, IN), with a sequencing depth of 40x. The obtained read datasets were preprocessed, assembled *de novo* and mapped to Gene Ontology (GO) and KEGG databases (Kanehisa 2002). Additionally, low quality reads were discarded, and adaptors and low quality ends were trimmed before *de novo* assembly using MIRA v. 3.9.16

(Chevreux B. 1999). Both normalized and SSH read datasets are available at NCBI's Bioproject database under the accession number PRJNA167773.

The generated contigs were annotated using BLAST (blastx) against the non-redundant protein sequence database (nr), setting a threshold e-value of $1e^{-6}$ (Altschul et al. 1997). Contigs were subsequently annotated with GO terms using the Blast2GO suite (Conesa et al. 2005, Gotz et al. 2008), including those terms obtained from InterPro and Annex analyses (Apweiler et al. 2001, Myhre et al. 2006).

Custom-made microarray construction and differential expression analysis

The sequencing and assembly of normalized and SSH libraries allowed to design specific probes targeting many of the transcripts identified. Accordingly, an Agilent oligonucleotide microarray encompassing 51,300 probes was constructed using the eArray™ design tool (Agilent Technologies, Santa Clara, CA) following a two-color Microarray-Based Gene Expression Analysis v. 6.5 Agilent-specific protocol with dye swap. Two biological replicates per tissue sample were analyzed in microarray experiments. Expression analyses were conducted using the R package limma from the Bioconductor repository (Smyth 2005). Results are organized based on the magnitude of the observed change in expression or Fold Change in a logarithmic scale (logFC) and the statistical significance of the observed change in expression represented by an adjusted p-value or False Discovery Rate by the Benjamini-Hochberg procedure (FDR). Probes showing an FDR < 0.05 were considered as differentially expressed. The correlation between logFC values of differentially expressed transcripts commonly observed in both digestive gland and gill tissues was analyzed using a linear regression based on Pearson's

coefficient of determination. The GO terms for the most representative biological processes in both upregulated and downregulated groups of transcripts were determined using topGO with statistical significance (p-values) calculated according to the weight algorithm (Alexa and Rahnenfuhrer 2010). Lastly, contigs were also mapped to the KEGG database for pathway analysis (Kanehisa 2002).

Results and Discussion

Characterization of OA-specific cDNA libraries in the mussel Mytilus

The analysis of OA in pooled digestive gland tissue of exposed individuals revealed a concentration of 18.27 ng of OA per gram of fresh tissue in exposed individuals (OA content in controls individuals is below detection limit), an order of magnitude below the legal OA limit established for safe consumption of shellfish in the European Union (Reguera et al. 2014). This result reinforces the focus of the present study on early stages of DSP HAB episodes, at a moment when mussels start accumulating OA in their tissues but their commercialization is still allowed by law. The construction of normalized (norm) cDNA libraries yielded 919,177 good quality reads, 514,276 for the exposed group (mgt) and 404,901 for the control group (mgc). After assembly, a total of 24,624 and 16,395 consensus sequences (contigs) were obtained, respectively. Complementary, the SSH libraries produced a set of 1,221,928 good quality reads (SSH) with 469,795 corresponding to the forward (fwd) library and 752,133 to the reverse (rev) library. Once assembled, a total of 21,591 contigs and 33,437 contigs were obtained, respectively.

Overall, blastx searches against the nr database resulted in the identification of 17,952 contigs from normalized libraries and 25,001 contigs from SSH libraries (see details in Table 1).

Given the high level of redundancy among *de novo* assembled libraries (Figure 1), contigs were combined into unigenes according to their annotation and were considered equivalent to the annotated transcripts (unigenes are therefore considered a set of uniquely identified transcripts). The normalized and SSH libraries constructed expand and complement partial sequence data previously released by our group in the Chromevaloa database (Suarez-Ulloa et al. 2013a). By combining both sets of sequences, the present work was able to produce a microarray tool increasing the coverage of OA-specific transcriptome in the mussel *Mytilus*, improving the unbiased analysis of the differences in gene expression.

Microarray-based analysis of transcriptomic responses to OA

The present work expanded previous analysis of the mussel's response to OA exposure using an omic approach using an oligonucleotide microarray designed from the sequences identified in pyrosequencing libraries. Accordingly, a medium-high coverage Agilent microarray (51,300 probes) was designed and developed using the sequences (contigs) obtained from the cDNA libraries constructed in this work. The hybridization of the microarray with RNA samples from exposed and control groups revealed a total number of 14,160 probes (digestive gland) and 6,913 probes (gill) differentially expressed (Figure 2).

The consistency between expression profiles in digestive gland and gill was assessed performing a linear regression of the logFC values of differentially expressed transcripts common for both tissues, (i.e., those showing FDR < 0.05 in both cases), showing a good correlation between both sets of transcripts (Figure 3). The detailed description of the transcripts displaying the highests differences in expression levels in both tissues, along with the maximum observed logFC value in the microarray analysis, is indicated in Supplementary Materials S2 and S3.

The present microarray analysis identified a set of transcripts displaying sharp expression differences between exposed and control treatments (Table 2 and Supplementary Materials S1 and S2), expanding the list of transcripts potentially involved in the response to OA (Manfrin et al. 2010, Suarez-Ulloa et al. 2013a). This was primarily facilitated by a larger coverage in the transcriptomic assessment, but also by the increase in bivalve genomic information that has been incorporated to molecular databases in recent years (Suarez-Ulloa et al. 2013b, Gerdol et al. 2014). Differentially expressed transcripts identified in this study (Supplementary Materials S1 and S2) include heat shock 70 kda protein 12b, proteases like cathepsins b and d, polyubiquitin and and proteasomal subunit beta type-4, commonly associated with an accumulation of misfolded or oxidized proteins observed under different types of environmental stress (Gotze et al. 2014). A subset of the identified transcripts showing the highest fold-change classified according to their main functional role is presented in Table 2.

Our results corroborate previous analyses describing the responses of *Mytilus galloprovincialis* to OA stress (Manfrin et al. 2010), particularly the strong upregulation of vdg3 and elongation factor 2. Vdg3 has only been identified in bivalves, being

particularly abundant in the digestive gland. It is associated with dramatic developmental changes in the mussel's life cycle at the benthic settlement stage. (He et al. 2014). On the other hand, elongation factor 2 (EEF-2) is widely ubiquitous across eukaryotic taxa and has an essential role in protein synthesis with specific regulatory mechanisms. It has been traditionally considered a housekeeping gene and used in qPCR analyses as an internal control. Although the functional implications of vdg3 and EEF-2 in the context of this study are still unclear, it is worth noting that our results add up to previous reports discouraging the use of EEF-2 as an internal control for qPCR analyses on bivalves without further validation (Du et al. 2013).

Opposite to these findings, a number of transcripts hypothesized to be responsive to OA failed to show substantial expression modifications under the conditions of this study. Notably, the Serine/Threonine phosphatases PP1 and PP2A, specific targets in OA toxicity mechanisms, did not show significant expression changes between treatments. OA is a well known selective inhibitor of the enzymatic activity of PP1 and PP2A phosphatases with critical consequences for the cell's fate (Shenolikar 1994), However, our results suggest that the upregulation of the PP1 and PP2A genes is not a relevant strategy versus the antagonist effects of OA. Similarly, Multi-Xenobiotic Resistance proteins (MXRs), good candidates to explain the high tolerance of bivalves versus pollution (Contardo-Jara et al. 2008), failed to show significant changes in expression. It is possible that their attributed role in OA uptake could be supplied by other proteins (e.g., the highly upregulated nose resistant to fluoxetine protein 6, a transport mediator of xenobiotics accross tissues).

Indeed, lysosomal uptake has been suggested as a possible explanation for the extraordinary tolerance of mussels to the effects of DSP pollution (Svensson et al. 2003).

In addition to transcripts previoulsy linked to OA responses, our results found an upregulation of an antimicrobial peptide (mytimacin) as well as an antifungal peptide (mytimycin) specific from mussels (Sonthi et al. 2011, Gerdol et al. 2012). Interestingly, mytimacin-5 (partial) was identified as one of the most upregulated transcripts in both gill and digestive gland. This peptide is especially interesting among the mytimacin family due to two additional cysteines in conserved positions predicted to form an extra disulfide bridge with yet unknown functional implications (Gerdol et al. 2012). C1q domain-containing proteins 1q3 and 1q25 showed a strong upregulation in the digestive gland. C1q is involved in the mammalian classical component pathway, playing an important role in innate immunity. Although no clear homologues to the vertebrate C1q complex subunits have been found in invertebrates yet, the C1q domain-containing protein family underwent a massive expansion in bivalves, including *Mytilus* spp. (Gerdol et al. 2015). C1q domain-containing proteins are very versatile and may display a wide range of ligand interactions and functions such as clearance of apoptotic cells through direct binding (Kishore et al. 2004). They have been found upregulated in molluscs challenged with different pathogens (Perrigault et al. 2009, Taris et al. 2009). Although their specific function remains unclear, the substantial upregulation found in the present work may be indicative of a relevant role during environmental stress responses.

These results provide valuable insights into the molecular effects of OA in the mussel, however, some considerations must be taken into account for their correct

interpretation. The levels of expression at protein and transcript levels may differ substantially due to kinetic effects as well as upper-layer regulatory mechanisms such as non-coding RNAs and post-translational modifications (Bai et al. 2015). Additionally, some of the transcripts showing a strong upregulation include endo-beta xylanases and endo-beta glucanases. Although this may be coherent with a faster-paced carbohydrate metabolism for energy production, it is likely that these differences between exposed and control groups arise artifactually from differences in the cell wall composition of the dinoflagellates fed to each group of mussels (i.e., *P. lima* is fed to the exposed group only). Similarly, the presence of antimicrobial and antifungal peptides may be due to the presence of infiltrated hemocytes in the digestive gland tissue. Noting these considerations, these results provide a general picture of the molecular response of mussels under DSP conditions but further validation is required to clarify the link between specific transcripts and the effects of OA in the digestive gland and gill tissues.

Expression and function profiles of transcripts differentially expressed in response to OA

The GO term annotation of transcripts differentially expressed in response to OA allowed the analysis of the biological processes in which their enconding genes are involved. A comparison of the functional profile for the two tissues studied is shown in Figure 4. These profiles are based on the levels of representation for the most general sub-categories in GO stemming from Biological Process (Ashburner et al. 2000).

Although absolute differences in magnitude are evident between digestive gland and gill (Figure 2), no major functional differences were found when comparing the profiles for both tissues (Figures 3 and 4). Nonetheless, such comparison may be

hampered by sample size differences (e.g., subtle tissue-specific differences could remain undetected) and the fact that the microarray could lack gill-specific transcripts. Indeed, recent reports suggest that OA may display tissue-specific effects. Accordingly, different cytotoxic effects of OA specific for different human cell types had been demonstrated *in vitro* (Rubiolo et al. 2011). Furthermore, it has been reported that mussel gills display higher sensitivity to OA than hemocytes after one hour exposure (Prego-Faraldo et al. 2015). This tissue specificity can be detected to some extent when comparing the enriched GO terms in the groups of up and down-regulated transcripts for both tissues (Table 3).

Most notably, terms related to the regulation of transcription and the cell cycle appear as enriched in the downregulated set of the digestive gland (e.g., transcription from RNA polymerase II promoter, histone acetylation, mitotic nuclear division, mitotic S phase), whereas for the gill these appear mostly represented in the upregulated set (e.g., positive regulation of cell growth, cell cycle, positive regulation of transcription, DNA-templated). This may suggest that the digestive gland is suffering to the point of trading cell development for detoxification and repair mechanisms, while the gill is overall less affected. In spite of this, both tissues consistently show enrichment in terms relative to the repair or degradation of damaged proteins. This could indicate that while the toxic mechanisms of OA are consistent across these tissues, different responses could be elicited depending on the level of the accumulation. However, the reduced statistical power in the enrichment analysis of the gill tissue due to a smaller sample size, hampers a fair comparison.

Further research will be required to clarify the extent in which the effects of OA are determined by the nature of the tissue, the time/dose or a combination of both.

Our results show an overall larger number of upregulated transcripts than downregulated transcripts, in agreement with previous reports although a strong dependence of the expresion profiles with time was demonstrated (Manfrin et al. 2010). Such findings are further supported by the analysis of the response of the Pacific oyster *Crassostrea gigas* to OA exposure using time-series (Romero-Geraldo et al. 2014), showing a strong dependence on time and dose. Altogether, it seems that expression profiles can hardly be extrapolated to other conditions different to those being studied. Given the highly dynamic nature of the transcriptome, only consistent patterns in expression can be informative of environmental stress conditions (Aardema and MacGregor 2002). This supports the use of expression signatures rather than individual biomarkers for biomonitoring purposes. Modeling systems of greater complexity including time and dose as variables would provide valuable information about the dynamics of the expression profiles.

The present work was completed by investigating the metabolic pathways associated with those enzymes identified as differentially expressed under OA exposure conditions (Supplementary Materials S3). Most of these pathways are involved in energy production (e.g., glycolysis/gluconeogenesis pathway, the citrate cycle, the pentose phosphate pathway and the oxidative phosphorylation pathway) as well as the regulation of the cell cycle and metabolism of drugs and xenobiotics. The observed functional profiles are consistent between tissues and also with observations in other organisms and types of abiotic stress. Accordingly, the role of metabolic functions was observed at the

proteomic level in the mussel *Perna viridis* exposed to OA pollution (Huang et al. 2015). The differential expression of enzymes involved in metabolic pathways such as Glycolisis, TCA and oxidative phosphorylation suggests that energy production becomes critical in situations of environmental stress. Such observations agree with the responses found in the Eastern oyster *Crassostrea virginica* exposed to different types of abiotic stress (Chapman et al. 2011). Furthermore, the role of the mTOR pathway as key regulator of the balance between energy consumption and cellular development was also evidenced in bivalves under environmental stress (Clark et al. 2013). An upregulation of enzymes PI3K, AMPK, LKB1 and ERK1/2 from this pathway (responsible for arresting the cell cycle when energy is required for resisting stress conditions) was found in the present work, suggesting that such mechanism is activated in the mussel as a response to OA toxicity. Lastly, the differential expression of enzymes involved in immunity-related pathways like biosynthesis of antibiotics further supports a link between environmental stress and changes in the immunity system (Malagoli et al. 2007).

Conclusions

The present work dissects the gene expression changes in different mussel tissues during early stages of DSP HAB episodes, suggesting that low concentrations of OA (below the legal OA limit established for safe consumption of shellfish) are enough to elicit sharp changes in the expression of genes involved in the response to this biotoxin. Prior to this work, a few studies attempted to investigate the transcriptomic changes in bivalves during HABs using high-throughput methods (Manfrin et al. 2010, Suarez-Ulloa et al. 2013a, Gerdol et al. 2014). However, the combined application of normalized and

SSH libraries together with the development of a custom-made OA-specific microarray in the present work, provides a more comprehensive characterization of the OA-specific transcriptome, improving the accuracy of the analysis of expression profiles compared to single-replicated RNA-Seq methods (Suarez-Ulloa et al. 2013a). The custom-made microarray platform generated in this work represents a convenient tool for long-term monitoring projects, offering a good level of standardization with lower requirements in computational resources comparing to the otherwise more informative RNA-Seq methodology (Guo et al. 2013). In addition, the transcriptomic coverage of this microarray is comparable to recent estimations for the size of the complete transcriptome in digestive gland of *Mytilus galloprovincialis* (Gerdol et al. 2014) as well as for the transcriptome of the Pacific oyster *Crassostrea gigas* (Zhang et al. 2012), thus representing a good approximation to an unbiased tool for expression analysis.

Our results suggest that the response to OA found in mussels is consistent with the model of intracellular response to stress previously reported for bivalve molluscs (Anderson et al. 2015). Accordingly, the activation of energy production mechanisms observed in the present work could be producing potentially harmful Reactive Oxygen Species (ROS), which unless controlled by chaperones or eliminated in the proteasomes, would induce apoptosis. An increase in ROS production has been recently reported for the mussels exposed to saxitoxins (i.e., neurotoxins responsible for the paralytic shellfish poisoning), supporting the applicability of this model to HABs exposure (Astuya et al. 2015). Indeed, our results show an upregulation in important chaperones (Hsp70) and proteases (cathepsins b and d) (Table 2) consistently with this model. Particularly the strong upregulation of cathepsins, known to be activated in the lysosomes (Kagedal et al.

2001), in conjunction with the activation of transport mechanisms suggested by our results (Table 3), offer support to the lysosomal uptake hypothesis proposed by Svensson et al. (Svensson et al. 2003). In addition, the upregulation of antimicrobial peptides suggests the activation of immunity mechanisms in conjunction with the general environmental stress response. However, it remains unclear whether this immune response is automatically triggered by abiotic factors or whether there is an opportunistic attack of pathogens present in the microbiota of the mussels. Current efforts are directed to clarify this question (De Rijcke et al. 2015).

Further work studying more restricted conditions with shorter periods of exposure and lower concentrations of dinoflagellates would better inform about the sensitivity of the transcriptomic approach for the detection of OA-pollution in the ocean. Complementary, long-term monitoring projects in combination with meta-analysis of publicly available data could provide valuable information on the basal trancriptomic changes constituting a general environmental response as well as on the specific transcriptomic signature of DSP toxicity stress.

References

Aardema, M. J. and J. T. MacGregor. 2002. Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies. Mutat Res 499:13-25.

Alexa, A. and J. Rahnenfuhrer. 2010. topGO: Enrichment analysis for Gene Ontology. Max Plank Institute Informatics.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.

Anderson, D. M. 2009. Approaches to monitoring, control and management of harmful algal blooms (HABs). Ocean Coast Manag 52:342-347.

Anderson, K., D. A. Taylor, E. L. Thompson, A. R. Melwani, S. V. Nair, and D. A. Raftos. 2015. Meta-analysis of studies using suppression subtractive hybridization and microarrays to investigate the effects of environmental stress on gene transcription in oysters. PLoS ONE 10:e0118839.

Apweiler, R., T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, L. Cerutti, F. Corpet, M. D. Croning, R. Durbin, L. Falquet, W. Fleischmann, J. Gouzy, H. Hermjakob, N. Hulo, I. Jonassen, D. Kahn, A. Kanapin, Y. Karavidopoulou, R. Lopez, B. Marx, N. J. Mulder, T. M. Oinn, M. Pagni, F. Servant, C. J. Sigrist, and E. M. Zdobnov. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res 29:37-40.

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, and G. Sherlock. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25-29.

Astuya, A., C. Carrera, V. Ulloa, A. Aballay, G. Nunez-Acuna, H. Hegaret, and C. Gallardo-Escarate. 2015. Saxitoxin modulates immunological parameters and gene transcription in *Mytilus chilensis* hemocytes. Int J Mol Sci 16:15235-15250.

Bai, Y., S. Wang, H. Zhong, Q. Yang, F. Zhang, Z. Zhuang, J. Yuan, X. Nie, and S. Wang. 2015. Integrative analyses reveal transcriptome-proteome correlation in biological pathways and secondary metabolism clusters in *A. flavus* in response to temperature. Scientific Reports 5:14582.

Banni, M., A. Negri, F. Mignone, H. Boussetta, A. Viarengo, and F. Dondero. 2011. Gene expression rhythms in the mussel *Mytilus galloprovincialis* (Lam.) across an annual cycle. PLoS ONE 6:e18904.

Bogdanova, E. A., E. V. Barsova, I. A. Shagina, A. Scheglov, V. Anisimova, L. L. Vagner, S. A. Lukyanov, and D. A. Shagin. 2011. Normalization of full-length-enriched cDNA. Methods Mol Biol 729:85-98.

Campos, A., S. Tedesco, V. Vasconcelos, and S. Cristobal. 2012. Proteomic research in bivalves: towards the identification of molecular markers of aquatic pollution. J Proteomics 75:4346-4359.

Chapman, R. W., A. Mancia, M. Beal, A. Veloso, C. Rathburn, A. Blair, A. F. Holland, G. W. Warr, G. Didinato, I. M. Sokolova, E. F. Wirth, E. Duffy, and D. Sanger. 2011. The transcriptomic responses of the eastern oyster, *Crassostrea virginica*, to environmental conditions. Mol Ecol 20:1431-1449.

Chevreux B., W. T., Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. Proceedings of the German Conference on Bioinformatics (GCB):45-56.

Clark, M. S., M. A. S. Thorne, A. Amaral, F. Vieira, F. M. Batista, J. Reis, and D. M. Power. 2013. Identification of molecular and physiological responses to chronic environmental challenge in an invasive species: the Pacific oyster, *Crassostrea gigas*. Ecology and Evolution 3:3283-3297.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674-3676.

Contardo-Jara, V., S. Pflugmacher, and C. Wiegand. 2008. Multi-xenobiotic-resistance a possible explanation for the insensitivity of bivalves towards cyanobacterial toxins. Toxicon 52:936-943.

De Rijcke, M., M. B. Vandegehuchte, J. V. Bussche, N. Nevejan, L. Vanhaecke, K. A. De Schamphelaere, and C. R. Janssen. 2015. Common European Harmful Algal Blooms affect the viability and innate immune responses of *Mytilus edulis* larvae. Fish Shellfish Immunol:175-181.

Diatchenko, L., Y. F. Lau, A. P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukyanov, K. Lukyanov, N. Gurskaya, E. D. Sverdlov, and P. D. Siebert. 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc Natl Acad Sci U S A 93:6025-6030.

Domenech, A., N. Cortes-Francisco, O. Palacios, J. M. Franco, P. Riobo, J. J. Llerena, S. Vichi, and J. Caixach. 2014. Determination of lipophilic marine toxins in mussels. Quantification and confirmation criteria using high resolution mass spectrometry. J Chromatogr A 1328:16-25.

Du, Y., L. Zhang, F. Xu, B. Huang, G. Zhang, and L. Li. 2013. Validation of housekeeping genes as internal controls for studying gene expression during Pacific oyster (*Crassostrea gigas*) development by quantitative real-time PCR. Fish Shellfish Immunol 34:939-945.

Fernandez-Tajes, J., F. Florez, S. Pereira, T. Rabade, B. Laffon, and J. Mendez. 2011. Use of three bivalve species for biomonitoring a polluted estuarine environment. Environ Monit Assess 177:289-300.

Gerdol, M., G. De Moro, C. Manfrin, A. Milandri, E. Riccardi, A. Beran, P. Venier, and A. Pallavicini. 2014. RNA sequencing and de novo assembly of the digestive gland transcriptome in *Mytilus galloprovincialis* fed with toxinogenic and non-toxic strains of *Alexandrium minutum*. BMC Res Notes 7:722.

Gerdol, M., G. De Moro, C. Manfrin, P. Venier, and A. Pallavicini. 2012. Big defensins and mytimacins, new AMP families of the Mediterranean mussel *Mytilus galloprovincialis*. Dev Comp Immunol 36:390-399.

Gerdol, M., P. Venier, and A. Pallavicini. 2015. The genome of the Pacific oyster *Crassostrea gigas* brings new insights on the massive expansion of the C1q gene family in Bivalvia. Dev Comp Immunol 49:59-71.

Gotz, S., J. M. Garcia-Gomez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talon, J. Dopazo, and A. Conesa. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420-3435.

Gotze, S., O. B. Matoo, E. Beniash, R. Saborowski, and I. M. Sokolova. 2014. Interactive effects of CO2 and trace metals on the proteasome activity and cellular stress response of marine bivalves *Crassostrea virginica* and *Mercenaria mercenaria*. Aquatic Toxicology 149:65-82.

Guo, Y., C. I. Li, F. Ye, and Y. Shyr. 2013. Evaluation of read count based RNAseq analysis methods. BMC Genomics 14:S2.

He, T. F., J. Chen, J. Zhang, C. H. Ke, and W. W. You. 2014. SARP19 and vdg3 gene families are functionally related during abalone metamorphosis. Dev Genes Evol 224:197-207.

Huang, L., Y. Zou, H.-w. Weng, H.-Y. Li, J.-S. Liu, and W.-D. Yang. 2015. Proteomic profile in *Perna viridis* after exposed to *Prorocentrum lima*, a dinoflagellate producing DSP toxins. Environmental Pollution 196:350-357.

Kagedal, K., U. Johansson, and K. Ollinger. 2001. The lysosomal protease cathepsin D mediates apoptosis induced by oxidative stress. FASEB J 15:1592-1594.

Kanehisa, M. 2002. The KEGG database. Novartis Found Symp 247:91-101; discussion 101-103, 119-128, 244-152.

Kishore, U., C. Gaboriaud, P. Waters, A. K. Shrive, T. J. Greenhough, K. B. Reid, R. B. Sim, and G. J. Arlaud. 2004. C1q and tumor necrosis factor superfamily: modularity and versatility. Trends Immunol 25:551-561.

Malagoli, D., L. Casarini, S. Sacchi, and E. Ottaviani. 2007. Stress and immune response in the mussel *Mytilus galloprovincialis*. Fish & Shellfish Immunology 23:171-177.

Manfrin, C., R. Dreos, S. Battistella, A. Beran, M. Gerdol, L. Varotto, G. Lanfranchi, P. Venier, and A. Pallavicini. 2010. Mediterranean mussel gene expression profile induced by okadaic acid exposure. Environ Sci Technol 44:8276-8283.

McNabb, P. S., A. I. Selwood, R. Van Ginkel, M. Boundy, and P. T. Holland. 2012. Determination of brevetoxins in shellfish by LC/MS/MS: single-laboratory validation. J AOAC Int 95:1097-1105.

Myhre, S., H. Tveit, T. Mollestad, and A. Laegreid. 2006. Additional gene ontology structure for improved biological reasoning. Bioinformatics 22:2020-2027.

Perrigault, M., A. Tanguy, and B. Allam. 2009. Identification and expression of differentially expressed genes in the hard clam, *Mercenaria mercenaria*, in response to quahog parasite unknown (QPX). BMC Genomics 10.

Prado-Alvarez, M., F. Florez-Barros, J. Mendez, and J. Fernandez-Tajes. 2013. Effect of okadaic acid on carpet shell clam (*Ruditapes decussatus*) haemocytes by in vitro exposure and harmful algal bloom simulation assays. Cell Biol Toxicol 29:189-197.

Prado-Alvarez, M., F. Florez-Barros, A. Sexto-Iglesias, J. Mendez, and J. Fernandez-Tajes. 2012. Effects of okadaic acid on haemocytes from *Mytilus galloprovincialis*: a comparison between field and laboratory studies. Mar Environ Res 81:90-93.

Prego-Faraldo, M. V., V. Valdiglesias, B. Laffon, J. M. Eirin-Lopez, and J. Mendez. 2015. *In vitro* analysis of early genotoxic and cytotoxic effects of okadaic acid in different cell types of the mussel *Mytilus galloprovincialis*. J Toxicol Environ Health A 78:814-824.

Prego-Faraldo, M. V., V. Valdiglesias, J. Mendez, and J. M. Eirin-Lopez. 2013. Okadaic Acid meet and greet: an insight into detection methods, response strategies and genotoxic effects in marine invertebrates. Mar Drugs 11:2829-2845.

Reguera, B., P. Riobo, F. Rodriguez, P. A. Diaz, G. Pizarro, B. Paz, J. M. Franco, and J. Blanco. 2014. Dinophysis toxins: causative organisms, distribution and fate in shellfish. Mar Drugs 12:394-461.

Romero-Geraldo, R. D. J., N. Garcia-Lagunas, and N. Y. Hernandez-Saavedra. 2014. Effects of In Vitro Exposure to Diarrheic Toxin Producer *Prorocentrum lima* on Gene Expressions Related to Cell Cycle Regulation and Immune Response in *Crassostrea gigas*. PLoS ONE 9.

Romero-Geraldo, R. d. J. and N. Hernandez-Saavedra. 2014. Stress Gene Expression in *Crassostrea gigas* (Thunberg, 1793) in response to experimental exposure to the toxic dinoflagellate *Prorocentrum lima* (Ehrenberg) Dodge, 1975. Aquaculture Research 45:1512-1522.

Rubiolo, J. A., H. Lopez-Alonso, F. V. Vega, M. R. Vieytes, and L. M. Botana. 2011. Okadaic acid and dinophysis toxin 2 have differential toxicological effects in hepatic cell lines inducing cell cycle arrest, at G0/G1 or G2/M with aberrant mitosis depending on the cell line. Arch Toxicol 85:1541-1550.

Sellner, K. G., G. J. Doucette, and G. J. Kirkpatrick. 2003. Harmful algal blooms: causes, impacts and detection. J Ind Microbiol Biotechnol 30:383-406.

Shenolikar, S. 1994. Protein Serine/Threonine phosphatases - New avenues for cell regulation. Annual Review of Cell Biology 10:55-86.

Smyth, G. K. 2005. Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions Using R and Bioconductor 1:397-420.

Sonthi, M., M. Toubiana, A. Pallavicini, P. Venier, and P. Roch. 2011. Diversity of coding sequences and gene structures of the antifungal peptide mytimycin (MytM) from the Mediterranean mussel, *Mytilus galloprovincialis*. Mar Biotechnol (NY) 13:857-867.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, C. Rivera-Casas, R. Gonzalez-Romero, J. Ausio, J. Mendez, J. Dorado, and J. M. Eirin-Lopez. 2013a. The CHROMEVALOA database: a resource for the evaluation of okadaic acid contamination in the marine environment based on the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis*. Mar Drugs 11:830-841.

Suarez-Ulloa, V., J. Fernandez-Tajes, C. Manfrin, M. Gerdol, P. Venier, and J. M. Eirin-Lopez. 2013b. Bivalve omics: state of the art and potential applications for the biomonitoring of harmful marine compounds. Mar Drugs 11:4370-4389.

Svensson, S., A. Sarngren, and L. Forlin. 2003. Mussel blood cells, resistant to the cytotoxic effects of okadaic acid, do not express cell membrane p-glycoprotein activity (multixenobiotic resistance). Aquat Toxicol 65:27-37.

Taris, N., R. P. Lang, P. W. Reno, and M. D. Camara. 2009. Transcriptome response of the Pacific oyster (*Crassostrea gigas*) to infection with *Vibrio tubiashii* using cDNA AFLP differential display. Animal Genetics 40:663-677.

Valdiglesias, V., M. Prego-Faraldo, E. Pásaro, J. Méndez, and B. Laffon. 2013. Okadaic acid: More than a diarrheic toxin. Marine Drugs 11:4328.

Zhang, G., X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que, Y. Xie, P. W. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, J. Meng, L. Yang, J. Liu, B. Wen, N. Zhang, Z. Huang, Q. Zhu, Y. Feng, A. Mount, D. Hedgecock, Z. Xu, Y. Liu, T. Domazet-Loso, Y. Du, X. Sun, S. Zhang, B. Liu, P. Cheng, X. Jiang, J. Li, D. Fan, W. Wang, W. Fu, T. Wang, B. Wang, J. Zhang, Z. Peng, Y. Li, N. Li, M. Chen, Y. He, F. Tan, X. Song, Q. Zheng, R. Huang, H. Yang, X. Du, L. Chen, M. Yang, P. M. Gaffney, S. Wang, L. Luo, Z. She, Y. Ming, W. Huang, B. Huang, Y. Zhang, T. Qu, P. Ni, G. Miao, Q. Wang, C. E. Steinberg, H. Wang, L. Qian, X. Liu, and Y. Yin. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49-54.

Figure 3. Venn diagram showing the extent of redundancy between the different libraries constructed in the present work: *norm_mgc*, normalized control library; *norm_mgt*, normalized exposed library; *ssh_fwd*, SSH forward library; *ssh_rev*, SSH reverse library.

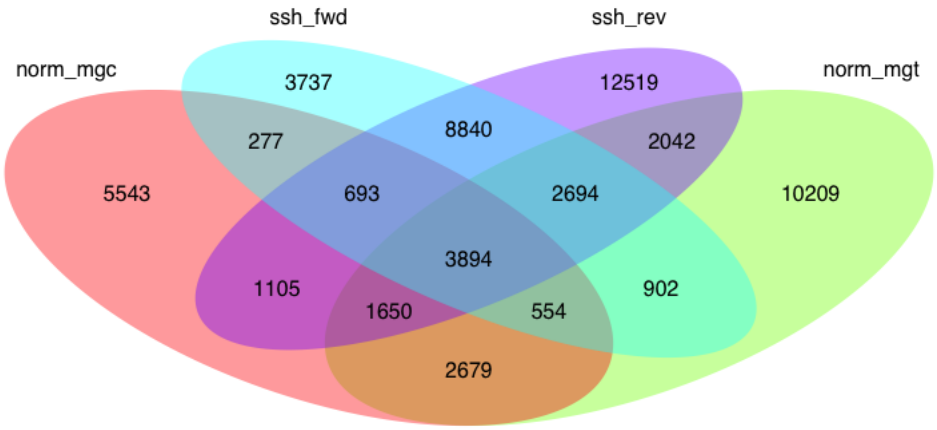


Figure 2. V-plots showing gene expression differences detected through microarray analysis in digestive gland (left) and gill (right) tissues. These differences are represented as net expression change (logFC) with statistical significance (FDR) indicated as a logarithmic scale. Probes highlighted in blue (FDR < 0.05) and purple (FDR < 0.05 and logFC > 2) represent the groups of transcripts displaying largest changes in gene expression between exposed and control treatments.

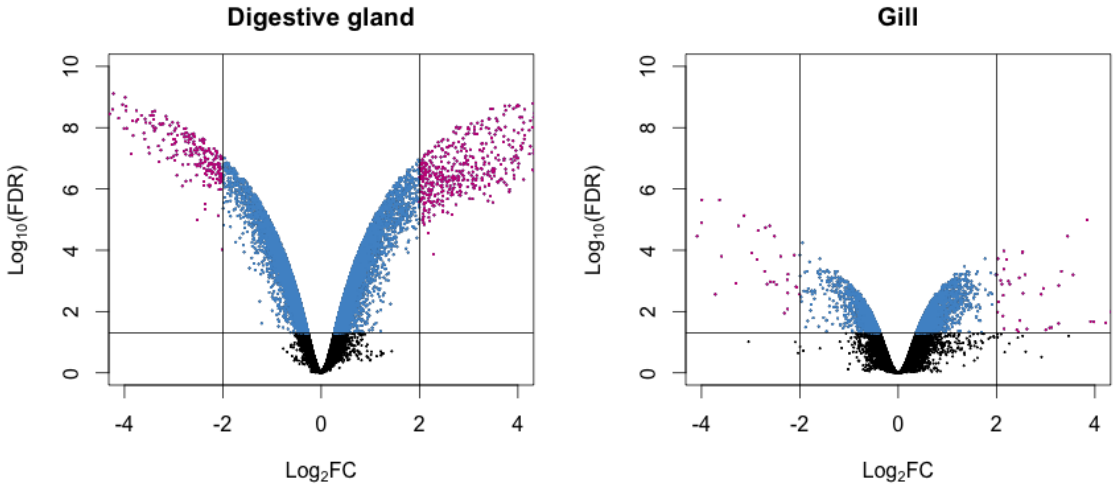


Figure 3. Correlation between paired logFC values calculated for transcripts identified in digestive gland and gill tissues between exposed and control treatments. Overall, a good level of agreement is found for gene expression changes ($R^2 \cong 0.6$).

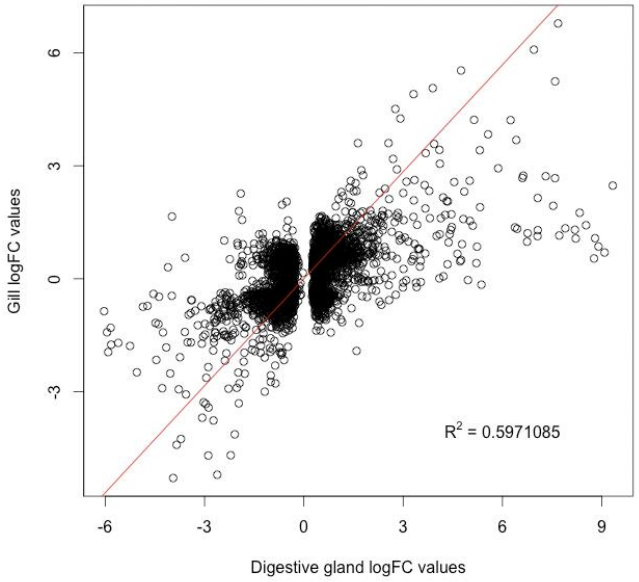


Figure 4. Graphical representation of the GO terms (general sub-categories in Biological Process ontology) most represented in transcripts differentially expressed for each mussel tissue according to the microarray analysis. The length of the bars is proportional to the number of sequences annotated for each specific GO term.

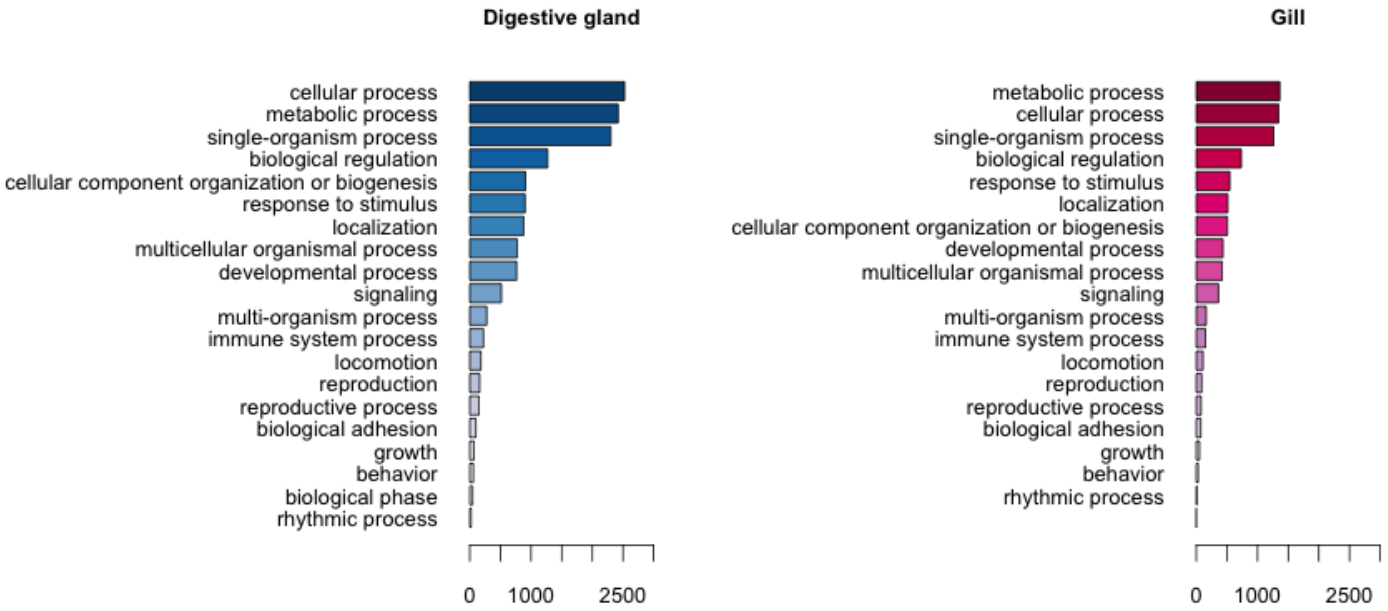


Table 1. Reads and annotated contigs obtained from the constructed cDNA libraries.

	Normalized Libraries		SSH Libraries	
	exposed	**control**	**forward**	**reverse**
Reads	514,276	404,901	469,795	752,133
Contigs	24,624	16,395	21,591	33,437
Annotated Contigs	10,617 (43%)	7,335 (45%)	6,448 (30%)	18,553 (55%)

Table 2. Selected subsets of differentially expressed transcripts identified by microarray analysis representative of the following functional categories: a) Protein repair or degradation, b) immune response, c) transport and energy production and d) cell cycle regulation.

a	**b**
Protein repair / degradation	Immune response
Heat shock 70 kda protein 12b	Mytimacin-5
Cathepsin d	c1q domain-containing protein 1q3
Cathepsin b	c1q domain-containing protein 1q25
Proteasome subunit beta type-4	Mytimicin precursor

c	**d**
Transport / Energy production	Cell cycle regulation
Nose resistant to fluoxetine protein 6	Bcl2 adenovirus e1b 19-kd protein-interacting
Interferon-inducible GTPase 5-like	Apoptosis inhibitor iap
Nadh dehydrogenase subunit	Jagged 1
Atpase H+ transporting lysosomal 21 kda v0 subunit	Oncoprotein-induced transcript 3 protein

Table 3. Enriched GO terms in sets of differentially expressed transcripts in both digestive gland and gill tissues: a.1) upregulated set in digestive gland, a.2) downregulated set in digestive gland, b.1) upregulated set in gill and b.2) downregulated set in gill. Data is sorted based on p-value in increasing order. P-values are calculated according to the weight algorithm in TopGO.

a.1

DIGESTIVE GLAND - UPREGULATED

GO term description	GO number	annotated	expected	p-value
vesicle-mediated transport	GO:0016192	176	60.79	1.00E-09
maintenance of protein localization in endoplasmic reticulum	GO:0035437	16	5.53	3.80E-08
cellular response to glucose starvation	GO:0042149	19	6.56	5.70E-08
cellular modified amino acid metabolic process	GO:0006575	72	24.87	7.40E-08
ER overload response	GO:0006983	15	5.18	1.10E-07
activation of signaling protein activity involved in unfolded protein response	GO:0006987	15	5.18	1.10E-07
cerebellum structural organization	GO:0021589	15	5.18	1.10E-07
regulation of cell migration	GO:0030334	47	16.23	2.40E-06
negative regulation of cellular response to growth factor stimulus	GO:0090288	23	7.94	2.80E-06
endosomal transport	GO:0016197	31	10.71	7.00E-06
pteridine-containing compound metabolic process	GO:0042558	17	5.87	7.00E-06
secretion by cell	GO:0032940	58	20.03	1.50E-05
coenzyme metabolic process	GO:0006732	117	40.41	1.70E-05
regulation of actin filament polymerization	GO:0030833	25	8.63	2.60E-05
cerebellar Purkinje cell layer development	GO:0021680	18	6.22	2.90E-05
cellular response to interleukin-4	GO:0071353	27	9.33	3.00E-05
regulation of protein ubiquitination	GO:0031396	68	23.49	4.40E-05
negative regulation of protein polymerization	GO:0032272	15	5.18	4.60E-05
aminoglycan metabolic process	GO:0006022	29	10.02	0.00015
regulation of lipid metabolic process	GO:0019216	16	5.53	0.00017

a.2

DIGESTIVE GLAND - DOWNREGULATED

GO term description	GO number	annotated	expected	p-value
DNA metabolic process	GO:0006259	157	102.77	2.60E-09
ribonucleoprotein complex biogenesis	GO:0022613	130	85.1	4.20E-08
mRNA processing	GO:0006397	70	45.82	1.80E-06
cilium morphogenesis	GO:0060271	53	34.69	4.10E-06
transcription from RNA polymerase II promoter	GO:0006366	186	121.76	4.90E-06
mitochondrial ATP synthesis coupled electron transport	GO:0042775	81	53.02	1.50E-05
mitotic nuclear division	GO:0007067	82	53.68	2.30E-05
inorganic cation transmembrane transport	GO:0098662	109	71.35	2.80E-05
chromosome organization	GO:0051276	167	109.32	7.20E-05
microtubule-based movement	GO:0007018	97	63.5	8.20E-05
cilium organization	GO:0044782	42	27.49	0.00018
positive regulation of ubiquitin-protein transferase activity	GO:0051443	32	20.95	0.00019
sodium ion transport	GO:0006814	46	30.11	0.00022
anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process	GO:0031145	31	20.29	0.00028
G1/S transition of mitotic cell cycle	GO:0000082	44	28.8	0.00029
mitotic S phase	GO:0000084	30	19.64	0.0004
chromatin remodeling	GO:0006338	39	25.53	0.0005
regulation of multi-organism process	GO:0043900	43	28.15	0.00058
cilium or flagellum-dependent cell motility	GO:0001539	17	11.13	0.00073
histone acetylation	GO:0016573	33	21.6	0.00078

b.1

GILL - UPREGULATED

GO term description	GO number	annotated	expected	p-value
biological_process	GO:0008150	926	516.95	1.60E-06
positive regulation of cell growth	GO:0030307	12	6.7	0.00087
carbohydrate metabolic process	GO:0005975	39	21.77	0.00149
cellular catabolic process	GO:0044248	30	16.75	0.04401
protein folding	GO:0006457	23	12.84	0.05753
protein polyubiquitination	GO:0000209	17	9.49	0.06629
lipid metabolic process	GO:0006629	24	13.4	0.09686
intracellular transport	GO:0046907	18	10.05	0.11899
nucleobase-containing compound catabolic...	GO:0034655	12	6.7	0.14577
proteolysis	GO:0006508	22	12.28	0.16789
single-organism developmental process	GO:0044767	94	52.48	0.19993
cellular macromolecular complex assembly	GO:0034622	11	6.14	0.20518
protein complex subunit organization	GO:0071822	11	6.14	0.20518
generation of neurons	GO:0048699	16	8.93	0.21463
vesicle-mediated transport	GO:0016192	21	11.72	0.21627
response to oxygen-containing compound	GO:1901700	10	5.58	0.28271
protein complex assembly	GO:0006461	10	5.58	0.31076
cell cycle	GO:0007049	17	9.49	0.31266
response to external stimulus	GO:0009605	17	9.49	0.31266
positive regulation of transcription, DN...	GO:0045893	12	6.7	0.32401

b.2

GILL - DOWNREGULATED

GO term description	GO number	annotated	expected	p-value
microtubule-based process	GO:0007017	25	7.31	1.40E-07
energy derivation by oxidation of organi...	GO:0015980	10	2.92	0.0012
heterocycle biosynthetic process	GO:0018130	35	10.23	0.0041
aromatic compound biosynthetic process	GO:0019438	35	10.23	0.0041
cellular nitrogen compound biosynthetic ...	GO:0044271	37	10.82	0.0086
organic cyclic compound biosynthetic pro...	GO:1901362	38	11.11	0.0121
biological_process	GO:0008150	926	270.74	0.0161
regulation of signal transduction	GO:0009966	18	5.26	0.0165
cellular protein modification process	GO:0006464	57	16.67	0.0537
nucleotide metabolic process	GO:0009117	11	3.22	0.0687
response to organic substance	GO:0010033	26	7.6	0.1048
single-organism transport	GO:0044765	66	19.3	0.1202
cell morphogenesis involved in different...	GO:0000904	10	2.92	0.1363
regulation of multicellular organismal p...	GO:0051239	10	2.92	0.1363
purine-containing compound metabolic pro...	GO:0072521	10	2.92	0.1363
single-organism biosynthetic process	GO:0044711	19	5.56	0.1603
cell surface receptor signaling pathway	GO:0007166	31	9.06	0.1635
regulation of biological quality	GO:0065008	17	4.97	0.202
anatomical structure morphogenesis	GO:0009653	33	9.65	0.2438
protein modification by small protein co...	GO:0032446	28	8.19	0.2836

CHAPTER IV

EFFECTS OF FLORIDA RED TIDES ON HISTONE VARIANT EXPRESSION AND DNA METHYLATION IN THE EASTERN OYSTER *CRASSOSTREA VIRGINICA*

Complementing the thorough transcriptomic characterization of the response to HAB conditions in a bivalve, the mussel *Mytilus*, carried out in the previous chapter, the current chapter focuses on what is the epigenetic response. In this case, the oyster *Crassostrea virginica* was exposed to a lab-controlled simulation of a natural Florida Red Tide (FRT) event. Specific epigenetic marks, including histone variants and DNA methylation were studied to ascertain their role in the response to brevetoxins, associated to FRTs.

This work was carried out in collaboration with Rodrigo Gonzalez-Romero[1], Javier Rodriguez Casariego[1], Daniel Garcia-Souto[2], Gabriel Diaz[1], Abraham Smith[1], Juan Jose Pasantes[2], Gary Rand[1] and Jose M. Eirin-Lopez[1], under affiliations to Florida International University (U.S.A.)[1] and University of Vigo (Spain)[2].

Abstract

Massive algal proliferations known as Harmful Algal Blooms (HABs) represent one of the most important threats to coastal areas. Among them, the so-called Florida Red Tides (FRTs, caused by blooms of the dinoflagellate *Karenia brevis* and associated brevetoxins) are particularly detrimental in the southeastern U.S., causing high mortality rates and annual losses in excess of $40 million. The ability of marine organisms to cope with environmental stressors (including those produced during HABs) is influenced by genetic and epigenetic mechanisms, the latter resulting in phenotypic changes caused by heritable modifications in gene expression, without involving changes in the genetic (DNA) sequence. Yet, studies examining cause-effect relationships between environmental stressors, specific epigenetic mechanisms and subsequent responses are

still lacking. The present work contributes to increase this knowledge by investigating the effects of Florida Red Tides on two types of mechanisms participating in the epigenetic memory of Eastern oysters: histone variants and DNA methylation. For that purpose, a HAB simulation was conducted in laboratory conditions, exposing oysters to increasing concentrations of *K. brevis*. The obtained results revealed, for the first time, the existence of H2A.X, H2A.Z and macroH2A genes in this organism, encoding histone variants potentially involved in the maintenance of genome integrity during responses to the genotoxic effect of brevetoxins. Additionally, an increase in H2A.X phosphorylation (γH2A.X, a marker of DNA damage) and a decrease in global DNA methylation were observed as the HAB simulation progressed. Overall, the present work provides a basis to better understand how epigenetic mechanisms participate in responses to environmental stress in marine invertebrates, opening new avenues to incorporate environmental epigenetics approaches into management and conservation programs.

Introduction

The impacts of global change on marine ecosystems are dramatic, as evidenced by growing pollution and acidification, sea level rise, and changes in ocean temperature and currents (Curry and Mauritzen 2005). Among them, the increase in the frequency and toxicity of episodes of massive algal proliferations known as Harmful Algal Blooms (HABs) represents one of the most important threats to fisheries, aquaculture-based industries, and human populations in coastal areas (Mat et al. 2013, Cook et al. 2015). During HABs, large amounts of potentially harmful biotoxins are produced and recruited into the human food chain through their accumulation by edible marine organisms

(Cardozo et al. 2007, Plakas et al. 2008). These biotoxins, among other relevant effects, have the ability to cause DNA damage, both in marine organisms and human consumers of contaminated shellfish (Valdiglesias et al. 2013, Prego-Faraldo et al. 2015, Prego-Faraldo et al. 2016).

The southeastern U.S. is particularly affected by HABs (Flewelling et al. 2005), notably by the so-called Florida Red Tides [FRTs, blooms of the dinoflagellate *Karenia brevis* (Davis 1948, Brand and Compton 2007)] and associated brevetoxins (PbTx). FRTs are responsible for high mortality rates of marine invertebrates, fishes and marine mammals (Brand et al. 2012), causing annual losses in excess of $40 million (Twiner et al. 2007). The consumption of brevetoxin-laden shellfish causes Neurotoxic Shellfish Poisoning (NSP) syndrome (McFarren et al. 1965), and exposure to aerosolized brevetoxins is responsible for respiratory distress (Abraham et al. 2005) and severe allergic reactions in humans (Fleming et al. 2007). Furthermore, brevetoxins convey critical disruptive effects at the most fundamental level, as they can induce DNA damage and apoptosis (Radwan and Ramsdell 2008, Murrell and Gibson 2009, 2011). Still, while the harmful effects of brevetoxins are well documented, the role these biotoxins play in *K. brevis* is still uncertain (Sunda et al. 2013, Errera et al. 2014).

Coastal areas bear the brunt of HABs, impacting marine communities and commercial resources, especially the aquaculture industry. Consequently, bivalve molluscs are generally used as sentinel organisms to study HAB pollution (Collin et al. 2010, Campos et al. 2012, Fernandez-Tajes et al. 2012, Luchmann et al. 2012, Milan et al. 2013, Suarez-Ulloa et al. 2013, Prego-Faraldo et al. 2015, Prego-Faraldo et al. 2016). Eastern oysters (*Crassostrea virginica)* are particularly relevant due to their abundance,

economic importance and major role in the function of estuary ecosystems in areas affected by FRTs (Dame 1972). During these episodes Eastern oysters rapidly accumulate brevetoxins (Plakas et al. 2002), causing mortality and interfering with larval development (Leverone et al. 2006, Rolton et al. 2014). At the subcellular level, the exposure to high brevetoxin concentrations was previously linked to the activation of specific molecular mechanisms involved in defense, detoxification and stress response in oysters (Mello et al. 2012), including advanced apoptosis and apoptosis-regulating systems (Zhang et al. 2014). Similarly, the presence of oxidative stress was described in other marine organisms including corals, turtles and manatees, as a consequence of exposure to *K. brevis* during FRTs (Ross et al. 2010, Perrault et al. 2014, Walsh et al. 2015).

Traditionally, the ability of organisms to cope with adverse environmental conditions such as those imposed by HABs was attributed to specifically adapted genotypes (Hoffmann and Willi 2008). However, we now know that responses to environmental changes are largely dependent on epigenetic regulatory mechanisms, in other words, heritable changes in gene expression resulting from modifications in chromatin structure, without involving changes in the genetic information stored in DNA (Allis et al. 2007, Feil and Fraga 2012). The study of the epigenetic mechanisms mediating exposure-response relationships constitutes the basis for environmental epigenetic analyses (Baccarelli and Bollati 2009, Bollati and Baccarelli 2010), providing information about how different environmental factors influence phenotypic variation (Cortessis et al. 2012, Suarez-Ulloa et al. 2015, Etchegaray and Mostoslavsky 2016). Chromatin, the association between DNA and chromosomal proteins, provides a

framework for the study of epigenetic modifications, including DNA methylation, histone variants [a set of minority histone proteins encompassing specialized functions in chromatin metabolism (Henikoff and Smith 2015)] and their post-translational modifications (PTMs), as well as non-coding RNAs, among others (Kouzarides 2007, Ptashne 2007, Arya et al. 2010, Talbert and Henikoff 2010, Mercer and Mattick 2013). These modifications define specific epigenomic states throughout the genome that are susceptible of being transmitted trans-generationally (both mitotically and meiotically), setting the basis for short-term acclimatization and long-term adaptation (Migicovsky and Kovalchuk 2011, Fernandez et al. 2014).

Among marine invertebrates, bivalve molluscs constitute emerging models in environmental epigenetics, as illustrated by recent studies examining the role of DNA methylation in the Pacific oyster (Gavery and Roberts 2010, Gavery and Roberts 2013, Diaz-Freije et al. 2014, Gavery and Roberts 2014, Olson and Roberts 2014) and the characterization of chromatin structure and histone variants in mussels (González-Romero et al. 2012, Rivera-Casas et al. 2016a, Rivera-Casas et al. 2016b). Yet, studies examining cause-effect relationships between environmental stressors, specific epigenetic mechanisms and subsequent responses in marine invertebrates are still lacking. That is primarily due to the absence of genomic information for many environmentally and ecologically relevant organisms, along with a scarce knowledge about their chromatin structure and dynamics (Arenas-Mena et al. 2007, Schulmeister et al. 2007, Suarez-Ulloa et al. 2015). The present work fills that gap by studying the role of two different epigenetic mechanisms during Eastern oyster responses to FRTs. First, given the genotoxic effect of brevetoxins produced during FRTs (Radwan and Ramsdell 2008,

Murrell and Gibson 2009, 2011), this study targets a group of three histone variants known to be involved in the maintenance of genome integrity, namely H2A.X, H2A.Z, and macroH2A (Li et al. 2005, Ivashkevich et al. 2012, Shi and Oberdoerffer 2012, Papamichos-Chronakis and Peterson 2013, Khurana et al. 2014). Although the presence of these variants has not been yet described in oysters, our previous results on other mollusc species support that notion (González-Romero et al. 2012, Rivera-Casas et al. 2016a, Rivera-Casas et al. 2016b). Second, the study of environmental epigenetic responses is completed by the analysis of genome-wide DNA methylation patterns during FRTs. The obtained results are consistent with a role for the phosphorylated histone γH2A.X and DNA methylation during exposure to the brevetoxin-producing dinoflagellate *K. brevis*.

Methods

Specimen collection and laboratory acclimatization

Eastern oyster specimens were collected at Rookery Bay National Estuarine Research Reserve, Naples FL. Collected organisms were transported to the laboratory and acclimatized to controlled conditions of temperature (25 ± 2 °C), aeration (dissolved oxygen > 6 mg/L), salinity (30 ± 2 ppt), light cycle and substrate in a recirculating seawater system for 5 weeks. During this period, oyster specimens were fed twice a day using a commercial mix of marine microalgae (1 mL per 30L tank) including *Isochrysis* sp., *Pavlova* sp., *Tetraselmis* sp., *Chaetoceros calcitrans*, *Thalassiosira weissflogii* and *T. pseudonana* (Shellfish Diet 1800, REED MARICULTURE). The concentrated diet was diluted with filtered seawater (1:10) to ensure separation of microalgae cells.

Experimental HAB simulation

The present work simulated the toxic effect resulting from the exposure of Eastern oysters to an incoming patch of *K. brevis* HAB that disappears within 24 hours (Stumpf et al. 2003). HAB simulations were carried out in the laboratory using an exponential culture of the brevetoxin-producing dinoflagellate *K. brevis* (approx. 45,000 cells/mL, courtesy of Dr. K. Rein, Florida International University). Toxin content in microalgae cultures was estimated using LC-MS as 5.0-6.8 pg/cell (Sun et al. 2016). Given that the present work aimed to simulate a natural FRT episode, *K. brevis* exposures were chosen over exposures to purified brevetoxins. The latter have major limitations for that purpose, including the existence of several brevetoxins subtypes (not all can be purified) and their high hydrophobicity, hampering homogeneous solution in seawater. Two groups of oysters (control and treated) were defined in the simulation, each consisting of 3 experimental units (biological replicates) with 20 individuals/unit (80 individuals per group). Each experimental unit consisted of a static 20 L tank where oysters were fed and acclimatized to controlled conditions for 24 h prior to exposure to *K. brevis*. Treated experimental units were subsequently exposed to increasing concentrations of *K. brevis*, starting at 5 cell/mL and reaching 1000 cell/mL after 5 h exposure (Fig. 1A). These conditions mirror *K. brevis* levels during a typical medium-high intensity FRT episode (Stumpf et al. 2003).

Specimens were initially fed with lower microalgae volumes (0.1 mL per 20 L tank) in order to promote feeding behavior. Salinity, pH, dissolved oxygen and conductivity conditions were monitored throughout the experiment to ensure water

quality maintaining optimal levels throughout the experiment, and thus discarding significant effects of increased algal levels in exposed individuals. Cell concentrations of *K. brevis* were quantified hourly using a Multisizer 4 Coulter Counter instrument (BECKMAN-COULTER). Dosage was adjusted to maintain homogeneous microalgae concentrations across replicates (Fig. 1B). Specimens (n=2 oysters per biological replicate) were collected at 4 different time points: T0, before exposure begins; T1, after 3 h exposure; T2, after 5 h exposure; T3, after 24 h exposure. Upon sampling, oysters were shucked, and gills were dissected and immediately flash-frozen in liquid nitrogen. The choice of gill tissue as model system in the present work was motivated by previous reports indicating that gills are the part of the organism first contacted by biotoxins during HABs, experiencing substantial DNA damage (Prego-Faraldo et al. 2015, Prego-Faraldo et al. 2016).

Gene expression analysis

Total RNA was extracted from gill cells using Ribozol Reagent (AMRESCO) and digested with PerfeCTa DNase I (QUANTA BIOSCIENCES) to eliminate residual genomic DNA. cDNA was synthesized using qScript cDNA Supermix (QUANTA BIOSCIENCES). Expression analyses were performed by means of quantitative PCR (qPCR) experiments. Accordingly, specific primers for Eastern oyster H2A.X, H2A.Z, and macroH2A histone genes were designed based on sequences retrieved from GenBank (Table 1) using Primer-BLAST software (Ye et al. 2012). Primers were also generated for GAPDH and RPL13, used as reference genes for normalization purposes.

Primer efficiency was calculated based on the slope of calibration curves constructed using ten-fold dilution steps, according to the formula $E = 10^{-1/slope}$.

Gene expression profiles were examined in gills from Eastern oyster by measuring SYBR green incorporation, using FastStart Essential DNA Green Master (ROCHE) in a LightCycler 96 System (ROCHE). A pre-incubation step of 10 min at 95 ºC was included. cDNA amplification was carried out in 45 cycles under the following conditions: denaturation for 10 s at 95 ºC, annealing for 10 s at 60 ºC, and elongation for 10 s at 72 ºC, including a final melting gradient up to 97 ºC using a ramp of 4.4 ºC/s to check primer specificity. Each individual reaction was carried out in triplicate including negative controls, No Template Control (NTC) and Non-Reverse Transcription Control (NRTC). Results were recorded as normalized ratio values by the LightCycler 96 Software version 1.1 following the Pfaffl method (Pfaffl 2001). Statistical analyses were carried out using ANOVA after observing an approximately normal distribution with QQ-plots, including a post-hoc Tukey HSD test for multiple comparisons. All analyses were carried out using R-Bioconductor (Gentleman et al. 2004).

Histone protein extraction, separation and western blot analysis

Histone protein isolation was performed as described elsewhere (Ausio and Moore 1998), adapting the protocol to oyster gill tissues in the present work. Accordingly, gills were homogenized with a Dounce Tissue Grinder in 100 mM KCl, 50 mM Tris-HCl (pH 7.5), 1 Mm $MgCl_2$, and 0.5% Triton X-100 buffer containing a protease inhibitor mixture. After homogenization and incubation on ice for 5 min, samples were centrifuged for 10 min at 12,000 g (4 ºC). The resulting pellets were

resuspended in 0.6 N HCl, homogenized and centrifuged again. The HCl supernatant extracts were precipitated overnight with six volumes of acetone at -20 ºC and centrifuged for 10 min at 12,000 g (4 ºC). Acetone pellets were dried using a speedvac concentrator and stored at -80 ºC.

The extracted histone proteins were separated using Sodium Dodecyl-Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) in ClearPAGE SDS gels 4-20% (C.B.S. SCIENTIFIC). Gels were stained with 0.2% (w/v) Coomassie blue in 25% (v/v) 2-propanol, 10% (v/v) acetic acid, and destained in 10% (v/v) 2-propanol, 10% (v/v) acetic acid. Western-blot analyses were performed using commercial human-specific antibodies, including anti-H2A.X (ABCAM), anti-γH2A.X (ROCKLAND), anti-H2A.Z (THERMOFISHER SCIENTIFIC); as well as antibodies raised in-house against mollusc macroH2A and H4 histones. Gels were electro-transferred to a nitrocellulose membrane (C.B.S. SCIENTIFIC) and processed as described elsewhere (Finn et al. 2008). Membranes were incubated with a secondary goat anti-rabbit antibody (ROCKLAND) whose signal was subsequently detected using enhanced chemiluminescence (Amershan ECL Prime Western Blotting Detection Reagent, GE HEALTHCARE LIFE SCIENCES) and a CCD based imager (ChemiDoc-It TS2 Imager, UVP).

Genome-wide DNA methylation analysis

Genomic DNA was purified from gill tissue as described elsewhere (Fernandez-Tajes et al. 2007), adapting the protocol to Eastern oysters in the present work. DNA methylation was analyzed using a Methylation Sensitive Amplified Polymorphism (MSAP) protocol, adapted from (Reyna-Lopez et al. 1997). This method is based on the

differential cleavage reactivity to Cytosine methylation (CCGG sites) between the isoesquizomeric endonucleases HpaII and MspI. Accordingly, while HpaII is sensitive to internal Cytosine methylation (i.e., 5'-C^{m}CGG-3' / 3'-GG^{m}CC-5') or hypermethylation states (i.e., 5'-^{m}C^{m}CGG-3' / 3'-GG^{m}C^{m}C-5'), MspI is sensitive to external Cytosine methylation, including hemimethylation states where methylation is located on the external Cytosine of the CCGG pattern but on a single DNA strand (i.e., 5'-^{m}CCGG-3' / 3'-GGCC-5') and hypermethylation states. Thus, by comparing both restriction profiles, MSAP analyses provide a basis for establishing global Cytosine methylation patterns and to perform comparisons among different samples (Diaz-Freije et al. 2014). Briefly, gill genomic DNA samples were separately digested with EcoRI/HpaII and EcoRI/MspI in parallel reactions. The resulting fragments were ligated to EcoRI adapters (5'-CTCGTAGACTGCGTACC-3', 3'-AATTGGTACGCAGTCTAC-5') and HpaII/MspI adapters (5'-GACGATGAGTCTAGAA-3', 3'-CGTTCTAGACTCATC-5'). Digestion-ligation reactions were performed simultaneously for 2 h at 37 °C in a solution consisting of 10 nM DNA, 4 U of EcoRI (NEB), 1 U of either HpaII (NEB) or MspI (NEB), 1 U T4 DNA ligase (NEB), 1X ligase buffer (NEB), 1X CutSmart Buffer (NEB). Restriction fragments were selectively amplified through two consecutive PCR reactions as follows: a first reaction containing 1 μL of 1/10 diluted restriction-ligation product, 20 pM of each HpaII/MspI+T (5´-GATGAGTCTAGAACGGT-3´) and EcoRI+A (5´-GACTGCGTACCAATTCA-3´) primers, 1X PCR buffer, 0.5 mM dNTPs (THERMO FISHER SCIENTIFIC), 2.5 mM $MgCl_2$, and 1 U BIOTAQ DNA polymerase (BIOLINE). The PCR conditions included an initial denaturation step of 2 min at 95 °C followed by 20 amplification cycles (20 s at 95 °C; 30 s at 56 °C; 2 min at 72 °C) and a

final extension step of 30 min at 72 °C. The second reaction used 0.5 µL of 1/10 of the pre-selective PCR product, 0.83 pM of each 6-FAM labelled selective primer, HpaII/MspI+TAG (5´-GATGAGTCTAGAACGGTCC-3´) and HpaII/MspI+TCC (5´-GATGAGTCTAGAACGGTAG-3´), 1X PCR buffer, 0.5 mM dNTPs, 2.5 mM $MgCl_2$ and 1 U DNA polymerase (BIOTAQ). Conditions included an initial denaturation step of 2 min at 95 °C followed by 10 amplification cycles (20 s at 95 °C; 30 s at 66 °C; 2 min at 72 °C) and a final extension step of 30 min at 72 °C. Amplified products were run in a ABI Prism 310 Genetic Analyzer (APPLIED BIOSYSTEMS) with a GeneScan 500 ROX in the DNA Core facility at Florida International University.

MSAP restriction profiles were scored using the GeneMapper v.3.7 software (APPLIED BIOSYSTEMS) and the resulting absence-presence matrix was analyzed using the R package MSAP (Perez-Figueroa 2013). For that purpose, the presence of both EcoRI-HpaII and EcoRI-MspI fragments (1/1) is used as an indicator of unmethylated states; Hemimethylated and internal Cytosine methylated loci are represented by EcoRI–HpaII (1/0) and EcoRI–MspI (0/1) fragments, respectively. Lastly, the absence of both fragments (0/0) mirrors either a target mutation (i.e., recognition site CCGG is not present any more) or hypermethylation, being therefore considered uninformative. Differences among methylation profiles were studied throughout the experiment using Analysis of MOlecular VAriance (AMOVA) and Principal Component Analysis (PCA). Additionally, Fisher's exact tests were carried out (Benjamini & Hochberg multi-test corrections, adjusted $p < 0.05$) based on counts for the four pattern categories (unmethylated, hemimethylated, hypermethylated, and internal Cytosine methylation), identifying loci with non-random distribution of methylation states. Relationship

estimates were computed using Gower's Coefficient of Similarity, and the resulting distance matrix was clustered using UPGMA and visualized as a heatmap with the "ComplexHeatmap" R package (Gu et al. 2016).

Results and Discussion

Florida Red Tide simulation and feeding response in Eastern oysters

FRTs are usually patchy and highly mobile (i.e., driven by currents and wind), displaying heterogeneous dinoflagellate concentrations as these episodes develop (Carvalho et al. 2011). In order to improve HAB simulation, the present work focused on the first 24 h of a FRT episode, where dinoflagellate concentrations can be predicted more accurately. Accordingly, Eastern oyster individuals were exposed to an exponential increase in the concentration of the brevetoxin-producing dinoflagellate *K. brevis* (Fig. 1A), simulating a typical medium-high intensity FRT episode (Stumpf et al. 2003). The presence of active feeding in exposed oysters was corroborated by the reduction in the number of *K. brevis* cells in experimental tanks after each hourly dosage application (Fig. 1B), as well as by additional indicators including oyster valve opening and production of feces. These results, together with the absence of oyster mortality and the stability in water quality parameters throughout the experiment and across treatment groups, support the effectiveness of HAB simulation in ensuring exposure of Eastern oysters to brevetoxins through *K. brevis* ingestion.

Eastern oyster exposure to K. brevis does not trigger specific modifications in the expression of H2A.X, H2A.Z and macroH2A

Different epigenetic mechanisms participate in environmental responses, including histone variants encompassing highly specialized functions (Talbert and Henikoff 2014). Among the different histone families constituting the nucleosome, H2A and H3 display the highest diversity of variants, including H2A.X, H2A.Z, macroH2A, and H2A.Bbd (H2A family); as well as H3.3 and cenH3 (H3 family), among several others (see references for details) (Cheema and Ausio 2015, Henikoff and Smith 2015). Among H2A variants, H2A.X, H2A.Z and macroH2A stand out due to their involvement (along with PTMs) in the maintenance of genome integrity (Li et al. 2005, Ivashkevich et al. 2012, Shi and Oberdoerffer 2012, Papamichos-Chronakis and Peterson 2013, Khurana et al. 2014). Given the ability of brevetoxins to induce DNA damage (Radwan and Ramsdell 2008, Murrell and Gibson 2009, 2011), it is hypothesized here that genes encoding H2A.X, H2A.Z and macroH2A will be differentially regulated in response to these stressors. Consequently, the expression of these variants was analyzed during HAB simulation, constituting the first report describing their presence in oysters. The obtained results revealed an absence of significant changes in gene expression across different time points (adjusted p-value > 0.05) and a substantial amount of inter-individual variation (Fig. 2). That is best exemplified by the expression of H2A.X at time point T1 (Fig. 2A), where individuals under similar experimental conditions display a two-fold factor difference in gene expression.

These results suggest that oyster responses to *K. brevis* exposure do not involve specific modifications in H2A.X, H2A.Z and macroH2A transcription. Yet, it still might be possible that specific cause-effect relationships are masked by the high levels of variation observed.

The role of histone variants was further investigated by studying H2A.X, H2A.Z and macroH2A protein expression levels (Fig. 3). This work is the first report describing the presence of these proteins in oysters, as well as the validity of commercial antibodies to detect histone variants in this group. Two of these antibodies constitute commercial products specifically raised against human H2A.X and H2A.Z histones. Their suitability for the study of oyster variants is supported by the high level of evolutionary conservation displayed by these proteins in eukaryotes (Malik and Henikoff 2003, Eirín-López et al. 2009), as well as by previous reports supporting the experimental application of these antibodies in molluscs (González-Romero et al. 2012, Rivera-Casas et al. 2016b). In the case of macroH2A, an antibody specifically raised against mussel macroH2A (able to specifically detect this variant in invertebrates) was used, further supporting its suitability for oysters (Rivera-Casas et al. 2016a). Similarly to gene expression analyses, western blot hybridizations did not show substantial modifications in H2A.X, H2A.Z or macroH2A protein levels throughout HAB simulation (Fig. 3B). Altogether, these results suggest that oyster exposure to FRTs does not result in specific transcriptional or translational modifications in the studied histone variants. Nonetheless, an active role for histone variants during environmental responses, without invoking modifications in their expression levels, could still be possible. Accordingly, the dynamic

mobilization of pre-existing histone variants to specific chromatin regions was described during environmental responses (Morrison and Shen 2005, Weber and Henikoff 2014), including responses to DNA damage (Lopez et al. 2012). Long-term exposure simulations combined with biopsy sampling of specific individuals (Acosta-Salmón and Southgate 2004) will complement the present results, helping ascertain the nature of the role played by histone variants during environmental responses (Romero-Geraldo et al. 2014, Romero-Geraldo et al. 2016).

γH2A.X increases in oysters exposed to growing concentrations of K. brevis

The phosphorylation of a Serine residue at the C-terminal domain of the H2A.X protein (γH2A.X) and its subsequent accumulation at damaged chromatin regions (γH2A.X foci) rapidly marks double-stranded DNA breaks for repair, constituting a biomarker of DNA damage widely used in standardized assays (Kuo and Yang 2008, Turinetto and Giachino 2015). Since DNA breaks stand out as the most severe effects caused by brevetoxin exposure (Altaf et al. 2007, Radwan and Ramsdell 2008, Murrell and Gibson 2009, 2011), protein expression analyses were complemented with the study of γH2A.X dynamics during HAB simulation. The obtained results revealed an increase in γH2A.X, concomitantly with the exposure of Eastern oysters to increasing concentrations of *K. brevis* (T1, T2), and followed by a slight decrease during the recovery phase (T3, Fig. 3C). Based on these results, it seems plausible that γH2A.X formation is triggered by the genotoxic effect of brevetoxins on Eastern oysters (Li et al. 2005, Ivashkevich et al. 2012). This finding has a dual relevance: first, it corroborates the

specialization of histone variants and PTMs in the chromatin of molluscs (González-Romero et al. 2012, Rivera-Casas et al. 2016a, Rivera-Casas et al. 2016b), supporting a role for γH2A.X during environmental responses in invertebrates (Suarez-Ulloa et al. 2015) and the evolutionary conservation of this mechanism (Kinner et al. 2008, Lee et al. 2014). Second, it validates the application of commercial antibodies to detect γH2A.X in bivalves, opening new avenues for monitoring DNA damage and health in populations of marine invertebrates.

Global DNA methylation decreases in oysters during HAB simulation

The regulatory role of DNA methylation during environmental responses was previously investigated in marine invertebrates, linking modifications in methylation of stress-responsive genes with phenotypic plasticity and adaptation (Gavery and Roberts 2010). The present work builds on those results, monitoring DNA methylation in Eastern oysters during HAB simulation. Methylation Sensitive Amplified Polymorphism (MSAP) analyses revealed a total of 428 bands, including 295 methylation-susceptible loci. Among those, 204 (69%) were considered polymorphic (i.e., they have at least two occurrences of both states). The analysis of different methylation states over time revealed a slight increase in unmethylated and hemimethylated states and a loss of fully methylated bands as HAB simulation progressed (T2 and T3, Table 2). More specifically, comparisons among different time points revealed significant differences in DNA methylation ($\Phi ST = 0.1296$, $p < 0.05$, Table 3), notably between earlier (T0, T1) and later (T2, T3) exposure, as well as between T2 and T3. Such differences were further evidenced by Principal Component Analysis (PCA), comparing samples in a space of

reduced dimensionality (Fig. 4A). Accordingly, the first coordinate (C1) accounts for 16% of the observed variation in DNA methylation, allowing to discriminate between two major groups: one of them corresponding to the early time points (T0 and T1), and the other one to the overexposed oyster samples from T2 and T3. The observed differences in DNA methylation were further supported by Fisher exact tests, identifying 10 MSAP loci showing significant differences ($p < 0.05$) between early (T0 and T1) and late (T2 and T3) exposure times (Fig. 4B). These loci were characterized based on methylation profiles, defining two groups: a first one predominantly showing internal Cytosine methylation at T0 and T1, but mostly unmethylated at T2 and T3 (left cluster); and a second group consisting of fully methylated fragments undergoing demethylation towards later exposure times (right cluster). Overall, MSAP results were consistent with a decrease in genome-wide DNA methylation levels as *K. brevis* concentration increased.

Specific DNA methylation profiles contribute to environmentally induced phenotypes, displaying different levels of heritability in invertebrates (Vandegehuchte et al. 2009). In the case of molluscs, DNA methylation participates in the epigenetic regulation of gene expression during development (Riviere et al. 2013, Diaz-Freije et al. 2014), as well as during responses to environmental stress (Gavery and Roberts 2014). Furthermore, it has been recently suggested that a reduction in global DNA methylation during early invasive episodes could compensate for low genetic variation caused by founder effects in the pygmy mussel (Ardura et al. 2017). Particularly, DNA hypomethylation was described in different organisms during responses to environmental pollution (Vandegehuchte et al. 2009, Chen et al. 2012, Fang et al. 2013, Dimond and Roberts 2016), including exposure to agents inducing oxidative stress (Mirbahai and

Chipman 2014). Given the ability of brevetoxins to induce oxidative damage, it seems plausible that the observed reduction in global DNA methylation in Eastern oysters is a direct consequence of their exposure to the brevetoxin-producing *K. brevis* during HAB simulation.

Conclusions

The present work investigated the effects of FRTs on two types of mechanisms potentially participating in the epigenetic memory of Eastern oysters: histone variants and DNA methylation. This research revealed, for the first time, the existence of H2A.X, H2A.Z and macroH2A genes in oysters along with their active transcription and translation, validating the application of specific qPCR primers, as well as commercial and invertebrate-specific antibodies for their study *in vivo*. While the expression of these variants does not seem to be specifically altered during FRTs, the observed increase in γH2A.X would be consistent with the presence of oxidative DNA damage in Eastern oysters, as a consequence of their exposure to brevetoxins produced by *K. brevis*. Based on these results, the application of γH2A.X as a biomarker of oxidative stress seems plausible, with obvious implications for the conservation and management of Eastern oysters as well as other marine invertebrates. Still, further analyses clarifying the relationship between the magnitude of γH2A.X and the levels of DNA damage will be required to fully elucidate the applicability of this histone variant modification in this group of organisms. Additionally, the study of DNA methylation patterns revealed significant differences between early (T0, T1) and late (T2, T3) stages of the HAB simulation. More precisely, a decrease in genome-wide DNA methylation levels was

observed as the simulation progressed. These results are consistent with the documented connection between DNA hypomethylation and environmental responses (Vandegehuchte et al. 2010, Chen et al. 2012, Fang et al. 2013, Dimond and Roberts 2016), thus supporting an active role for this epigenetic mechanism during oyster responses to *K. brevis* exposure (i.e., modulation of genes specifically involved in responses to brevetoxin stress). Overall, the present work provides a basis to better understand how epigenetic mechanisms participate in responses to environmental stress in marine invertebrates. By doing so, it opens new avenues to incorporate environmental epigenetics approaches into management and conservation programs.

References

Abraham, W. M., A. J. Bourdelais, J. R. Sabater, A. Ahmed, T. A. Lee, I. Serebriakov, and D. G. Baden. 2005. Airway responses to aerosolized brevetoxins in an animal model of asthma. Am J Respir Crit Care Med 171:26-34.

Acosta-Salmón, H. and P. C. Southgate. 2004. Use of a biopsy technique to obtain gonad tissue from the blacklip pearl oyster *Pinctada margaritifera* (L.). Aquaculture Research 35:93-96.

Allis, C. D., T. Jenuwein, and D. Reinberg. 2007. Epigenetics. Cold Spring Harbor Laboratory Press, New York.

Altaf, M., N. Saksouk, and J. Cote. 2007. Histone modifications in response to DNA damage. Mutat Res 618:81-90.

Ardura, A., A. Zaiko, P. Moran, S. Planes, and E. Garcia-Vazquez. 2017. Epigenetic signatures of invasive status in populations of marine invertebrates. Sci Rep 7:42193.

Arenas-Mena, C., K. S. Wong, and N. R. Arandi-Foroshani. 2007. Histone H2A.Z expression in two indirectly developing marine invertebrates correlates with undifferentiated and multipotent cells. Evol Dev 9:231-243.

Arya, G., A. Maitra, and S. A. Grigoryev. 2010. A structural perspective on the where, how, why, and what of nucleosome positioning. J Biomol Struct Dyn 27:803-820.

Ausio, J. and S. C. Moore. 1998. Reconstitution of chromatin complexes from high-performance liquid chromatography-purified histones. Methods 15:333-342.

Baccarelli, A. and V. Bollati. 2009. Epigenetics and environmental chemicals. Curr Opin Pediatr 21:243-251.

Bollati, V. and A. Baccarelli. 2010. Environmental epigenetics. Heredity 105:105-112.

Brand, L. E., L. Campbell, and E. Bresnan. 2012. *Karenia*: the biology and ecology of a toxic genus. Harmful Algae 14:156-178.

Brand, L. E. and A. Compton. 2007. Long-term increase in *Karenia brevis* abundance along the Southwest Florida Coast. Harmful Algae 6:232-252.

Campos, A., S. Tedesco, V. Vasconcelos, and S. Cristobal. 2012. Proteomic research in bivalves: towards the identification of molecular markers of aquatic pollution. J Proteomics 75:4346-4359.

Cardozo, K. H., T. Guaratini, M. P. Barros, V. R. Falcao, A. P. Tonon, N. P. Lopes, S. Campos, M. A. Torres, A. O. Souza, P. Colepicolo, and E. Pinto. 2007. Metabolites from algae with economical impact. Comp Biochem Physiol C Toxicol Pharmacol 146:60-78.

Carvalho, G. A., P. J. Minnett, V. F. Banzon, W. Baringer, and C. A. Heil. 2011. Long-term evaluation of three satellite ocean color algorithms for identifying harmful algal blooms (*Karenia brevis*) along the west coast of Florida: A matchup assessment. Remote Sens Environ 115:1-18.

Cheema, M. S. and J. Ausio. 2015. The Structural Determinants behind the Epigenetic Role of Histone Variants. Genes (Basel) 6:685-713.

Chen, T., T. D. Williams, A. Mally, C. Hamberger, L. Mirbahai, K. Hickling, and J. K. Chipman. 2012. Gene expression and epigenetic changes by furan in rat liver. Toxicology 292:63-70.

Collin, H., A. L. Meistertzheim, E. David, D. Moraga, and I. Boutet. 2010. Response of the Pacific oyster *Crassostrea gigas*, Thunberg 1793, to pesticide exposure under experimental conditions. J Exp Biol 213:4010-4017.

Cook, P. F., C. Reichmuth, A. A. Rouse, L. A. Libby, S. E. Dennison, O. T. Carmichael, K. T. Kruse-Elliott, J. Bloom, B. Singh, V. A. Fravel, L. Barbosa, J. J. Stuppino, W. G. Van Bonn, F. M. D. Gulland, and C. Ranganath. 2015. Algal toxin impairs sea lion memory and hippocampal connectivity, with implications for strandings. Science 350:1545-1547.

Cortessis, V. K., D. C. Thomas, A. J. Levine, C. V. Breton, T. M. Mack, K. D. Siegmund, R. W. Haile, and P. W. Laird. 2012. Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. Hum Genet 131:1565-1589.

Curry, R. and C. Mauritzen. 2005. Dilution of the northern North Atlantic Ocean in recent decades. Science 308:1772-1774.

Dame, R. F. 1972. The ecological energies of growth, respiration, and assimilation in the intertidal American oyster *Crassostrea virginica*. Mar. Biol. 17:243-250.

Davis, C. 1948. *Gymnodinium breve*: a cause of discolored water and animal mortality in the Gulf of Mexico. Bot. Gaz. 109:358-360.

Diaz-Freije, E., C. Gestal, S. Castellanos-Martinez, and P. Moran. 2014. The role of DNA methylation on *Octopus vulgaris* development and their perspectives. Front Physiol 5.

Dimond, J. L. and S. B. Roberts. 2016. Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change. Mol Ecol 25:1895-1904.

Eirín-López, J. M., R. González-Romero, D. Dryhurst, J. Méndez, and J. Ausió. 2009. Long-term evolution of histone families: old notions and new insights into their diversification mechanisms across eukaryotes. Pages 139-162 *in* P. Pontarotti, editor. Evolutionary Biology: Concept, Modeling, and Application. Springer-Verlag, Berlin Heidelberg.

Errera, R. M., S. Yvon-Lewis, J. D. Kessler, and L. Campbell. 2014. Reponses of the dinoflagellate *Karenia brevis* to climate change: pCO2 and sea surface temperatures. Harmful Algae 37:110-116.

Etchegaray, J. P. and R. Mostoslavsky. 2016. Interplay between Metabolism and Epigenetics: A Nuclear Adaptation to Environmental Changes. Molecular Cell 62:695-711.

Fang, X., C. Thornton, B. E. Scheffler, and K. L. Willett. 2013. Benzo[a]pyrene decreases global and gene specific DNA methylation during zebrafish development. Environ Toxicol Pharmacol 36:40-50.

Feil, R. and M. F. Fraga. 2012. Epigenetics and the environment: emerging patterns and implications. Nat Rev Genet 13:97-109.

Fernandez, A. F., E. G. Torano, R. G. Urdinguio, A. G. Lana, I. A. Fernandez, and M. F. Fraga. 2014. The epigenetic basis of adaptation and responses to environmental change: perspective on human reproduction. Adv Exp Med Biol 753:97-117.

Fernandez-Tajes, J., A. Arias-Perez, M. Fernandez-Moreno, and J. Mendez. 2012. Sharp decrease of genetic variation in two Spanish localities of razor clam *Ensis siliqua*: natural fluctuation or Prestige oil spill effects? Ecotoxicology 21:225-233.

Fernandez-Tajes, J., M. Gaspar, D. Martinez-Patino, N. McDonough, D. Roberts, A. M. González-Tizón, A. Martinez-Lage, and J. Mendez. 2007. Genetic variation of the razor clam *Ensis siliqua* (Jeffreys, 1875) along the European coast based on random amplified polymorphic DNA markers. Aquac. Res. 38:1205-1212.

Finn, R. M., K. Browne, K. C. Hodgson, and J. Ausio. 2008. sNASP, a histone H1-specific eukaryotic chaperone dimer that facilitates chromatin assembly. Biophys J 95:1314-1325.

Fleming, L. E., B. Kirkpatrick, L. C. Backer, J. A. Bean, A. Wanner, A. Reich, J. Zaias, Y. S. Cheng, R. Pierce, J. Naar, W. M. Abraham, and D. G. Baden. 2007. Aerosolized red-tide toxins (brevetoxins) and asthma. Chest 131:187-194.

Flewelling, L. J., J. P. Naar, J. P. Abbott, D. G. Baden, N. B. Barros, G. D. Bossart, M. Y. Bottein, D. G. Hammond, E. M. Haubold, C. A. Heil, M. S. Henry, H. M. Jacocks, T. A. Leighfield, R. H. Pierce, T. D. Pitchford, S. A. Rommel, P. S. Scott, K. A. Steidinger, E. W. Truby, F. M. Van Dolah, and J. H. Landsberg. 2005. Brevetoxicosis: red tides and marine mammal mortalities. Nature 435:755-756.

Gavery, M. R. and S. B. Roberts. 2010. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). BMC Genomics 11:483.

Gavery, M. R. and S. B. Roberts. 2013. Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. PeerJ 1.

Gavery, M. R. and S. B. Roberts. 2014. A context dependent role for DNA methylation in bivalves. Brief Funct Genomics 13:217-222.

Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80.

González-Romero, R., C. Rivera-Casas, L. J. Frehlick, J. Méndez, J. Ausió, and J. M. Eirín-López. 2012. Histone H2A (H2A.X and H2A.Z) variants in molluscs: molecular characterization and potential implications for chromatin dynamics. PLoS ONE 7:e30006.

Gu, Z., R. Eils, and M. Schlesner. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32:2847-2849.

Henikoff, S. and M. M. Smith. 2015. Histone variants and epigenetics. Cold Spring Harb Perspect Biol 7:a019364.

Hoffmann, A. A. and Y. Willi. 2008. Detecting genetic responses to environmental change. Nat Rev Genet 9:421-432.

Ivashkevich, A., C. E. Redon, A. J. Nakamura, R. F. Martin, and O. A. Martin. 2012. Use of the gamma-H2AX assay to monitor DNA damage and repair in translational cancer research. Cancer Lett 327:123-133.

Khurana, S., M. J. Kruhlak, J. Kim, A. D. Tran, J. Liu, K. Nyswaner, L. Shi, P. Jailwala, M. H. Sung, O. Hakim, and P. Oberdoerffer. 2014. A macrohistone variant links dynamic chromatin compaction to BRCA1-dependent genome maintenance. Cell Rep 8:1049-1062.

Kinner, A., W. Wu, C. Staudt, and G. Iliakis. 2008. Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. Nucleic Acids Res 36:5678-5694.

Kouzarides, T. 2007. Chromatin modifications and their function. Cell 128:693-705.

Kuo, L. J. and L. X. Yang. 2008. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. In Vivo 22:305-309.

Lee, C. S., K. Lee, G. Legube, and J. E. Haber. 2014. Dynamics of yeast histone H2A and H2B phosphorylation in response to a double-strand break. Nat Struct Mol Biol 21:103-109.

Leverone, J. R., N. J. Blake, R. H. Pierce, and S. E. Shumway. 2006. Effects of the dinoflagellate *Karenia brevis* on larval development in three species of bivalve mollusc from Florida. Toxicon 48:75-84.

Li, A., J. M. Eirín-López, and J. Ausió. 2005. H2AX: tailoring histone H2A for chromatin-dependent genomic integrity. Biochem. Cell Biol. 83:505-515.

Lopez, M. F., J. Tollervey, B. Krastins, A. Garces, D. Sarracino, A. Prakash, M. Vogelsang, G. Geesman, A. Valderrama, I. K. Jordan, and V. V. Lunyak. 2012. Depletion of nuclear histone H2A variants is associated with chronic DNA damage signaling upon drug-evoked senescence of human somatic cells. Aging (Albany NY) 4:823-842.

Luchmann, K. H., J. J. Mattos, M. N. Siebert, T. S. Dorrington, G. Toledo-Silva, P. H. Stoco, E. C. Grisard, and A. C. Bainy. 2012. Suppressive subtractive hybridization libraries prepared from the digestive gland of the oyster *Crassostrea brasiliana* exposed to a diesel fuel water-accommodated fraction. Environ Toxicol Chem 31:1249-1253.

Malik, H. S. and S. Henikoff. 2003. Phylogenomics of the nucleosome. Nat. Struct. Biol. 10:882-891.

Mat, A. M., H. Haberkorn, J. P. Bourdineaud, J. C. Massabuau, and D. Tran. 2013. Genetic and genotoxic impacts in the oyster *Crassostrea gigas* exposed to the harmful alga *Alexandrium minutum*. Aquat Toxicol 140-141:458-465.

McFarren, E. F., F. J. Silva, H. Tanabe, W. B. Wilson, J. E. Campbell, and K. H. Lewis. 1965. The occurrence of a ciguatera-like poison in oysters, clams, and *Gymnodinium breve* cultures. Toxicon 3:111-123.

Mello, D. F., E. S. de Oliveira, R. C. Vieira, E. Simoes, R. Trevisan, A. L. Dafre, and M. A. Barracco. 2012. Cellular and transcriptional responses of *Crassostrea gigas* hemocytes exposed in vitro to brevetoxin (PbTx-2). Mar Drugs 10:583-597.

Mercer, T. R. and J. S. Mattick. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol 20:300-307.

Migicovsky, Z. and I. Kovalchuk. 2011. Epigenetic memory in mammals. Front Genet 2:28.

Milan, M., M. Pauletto, T. Patarnello, L. Bargelloni, M. G. Marin, and V. Matozzo. 2013. Gene transcription and biomarker responses in the clam *Ruditapes philippinarum* after exposure to ibuprofen. Aquat Toxicol 126:17-29.

Mirbahai, L. and J. K. Chipman. 2014. Epigenetic memory of environmental organisms: a reflection of lifetime stressor exposures. Mutat Res Genet Toxicol Environ Mutagen 764-765:10-17.

Morrison, A. J. and X. Shen. 2005. DNA repair in the context of chromatin. Cell Cycle 4:568-571.

Murrell, R. N. and J. E. Gibson. 2009. Brevetoxins 2, 3, 6, and 9 show variability in potency and cause significant induction of DNA damage and apoptosis in Jurkat E6-1 cells. Arch Toxicol 83:1009-1019.

Murrell, R. N. and J. E. Gibson. 2011. Brevetoxin 2 alters expression of apoptotic, DNA damage, and cytokine genes in Jurkat cells. Hum Exp Toxicol 30:182-191.

Olson, C. E. and S. B. Roberts. 2014. Genome-wide profiling of DNA methylation and gene expression in *Crassostrea gigas* male gametes. Front Physiol 5:224.

Papamichos-Chronakis, M. and C. L. Peterson. 2013. Chromatin and the genome integrity network. Nat Rev Genet 14:62-75.

Perez-Figueroa, A. 2013. msap: a tool for the statistical analysis of methylation-sensitive amplified polymorphism data. Mol Ecol Resour 13:522-527.

Perrault, J. R., J. R. Schmid, C. J. Walsh, J. E. Yordy, and A. D. Tucker. 2014. Brevetoxin exposure, superoxide dismutase activity and plasma protein electrophoretic profiles in wild-caught Kemp's ridley sea turtles (*Lepidochelys kempii*) in southwest Florida. Harmful Algae 37:194-202.

Pfaffl, M. W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45.

Plakas, S. M., K. R. El Said, E. L. Jester, H. R. Granade, S. M. Musser, and R. W. Dickey. 2002. Confirmation of brevetoxin metabolism in the Eastern oyster (*Crassostrea virginica*) by controlled exposures to pure toxins and to *Karenia brevis* cultures. Toxicon 40:721-729.

Plakas, S. M., E. L. Jester, K. R. El Said, H. R. Granade, A. Abraham, R. W. Dickey, P. S. Scott, L. J. Flewelling, M. Henry, P. Blum, and R. Pierce. 2008. Monitoring of brevetoxins in the *Karenia brevis* bloom-exposed Eastern oyster (*Crassostrea virginica*). Toxicon 52:32-38.

Prego-Faraldo, M. V., V. Valdiglesias, B. Laffon, J. M. Eirin-Lopez, and J. Mendez. 2015. *In vitro* analysis of early genotoxic and cytotoxic effects of okadaic acid in different cell types of the mussel *Mytilus galloprovincialis*. J Toxicol Environ Health A 78:814-824.

Prego-Faraldo, M. V., V. Valdiglesias, B. Laffon, J. Mendez, and J. M. Eirin-Lopez. 2016. Early Genotoxic and Cytotoxic Effects of the Toxic Dinoflagellate *Prorocentrum lima* in the Mussel *Mytilus galloprovincialis*. Toxins (Basel) 8:1-14.

Ptashne, M. 2007. On the use of the word 'epigenetic'. Curr Biol 17:R233-236.

Radwan, F. F. and J. S. Ramsdell. 2008. Brevetoxin forms covalent DNA adducts in rat lung following intratracheal exposure. Environ Health Perspect 116:930-936.

Reyna-Lopez, G. E., J. Simpson, and J. Ruiz-Herrera. 1997. Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. Mol Gen Genet 253:703-710.

Rivera-Casas, C., R. Gonzalez-Romero, M. S. Cheema, J. Ausio, and J. M. Eirin-Lopez. 2016a. The characterization of macroH2A beyond vertebrates supports an ancestral origin and conserved role for histone variants in chromatin. Epigenetics 11:415-425.

Rivera-Casas, C., R. Gonzalez-Romero, A. Vizoso-Vazquez, M. S. Cheema, M. E. Cerdan, J. Mendez, J. Ausio, and J. M. Eirin-Lopez. 2016b. Characterization of mussel H2A.Z.2: a new H2A.Z variant preferentially expressed in germinal tissues from *Mytilus*. Biochem Cell Biol 94:480-490.

Riviere, G., G. C. Wu, A. Fellous, D. Goux, P. Sourdaine, and P. Favrel. 2013. DNA methylation is crucial for the early development in the Oyster *C. gigas*. Mar Biotechnol (NY) 15:739-753.

Rolton, A., J. Vignier, P. Soudant, S. E. Shumway, V. M. Bricelj, and A. K. Volety. 2014. Effects of the red tide dinoflagellate, *Karenia brevis*, on early development of the eastern oyster *Crassostrea virginica* and northern quahog *Mercenaria mercenaria*. Aquat Toxicol 155:199-206.

Romero-Geraldo, R., N. Garcia-Lagunas, and N. Y. Hernandez-Saavedra. 2016. *Crassostrea gigas* exposure to the dinoflagellate *Prorocentrum lima*: Histological and gene expression effects on the digestive gland. Mar Environ Res 120:93-102.

Romero-Geraldo, R. D. J., N. Garcia-Lagunas, and N. Y. Hernandez-Saavedra. 2014. Effects of In Vitro Exposure to Diarrheic Toxin Producer *Prorocentrum lima* on Gene Expressions Related to Cell Cycle Regulation and Immune Response in *Crassostrea gigas*. PLoS ONE 9.

Ross, C., R. Ritson-Williams, R. Pierce, J. B. Bullington, M. Henry, and V. J. Paul. 2010. Effects of the Florida red tide dinoflagellate, *Karenia brevis*, on oxidative stress and metamorphosis of larvae of the coral *Porites astreoides*. Harmful Algae 9:173-179.

Schulmeister, A., M. Schmid, and E. M. Thompson. 2007. Phosphorylation of the histone H3.3 variant in mitosis and meiosis of the urochordate *Oikopleura dioica*. Chromosome Res 15:189-201.

Shi, L. and P. Oberdoerffer. 2012. Chromatin dynamics in DNA double-strand break repair. Biochim Biophys Acta 1819:811-819.

Stumpf, R. P., M. E. Culver, P. A. Tester, M. Tomlinson, G. J. Kirkpatrick, B. A. Pederson, E. Truby, V. Ransibrahmanakul, and M. Soracco. 2003. Monitoring *Karenia brevis* blooms in the Gulf of Mexico using satellite ocean color imagery and other data. Harmful Algae 2:147-160.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, C. Rivera-Casas, R. Gonzalez-Romero, J. Ausio, J. Mendez, J. Dorado, and J. M. Eirin-Lopez. 2013. The CHROMEVALOA database: a resource for the evaluation of okadaic acid contamination in the marine environment based on the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis*. Mar Drugs 11:830-841.

Suarez-Ulloa, V., R. Gonzalez-Romero, and J. M. Eirin-Lopez. 2015. Environmental epigenetics: A promising venue for developing next-generation pollution biomonitoring tools in marine invertebrates. Mar Pollut Bull 98:5-13.

Sun, P., C. Leeson, X. Zhi, F. Leng, R. H. Pierce, M. S. Henry, and K. S. Rein. 2016. Characterization of an epoxide hydrolase from the Florida red tide dinoflagellate, *Karenia brevis*. Phytochemistry 122:11-21.

Sunda, W. G., C. Burleson, D. R. Hardison, J. S. Morey, Z. Wang, J. Wolny, A. A. Corcoran, L. J. Flewelling, and F. M. Van Dolah. 2013. Osmotic stress does not trigger brevetoxin production in the dinoflagellate *Karenia brevis*. Proc Natl Acad Sci U S A 110:10223-10228.

Talbert, P. B. and S. Henikoff. 2010. Histone variants--ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol 11:264-275.

Talbert, P. B. and S. Henikoff. 2014. Environmental responses mediated by histone variants. Trends Cell Biol.

Turinetto, V. and C. Giachino. 2015. Multiple facets of histone variant H2AX: a DNA double-strand-break marker with several biological functions. Nucleic Acids Res 43:2489-2498.

Twiner, M. J., M. Y. Bottein Dechraoui, Z. Wang, C. M. Mikulski, M. S. Henry, R. H. Pierce, and G. J. Doucette. 2007. Extraction and analysis of lipophilic brevetoxins from the red tide dinoflagellate *Karenia brevis*. Anal Biochem 369:128-135.

Valdiglesias, V., J. Fernandez-Tajes, J. Mendez, E. Pasaro, and B. Laffon. 2013. The marine toxin okadaic acid induces alterations in the expression level of cancer-related genes in human neuronal cells. Ecotoxicol Environ Saf 92:303-311.

Vandegehuchte, M. B., F. Lemiere, and C. R. Janssen. 2009. Quantitative DNA-methylation in *Daphnia magna* and effects of multigeneration Zn exposure. Comp Biochem Physiol C Toxicol Pharmacol 150:343-348.

Vandegehuchte, M. B., F. Lemiere, L. Vanhaecke, W. Vanden Berghe, and C. R. Janssen. 2010. Direct and transgenerational impact on *Daphnia magna* of chemicals with a known effect on DNA methylation. Comp Biochem Physiol C Toxicol Pharmacol 151:278-285.

Walsh, C. J., M. Butawan, J. Yordy, R. Ball, L. Flewelling, M. de Wit, and R. K. Bonde. 2015. Sublethal red tide toxin exposure in free-ranging manatees (*Trichechus manatus*) affects the immune system through reduced lymphocyte proliferation responses, inflammation, and oxidative stress. Aquat Toxicol 161:73-84.

Weber, C. M. and S. Henikoff. 2014. Histone variants: dynamic punctuation in transcription. Genes Dev 28:672-682.

Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134.

Zhang, L., L. Li, Y. Zhu, G. Zhang, and X. Guo. 2014. Transcriptome analysis reveals a rich gene set related to innate immunity in the Eastern oyster (*Crassostrea virginica*). Mar Biotechnol (NY) 16:17-33.

Figure 1. Florida Red Tide HAB simulation. A, Schematic representation of HAB simulation experiment, including cell concentration (cell/mL) of the brevetoxin-producing dinoflagellate *K. brevis* over time. Samples (n=2 oysters per biological replicate) were collected at different intervals (T0, 0 h; T1, 3 h; T2, 5 h; T3, 24 h) during the simulation. B, Monitoring of *K. brevis* cell concentrations during HAB simulation. Final cell concentrations used in the present study (5, 50, 100 and 1000 cell/mL) are represented by dotted lines. Each data point represents the average *K. brevis* cell count across biological replicates along with the corresponding standard error in water samples (grey squares), measured after application of *K. brevis* cell cultures to experimental tanks.

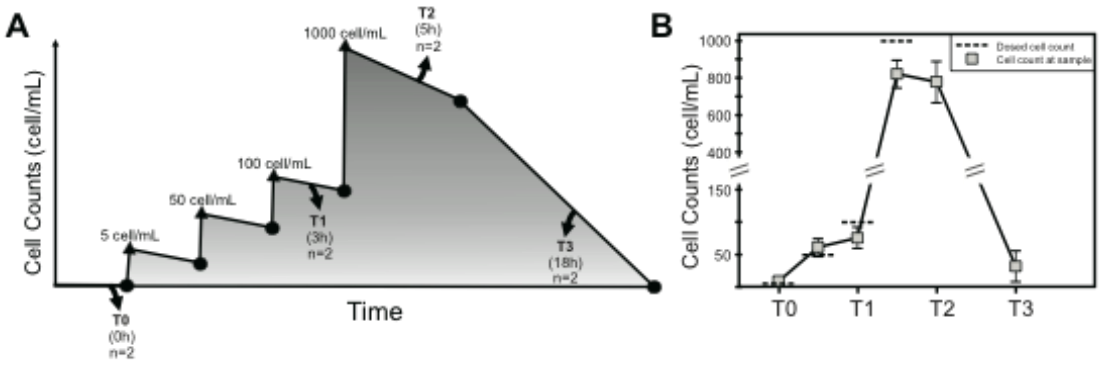


Figure 2. Gene expression of histone variants during exposure to *K. brevis*. Boxplots represent the quartile distribution of the expression levels of Eastern oyster genes encoding histone variants H2A.X (A), H2A. Z (B) and macroH2A (C). Results are categorized by exposure time (T0-T3) and represented as normalized ratios respect to the study calibrator (i.e., gene expression at T0). Boxes and whiskers show a large dispersion in the data for each group of samples, and no significant differences between groups.

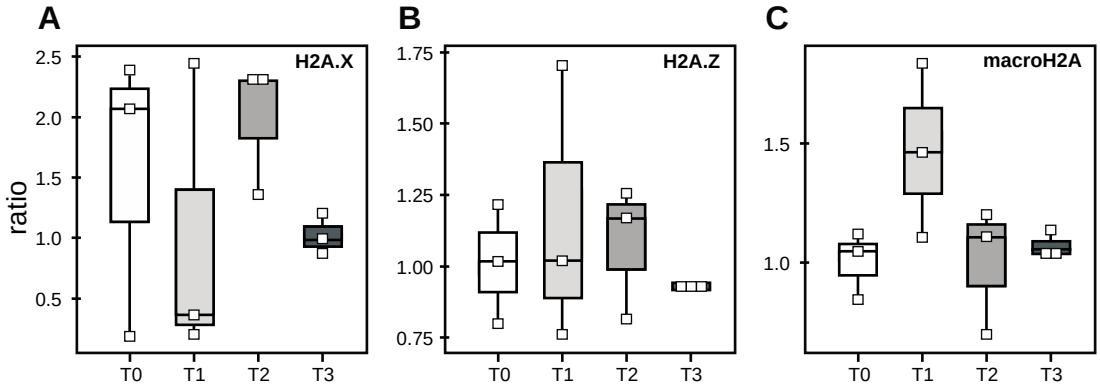


Figure 3. Histone variant protein expression and H2A.X phosphorylation (γH2A.X) during exposure to *K. brevis*. A, SDS gel showing normalized histones extracted from Eastern oyster individuals at different biological replicates (*r1*, *r2*, *r3*) and at different time points (T0, T1, T2, T3). B, Western blot hybridization revealing homogeneous levels of H2A.X, H2A.Z and macroH2A proteins throughout HAB simulation. C, Western blot hybridization showing an increase in γH2A.X formation concomitantly with exposure to increasing *K. brevis* concentrations (T1, T2), followed by a slight decrease during the recovery phase (T3). An H4 antibody was used for normalization purposes. Gels in different boxes denote independent hybridization experiments. M, ClearPAGE Two-Color Marker (C.B.S. SCIENTIFIC)

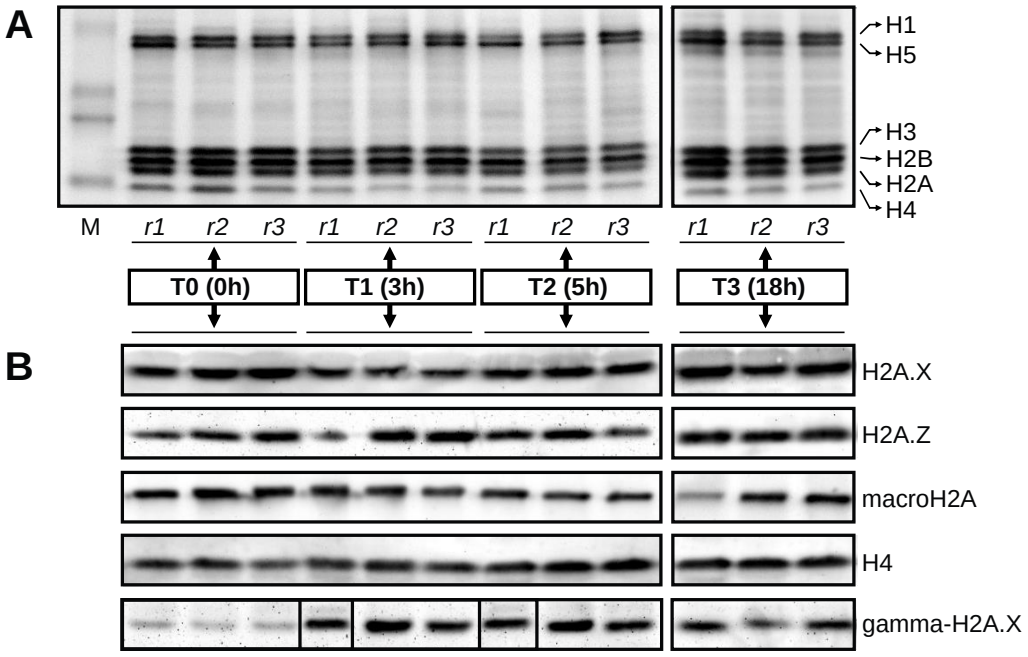


Figure 4. Global DNA methylation changes during HAB simulation. A) Principal Component Analysis of complete MSAP profiles representing the different treatment groups (i.e., time points), labeled in the centroid of each cluster (T0-T4). The first two principal components are shown, indicating the percentage of the global variance explained on the corresponding axis. Individual samples are represented as points and the variance within each group is represented with an ellipse. Results show that earlier (T0, T1) and later time points (T2, T3) segregate through the C1=0 axis, corresponding to differentiated genome-wide methylation patterns. B) Heatmap representing changes in DNA methylation in a group of 10 loci showing a non-random distribution of DNA methylation patterns ($p < 0.05$) throughout the HAB simulation. Methylation profiles observed at T0 and T1 display significant differences from patterns observed at T2 and T3. Rows (specimens) and columns (MSAP loci) were clustered using Gower's Coefficient of Similarity. Loci methylation status is indicated in the right margin of the figure: HMM, hemimethylated, HPM, hypermethylated, Internal Cytosine Methylated (ICM) or Unmethylated (NMT).

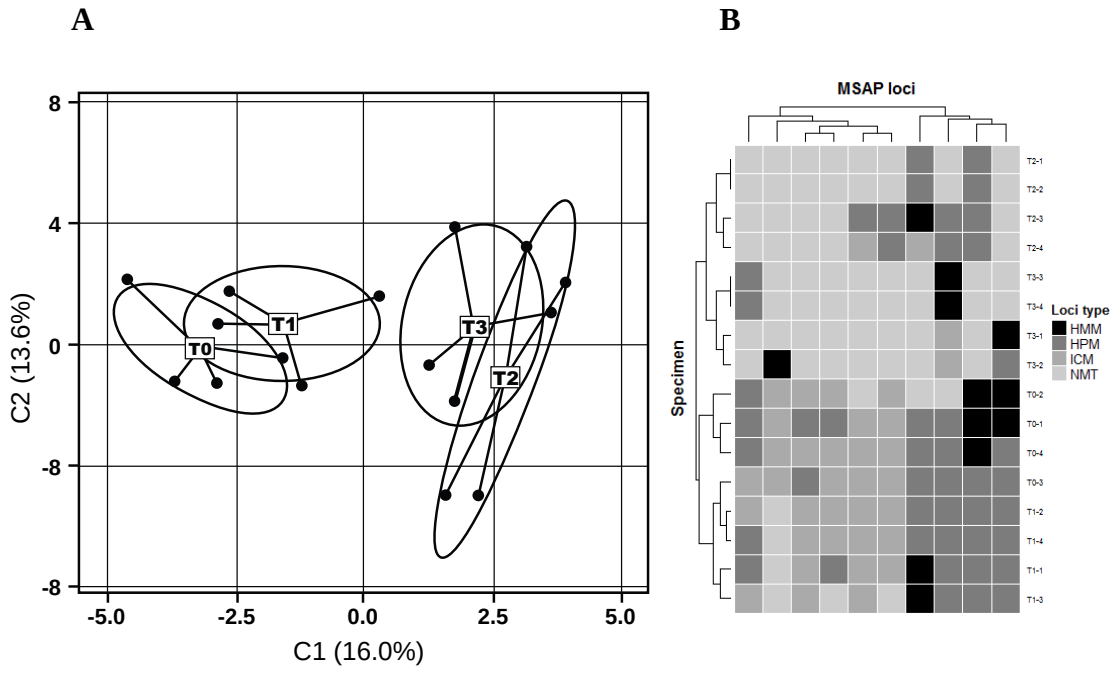


Fig. 4, Gonzalez-Romero et al. 2016

Table 1. qPCR primers used in gene expression analyses, specifically designed to amplify histone variant genes and reference genes in Eastern oyster.

Gene	Primer Name	Sequence (5' → 3')
H2A.X	Cv-H2A.X-Fw	AGTTACCATTGCCCAAGGAGG
	Cv-H2A.X-Rv	AAAATTCCTGGGACTGTGACGA
H2A.Z	Cv-H2A.Z-Fw	CGCCATCAGAGGAGACGAAG
	Cv-H2A.Z-Rv	AGCTGTTTTCTGTGTGCCCT
MacroH2A	Cv-mH2A-Fw	TCATTTCCGTATCGGAGCGG
	Cv-mH2A-Rv	CTCTTGCAGCATTTCCAGCC
GAPDH	Cv-GAPDH-Fw	GACAACAGTCCACGCCTACA
	Cv-GAPDH-Rv	GGATGACCTTACCCACTGCC
RPL13	Cv-RPL13-Fw	CCGGGCTCCCAGTAAAATGT
	Cv-RPL13-Rv	TGTCGTATGGGGGAGGGATT

Table 2. Frequency (%) of different methylation states at target sequences across different time points.

Band pattern	Methylation state	Time points			
		T0	T1	T2	T3
HPA+/MSP+	Unmethylated	0.1017	0.1085	0.2017	0.1856
HPA+/MSP-	Hemimethylated	0.0958	0.0991	0.1576	0.1763
HPA-/MSP+	Internal Cytosine methylation	0.1059	0.1525	0.0898	0.1398
HPA-/MSP-	Full methylation or absence of target	0.6966	0.6398	0.5508	0.4983

Table 3. Pairwise AMOVAs between exposure time points based on complete MSAP methylation profiles. ΦST values and π-values are indicated in the upper and lower diagonal, respectively. * $p < 0.05$.

	T0	**T1**	**T2**	**T3**
T0		-0.0087	0.3543	0.4131
T1	0.6489		0.4188	0.4167
T2	0.0311*	0.0255*		0.3333
T3	0.0291*	0.0284*	0.0290*	

CHAPTER V

TRANSCRIPTOMIC PATTERNS IN THE RESPONSE OF THE PACIFIC OYSTER *CRASSOSTREA GIGAS* TO ENVIRONMENTAL STRESS REVEALED BY CO-EXPRESSION NETWORK ANALYSIS

In order to understand what transcriptomic patterns lack the ability to inform about the corresponding triggering environmental conditions, the hypothesis of a core or minimal cellular response to environmental stress is revisited. Publicly available RNA-Seq data series from the oyster *Crassostrea gigas* exposed to an array of different stressors were analyzed using co-expression network analysis methods. A targeted analysis of genes with epigenetic relevance (i.e., chromatin-associated genes) is also included, assessing their potential as biomarkers of environmental stress.

This work was carried out with the collaboration of Camilo Valdes[1], Vanessa Aguiar Pulido[2], Giri Narasimhan[1] and Jose M. Eirin Lopez[1], from Florida International University (U.S.A.)[1] and Cornell University (U.S.A.)[2].

Abstract

The current post-genomic era is characterized by a dramatic increase in the production of transcriptomic data that can be used in meta-analyses, looking for consistent patterns of gene expression that provide a better understanding of how organisms respond to conditions like environmental stress. In particular, time series and stress-level series offer unique opportunities to make observations about the dynamics of these responses. With this goal in mind, co-expression network analyses represent a valuable approach given their ability to synthesize large amounts of information and to unravel hidden patterns.

The present work aims to provide a better understanding of the molecular strategies that the oyster *Crassotrea gigas* commonly uses to cope with changing environments and stress. Such information is critical to distinguish those changes in

expression and altered biological processes that are specific in the response to a particular environmental condition from those that are commonly responsive in circumstances of general stress. With this goal in mind, different environmental variables are included in the study such as temperature, salinity, exposure to the heavy metal Zinc and exposure to a pathogen of the *Vibrio* species. The sets of differentially expressed transcripts in gill tissue are analyzed showing a very small overlap. However, general similarities in their functional profiles support the notion of a core response to stress. Co-expression network analysis of time-series further reveals a dynamic response to stress allowing the chronological organization of responsive transcript modules.

Additionally, the potential of chromatin-associated genes as stress biomarkers is also assessed using the co-expression network approach, leading to the unexpected discovery of a member of the H1 histone family that displays relatively high and variable levels of expression under different stressors. These findings support the value of network analysis to promote new hypotheses and convey relevant insights for the development of stress biomarkers in the oyster *C. gigas*.

Introduction

Transcriptomic studies have contributed to the identification of specific groups of genes involved in the response and adaptation of bivalves to external conditions (e.g., immune response as a determinant of the resistance to summer mortality in *C. gigas*) (Chaney and Gracey 2011). Indeed, sequencing and data mining of ESTs are essential steps for the comparative identification of molecules and related pathways of response to specific stimuli.

Nowadays, this task is greatly facilitated by the availability of high throughput sequencing technologies, yielding unprecedented amounts of sequence data. However, managing massive datasets requires appropriate computational strategies. This kind of approach enables transforming data into information and generating new knowledge through human interpretation. Information can be extracted looking for patterns across multiple datasets using data mining techniques and visualization methods. Regarding this aspect, network-based analysis is capable of intuitively displaying not only the most relevant elements in a process (e.g., genes or transcripts) but also the relationships between them, representing a valuable tool for the interpretation of massive datasets (Ucar et al. 2007).

A common application of these computational methods to transcriptomic data is co-expression network analysis (Zhang and Horvath 2005). This approach calculates pairwise correlations of gene expression profiles across time or different conditions to build a correlation matrix that can be graphically represented as a network. Those genes that show strong co-expression (positive correlation) or those that display an opposite behavior (negative correlation) are clustered together, thus highlighting potentially relevant relationships between different genes and biological processes. This approach has been previously applied to clinical research in humans (Ren et al. 2015) as well as to environmental studies using different organisms (Schneider et al. 2014), most notably plants (Barah et al. 2013, Jung et al. 2013, Priest et al. 2014). In the case of bivalve molluscs, co-expression network analysis has also been successfully applied to microarray data of oysters following their recovery from heat shock, pinpointing key genes in the cascade response of oysters to thermal stress (Zhang et al. 2012b).

Although genome-wide co-expression network analysis is technically possible, it is a common practice to reduce the size of the correlation matrix by selecting a subset composed of genes of particular interest, which allows a better visualization of results. Such subset usually comprises those genes that show a statistically significant change in expression (Tantai et al. 2015, Safari-Alighiarloo et al. 2016). Alternatively, the focus may be placed on specific protein families expected to play a major role in the stimulus or conditions under study (Jung et al. 2013, Tang et al. 2013).

The present work describes the network-based analysis of time-series and stress-level series RNA-Seq data of the Pacific oyster *Crassostrea gigas* obtained from publicly available transcriptomic datasets. The relationships between gene expression profiles among differentially expressed transcripts as well as chromatin-associated transcripts with a role in epigenetic mechanisms (i.e., histones, histone variants and histone-modifying enzymes) unravel patterns of expression of great value for the development of biomarkers of environmental stress.

The analysis of differentially expressed transcripts represents the most common approach in transcriptomic studies. Such genes display sharp changes in expression levels that inherently make them a great source of biomarkers. Furthermore, this feature also makes them ideal for co-expression network analyses since their expression profiles are accentuated and distinct in contrast with other genes that, despite their potential biological significance, may show nearly-flat profiles and in consequence, less informative correlations.

On the other hand, the rationale underlying the epigenetic approach postulates that changes in gene expression will be modulated by structural transitions in chromatin,

facilitating or preventing the access of the transcriptional machinery to the genetic information. Thus, changes in the expression of chromatin-associated genes could mirror changes in the environment, thus representing a promising venue for the development of biomarkers valuable in monitoring efforts (Suarez-Ulloa et al. 2015b).

An important aspect in the process of developing transcriptional biomarkers is the identification of the biological processes that may be core to any non-specific environmental stress and those that are strongly associated to a determined stressor. The meta-analysis of microarray data and Suppression Subtractive Hybridization (SSH) libraries across different environmental conditions has resulted in a consensus model for a general environmental stress response in oysters, pinpointing biological processes involved in the responses to a variety of stressors (Anderson et al. 2015). In this regard, the comparison of networks constructed from RNA-Seq data characterizing transcriptional responses under different environmental conditions provides further insights into the core response of oysters under environmental stress.

Methods

Data collection

Datasets were retrieved from SRA database (NCBI) after an exhaustive search for RNA-Seq Bioprojects containing data from oysters of the species *Crassostrea gigas*, which has the most developed genome assembly available among bivalve molluscs. Among the identified Bioprojects, those containing series data produced in experiments with environmental stressors with consistent technology (Illumina, single-end) were selected. The Bioproject PRJNA146329, contains series datasets from *C.gigas* exposed

salinity, temperature and heavy metal pollution (Zn, 1 mg/L) (Zhang et al. 2012a). The Bioproject PRJNA194079 complements this set of abiotic stressors with data from an experiment with *C. gigas* exposed to pathogen challenge with *Vibrio* spp. (study not published). All RNA extracts were obtained from gill tissue while the heavy metal exposure experiment includes an additional series obtained from digestive gland. Tissue from 3 individuals was pooled for the different experiments within Bioproject PRJNA146329, while no details are accessible for samples exposed to *Vibrio* challenge (PRJNA194079). A summarized description of these datasets, including the correspondence of samples with time points or stressor levels, is included in Table 1.

Data processing

Each SRA dataset was downloaded from the NCBI repository and converted to fastq format using the SRA-toolkit. As a reference for alignment of the reads, the genome assembly of *Crassostrea gigas* GCA_000297895.1 was obtained from the Ensembl repository (metazoa.ensembl.org/Crassostrea_gigas) as well as the corresponding GFF3 annotation file. Alignment of single-end Illumina reads was carried out using Tophat 2.1.1 (Trapnell et al. 2009). Subsequent differential expression analysis based on fragments per kilobase of exon per million reads mapped (FPKM) metrics was calculated using Cuffdiff 2.2.1 (Trapnell et al. 2012). Full FPKM tables including all samples in an experimental series were used to construct the expression profiles and perform further analysis. Sets of differentially expressed transcripts were defined based on their p-value adjusted by the Benjamini-Hochberg method (FDR < 0.05) as long as significance is found in at least one of the samples of the series. Sets of chromatin-associated transcripts

were manually selected including annotated histones, histone variants, post-translational modifying enzymes and DNA methyltransferases, analyzed following the same downstream pipeline as the differentially expressed transcript sets.

Downstream analysis

FPKM tables were calculated for each experiment and used to build correlation matrices based on Pearson coefficients with a significance cutoff of 0.05. Clusters were defined using the Affinity Propagation method, with no requirement of estimated input parameters. These analysis steps as well as the construction of co-expression networks were carried out using the pipeline PLUMA (T. Cickovski 2016), adapted in this case for RNA-Seq applications. Generated network files were visualized using the software Cytoscape (Shannon et al. 2003).

The complete set of differentially expressed transcripts per experiment as well as selected clusters representative of the most condensed areas in the corresponding networks (i.e., transcript modules) were further analyzed using the analysis suite Blast2GO (Gotz et al. 2008) and the Bioconductor-R package topGO using the Fisher's exact test for Gene Ontology enrichment (Alexa and Rahnenfuhrer 2010).

Targeted analysis of chromatin-associated genes was carried out using the same pipeline for genome mapping and network construction. Additionally, the set of H1 homologs was aligned using Clustal W (Thompson et al. 1994) and a maximum likelihood tree was constructed using the TreeBeST tool on the ENSEMBL web portal (Vilella et al. 2009).

Results and Discussion

Analysis of differentially expressed transcripts under different stressors

Transcripts were considered differentially expressed as long as they showed a significant difference in expression relative to the control in at least one of the samples/treatments of the data series. List of differentially expressed transcripts were compiled and compared between experiments showing a relatively low overlap between stressors (Figure 1). Stress by salinity yielded a total number 486 differentially expressed transcripts. Temperature stress produced a total of 327 differentially expressed transcripts. Exposure to the heavy metal Zn produced 332 and 297 in gill and digestive gland tissues respectively. Lastly, exposure to the pathogen *Vibrio* yielded a total of 419 differentially expressed transcripts. Although in comparable numbers, most of the differentially expressed transcripts were unique in the response to a specific stressor, representing a source of potential biomarkers.

In this analysis, a closer attention was paid to transcripts annotated as miRNA as well as to transcripts coding for chromatin-associated proteins, including histones, histone variants, histone-modifying enzymes and DNA methyltransferases. These subsets of transcripts are particularly interesting given their role in epigenetic mechanisms and gene expression regulation. However, none of the miRNA showed a significant difference in expression under any of the different stressors. Similarly, most chromatin-associated transcripts failed to show a significant difference in expression across conditions. Only one transcript annotated as histone H1-delta (EKC17653) was identified as differentially expressed under temperature stress, as discussed later in this chapter.

A comparison of across the sets of differentially expressed transcripts based on the distribution of ranked low-level Gene Ontology (GO) terms (i.e., direct descendants of the three main GO categories: Biological Process [BP], Molecular Function [MF] and Cellular Component [CC]) shows similar functional profiles regardless of the stressor in play (Figure 2). There is a consistent dominance of cellular and metabolic processes, closely followed by processes of biological regulation. Binding and catalytic activity represent the most common molecular functions for these sets of transcripts, with other miscellaneous group of GO terms following showing a much smaller representation. Similarly, the cellular component categories are dominated by the general GO terms "cell part", "cell" and "organelle" in all cases. Other less represented GO terms vary their position in the ranking while showing negligible differences in their representation levels. The consensus model for stress responses in oysters formulated by Anderson et al. highlights the relevance of an enhanced metabolism, transport mechanisms and protection versus oxidative damage, which is coherent with the present results. A closer look into the most represented GO terms in these transcript sets provides a more detailed interpretation for these results.

The enrichment analysis of differentially expressed transcripts for each one of the considered stressors produces a distinct ranked list of specific GO terms that are most representative of the respective sets (Table 2). These lists of top 20 enriched GO terms reveal important similarities given the small overlap found between sets of differentially expressed transcripts. Consistently with Anderson's published consensus model, processes related to oxidative stress are enriched under different stress conditions represented by terms such as "Oxidation-reduction process" in all four considered

stressors. This example and others like “Response to oxidative stress” can be seen in Table 2. The regulation of protein folding and degradation is represented across stressors with terms such as “Protein ubiquitination” in salinity and metal exposure and “Negative regulation of proteolysis” in pathogen challenge. Surprisingly, processes related to protein folding and repair, commonly associated with Heat-shock proteins, do not appear among the 20 most significant GO terms for temperature stress response. However, scrolling down in the ranked list, the terms “Protein ubiquitination”, “Response to oxidative stress”, “Proteolysis” and “Protein folding” respectively occupy positions 29, 30, 32 and 34.

Regarding Anderson’s observations about a prevalence of metabolism of carbohydrates, our general GO analysis is not primarily related to energy-producing metabolic pathways. Rather, our results show enrichment in processes related to the metabolism of aminoacids and nucleotides, consistent with an overall increase in regulation of gene and protein expression.

Of particular interest is the appearance of GO terms such as “Immune response” or “Defense response to virus”, indicating a close link between a general response to stress and an immune reaction. This observation had been reported in the transcriptomic analysis of the mussel *Mytilus galloprovincialis*, included as Chapter 3 of the present dissertation (Suarez-Ulloa et al. 2015a), suggesting that this link between stress and immune responses is common to other bivalve species and extends to other environmental stressors like the exposure to toxic Harmful Algal Blooms. In order to explain this phenomenon, we must consider the physiology of the neuroendocrine-immune regulatory system in bivalve molluscs. It has been recently suggested that

processes such as immune response, regulation of apoptosis, redox reactions and protein folding in oysters may be under a single regulatory network stemming from the MAPK pathway (Liu et al. 2016). The results obtained during the production of the present dissertation further support this interesting hypothesis.

Network analysis and clustering of gene expression profiles

The availability of data series allows correlation analyses between transcriptomic expression profiles. In this work, we have constructed networks representing the strength of pairwise correlations between differentially expressed genes under different stress challenges. The product is an undirected graph representing each transcript as a node and each significant correlation as an edge joining pairs of nodes (Figure 3).

The expression profiles for those clusters of transcripts (i.e., modules) show strong variability across time or stressor intensity demonstrating the highly dynamic nature of the transcriptome under environmental changes. In fact, these modules display a sequential organization of gene expression, unraveling groups of transcripts that respond at early time points versus others that respond later on (Figure 3).

Within well-organized clusters, most of these responses involve punctual changes in expression under specific experimental conditions or time points. Genes with profiles consistently up or down-regulated are a minority. This pattern is maintained throughout the different stressors considered in this meta-analysis. However, opposite to previous reports by Anderson et al. claiming a significant general up-regulation under stress conditions, our results fail to show a significant difference between the total number of up-regulated transcripts and the number of down-regulated transcripts (Table 3)

(Anderson et al. 2015). This observation is limited to the set of differentially expressed transcripts with statistical significance only but all the times or conditions where they may show such up or down regulation versus the control are included. This means that if some specific transcript shows up as differentially expressed under different temperatures, salinity levels or time points in exposure experiments, it will be counted multiple times. Based on these results there is no evidence to claim a net increase or decrease in gene expression levels under stress conditions.

Functional dynamics of transcript modules

The sequential response of different groups of genes at different time points or different stressor levels, as shown by the profiles depicted on Figure 3, allows an organization of associated biological functions. This approach is particularly useful in time series data (i.e., metal or pathogen exposure) and therefore the current section focuses on this type of data, providing insights into the activity of the cell as time progresses. A similar method has been previously used with microarray data of oysters recovering from a heat shock, supporting the benefits of network-based approaches in transcriptomic studies (Zhang et al. 2012b). Indeed, the understanding of the dynamics in the response to environmental stress provides a new informative dimension that becomes critical in the search for useful biomarkers.

Table 4 shows the order in which biological processes are acting after initial exposure to the heavy metal Zn in gill tissue. In summary, shortly after the exposure period begins, cellular and metabolic processes are triggered, particularly peptide synthesis, suggesting an increase in protein production. Subsequently, processes of

biological regulation become most distinct, being the negative regulation of endopeptidase activity the most significant. This suggests that the protein synthesis must still be maintained after several days of exposure. 7 days after exposure, while regulation of expression keeps being a defining trait, transport mechanisms become more relevant as well as the metabolism of carbohydrates, possibly to sustain the increased cellular activity with energy production. Lastly, after 13 days of exposure, the metabolism of the endogenous polycarbohydrate chitin becomes the most significant biological process. This result principally stems from the up-regulation of a transcript homolog to the mammalian Chitotriosidase-1, which in humans is associated with defense against pathogens (Kanneganti et al. 2012). This observation further supports the tight link between responses to environmental stress and the activation of immune mechanisms.

In the case of exposure to the pathogen *Vibrio* (Table 5), the biological processes enriched in the ordered transcript modules are low-level and relatively uninformative, showing a convergence between the initial and the later hours of exposure. However, processes of protein phosphorylation are highlighted one day after initial exposure. This is related to a protein annotated as Serine/threonine-protein kinase Nek1, which has been suggested to participate in DNA damage signaling independently from the canonical repair pathway involving ATM (Chen et al. 2011). In fact, ATM is not identified as differentially expressed under any of the stressors included in this study. However, the Serine/threonine-protein kinase Nek1 is restricted to the response to pathogen challenge, being surprisingly absent from the set of differentially expressed transcripts for heavy metal exposure, more likely to cause DNA damage (Stohs and Bagchi 1995).

Comparison of responses in gills and digestive gland tissues

The availability of data from the same experiment allows the comparison of the genes and associated functions that are responsive to heavy metal exposure (Zn, 1 mg/L) in different tissues: Gills and digestive gland. The number of differentially expressed transcripts in both cases is relatively close (i.e., 332 for gills and 297 for digestive gland), suggesting that both tissues are affected in a comparable manner. It is easy to expect that different tissues will show different responses expressing tissue-specific genes but it is still striking that the comparison of their corresponding functional profiles displays more differences between two tissues exposed to the same stressor than the the same tissue (gill) exposed to different stressors, as it can be noticed when comparing Figure 2 and Figure 4.

Several biological processes represented in the set of differentially expressed transcripts from gill are absent from that of the digestive gland. These processes include "developmental process", "reproductive process" and "reproduction", which rather than an actual difference in the stress response could actually suggest a contamination of the gill samples with gonad tissue. In fact, the previously discussed upregulation of the Serine/threonine-protein kinase Nek1 may support this conjecture, since this kinase is typically expressed in germ cells playing a relevant role in meiosis (Letwin et al. 1992). Further analysis will be required to clarify whether Nek1 is actively playing a role in the response to stress in gill tissue or whether this observation is an artifact.

Regarding molecular functions, the GO term "transporter activity" becomes more represented in the digestive gland than in gills, which is probably related with the high levels of bioaccumulation and lysosomal uptake of xenobiotics that occur primarily in the

lipid-rich digestive gland of molluscs (Svensson et al. 2003). This explanation also fits the observation of the GO term "membrane" becoming the most abundant annotation in the cellular component category. Different responses to metal exposure have been previously described in gills and digestive gland of the oyster *Pinctada fucata*, acknowledging the importance of having into account different times of exposure, the level of the stressor and the type of tissue in the development and use of biomarkers (Jing et al. 2006).

On the other hand, both tissues show the differential expression of transcripts related to immunity. Based on the observations made in this meta-analysis together with those from the transcriptomic analysis of the mussel *Mytilus* exposed to okadaic acid (Chapter 3 of this dissertation) (Suarez-Ulloa et al. 2015a), we propose that the incorporation of immune-related processes would nicely complement the consensus model proposed by Anderson et al. (Anderson et al. 2015).

Targeted analysis of chromatin-associated patterns

The focus on chromatin-associated genes as targets for co-expression network analysis aims to unravel the dynamics and relationships within a relatively self-contained system with a fundamental role in gene expression (Bannister and Kouzarides 2011). Whether this set of genes is able to mirror changes in the environment or not has been questioned in the present work, including specifically genes of histones, histone variants, histone-modifying enzymes and DNA methyltransferases in the analysis. This study complements the work carried out in the previous chapter of this dissertation (Chapter 4) where the epigenetic response of the oyster *C. virginica* was assessed under stress

conditions caused by a Florida Red Tide simulation using a targeted approach (Suarez Ulloa 2017). Both in that work, using quantitative PCR (qPCR) and Western blot for gene and protein expression analyses, as well as in the present meta-analysis, the main outcome has been insufficient evidence to claim a significant change in expression for histone variants such as H2A.X, H2A.Z and macroH2A, despite of their expected role in the response to environmental stress by exposure to a genotoxic agent (Radwan and Ramsdell 2008). However, the lack of statistical significance does not necessarily imply a lack of biological significance. On the contrary, blind confidence in hard p-value cutoffs might lead to low reproducibility, making the use of statistical significance thresholds a current matter of debate in the scientific community (Nuzzo 2014).

On the other hand, co-expression network analysis becomes limited when the changes in expression across samples are far from significant or even so low that fall within the error of the sequencing/processing method, leading to spurious significance in correlations based on nearly (or completely) flat expression profiles. The networks constructed focusing on the chromatin-associated genes of *C. gigas* exemplify this situation, since in multiple cases had FPKM values of 0 or close across experimental conditions or time points. However, the ability of network-based methods to synthesize information in an efficient visual manner was able to highlight the presence of genes that followed a relevant non-flat expression profile. Figure 5 shows the chromatin-focused network calculated from the temperature challenge dataset where the size of nodes has been made proportional to the level of expression and the thickness of the edges proportional to the calculated pairwise correlation coefficients.

The detailed observation of networks such as this one may lead to new questions and hypotheses that can be explored going back to the raw data. For instance, not all nodes annotated with the same gene name follow the same expression profile, since not all of them appear as correlated (i.e., connected) in the graph. A good example of this observation is nodes annotated as histone H1-delta. We find that not all H1-delta correlate to each other (see Figure 5), which indicates that they follow different expression profiles. Also, their different node sizes suggest that not all of them are expressed at the same level, with the H1-delta labeled as A in Figure 4 showing a higher level of expression than the other H1-delta genes. This observation is confirmed when the expression profiles of the genes represented in this network are plotted as a function of the changing experimental conditions under the different stressors (Figure 6), demonstrating that this particular H1-delta (gene: CGI_10000402, transcript: EKC17653) is consistently expressed at higher levels than any other histone or histone variant currently described for *C. gigas* genome across different environmental conditions. Furthermore, it seems to be responsive to changes in such experimental conditions or to the effect of time of exposure. Indeed, the H1-delta EKC17653 has been identified as differentially expressed with statistical significance (FDR < 0.05) at least under one of the temperatures tested in the temperature challenge experiment, suggesting that this particular H1-delta can be induced by environmental stress.

A more detailed analysis of the sequence of this gene using a maximum likelihood phylogenetic analysis shows the divergence of H1-delta EKC17653 from the rest of H1-delta in *C. gigas*, and the alignment of their protein sequences evidences the aminoacidic differences between them (Figure 7).

Overall, these results suggest that histone H1-delta EKC17653 could be a newly identified member of the histone H1 family with particular functional roles that will require further research to be unraveled. Its apparent response to environmental changes shows great potential for this particular H1-delta as a stress biomarker. Interestingly, other H1 variants had been described as stress-responsive in previous studies, although these studies were so far limited to plants (Kim et al. 2010). This represents the first report of a H1 variant as a potential biomarker for environmental stress in metazoans, which in contrast with the lack of an obvious response by other histone genes like H2A.Z, H2A.X and macroH2A, establishes a distinction between the core set of histones and linker histones in stress response of bivalves.

Conclusions

The comparison between the sets of differentially expressed transcripts across different stressors shows small overlap, revealing only 4 transcripts as deregulated under all the considered experiments, which contrasts with the overall similarities found in their associated functional profiles. This information becomes particularly relevant for the development of biomarkers, where individual genes show little potential given their variability across conditions, time and tissues. Despite this lack of coincidence in specific responsive genes, certain biological processes are consistent with a general response to environmental stress in the line of what has been previously suggested as a consensus model. Such model includes protein degradation, heightened metabolism and transport as well as mechanisms involved in oxidative stress composing a core response to environmental stress in oysters (Anderson et al. 2015). Based on the present results, we

propose complementing this model by including processes related to the maintenance of immunity, which repeatedly appear to be activated regardless of the specific stressor and in the absence of external pathogenic threats (Suarez-Ulloa et al. 2015a). Although a link between stress and immune responses has been described in other organisms (Muralidharan and Mandrekar 2013), recent findings about the tightly linked neuroendocrine-immune regulatory system of bivalves provide a good basis to understand this phenomenon (Liu et al. 2016).

Alternatively, the search of epigenetic biomarkers of stress seems to be hampered by the relatively low level of expression that most histones, histone variants and histone-modifying enzymes display under changing environmental conditions. The role of these chromatin-associated proteins may be supported by alternative mechanisms such as recruitment and relocation in the chromatin without requiring dramatic changes in their expression. However, there seems to be an exception with a particularly divergent variant of the H1 linker family represented by the transcript EKC17653, currently annotated as H1-delta. The sequence analysis of the group of H1-delta proteins predicted by the current genome assembly shows substantial differences that suggest the need of revising such annotation. The H1-delta EKC17653 represents the only histone variant identified in this study as a promising biomarker of stress. The recent discovery of new histone variants in bivalve molluscs opens up an exciting field of research with critical relevance for epigenetics and evolution (Rivera-Casas et al. 2016a, Rivera-Casas et al. 2016b). In order to shed more light into the functional role and evolutionarily origin of this interesting H1-delta, a detailed characterization and functional assays will be required. Subsequently, its real potential as stress biomarker can be assessed.

These results support the value of network analysis of transcriptomic data for a better understanding of the biological processes that are triggered in response to environmental cues. The perspective of a growing pool of available transcriptomic resources paints a promising picture for the understanding of the relationship between traditional non-model organisms such as bivalves and a changing environment, with critical implications for the conservation of marine ecosystems.

References

Alexa, A. and J. Rahnenfuhrer. 2010. topGO: Enrichment analysis for Gene Ontology. Max Plank Institute Informatics.

Anderson, K., D. A. Taylor, E. L. Thompson, A. R. Melwani, S. V. Nair, and D. A. Raftos. 2015. Meta-analysis of studies using suppression subtractive hybridization and microarrays to investigate the effects of environmental stress on gene transcription in oysters. PLoS ONE 10:e0118839.

Bannister, A. J. and T. Kouzarides. 2011. Regulation of chromatin by histone modifications. Cell Res 21:381-395.

Barah, P., P. Winge, A. Kusnierczyk, D. H. Tran, and A. M. Bones. 2013. Molecular signatures in Arabidopsis thaliana in response to insect attack and bacterial infection. PLoS ONE 8:e58987.

Chaney, M. L. and A. Y. Gracey. 2011. Mass mortality in Pacific oysters is associated with a specific gene expression signature. Mol Ecol 20:2942-2954.

Chen, Y., C. F. Chen, D. J. Riley, and P. L. Chen. 2011. Nek1 kinase functions in DNA damage response and checkpoint control through a pathway independent of ATM and ATR. Cell Cycle 10:655-663.

Gotz, S., J. M. Garcia-Gomez, J. Terol, T. D. Williams, S. H. Nagaraj, M. J. Nueda, M. Robles, M. Talon, J. Dopazo, and A. Conesa. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420-3435.

Jing, G., Y. Li, L. Xie, and R. Zhang. 2006. Metal accumulation and enzyme activities in gills and digestive gland of pearl oyster (Pinctada fucata) exposed to copper. Comp Biochem Physiol C Toxicol Pharmacol 144:184-190.

Jung, K. H., H. J. Gho, M. X. Nguyen, S. R. Kim, and G. An. 2013. Genome-wide expression analysis of HSP70 family genes in rice and identification of a cytosolic HSP70 gene highly induced under heat stress. Funct Integr Genomics 13:391-402.

Kanneganti, M., A. Kamba, and E. Mizoguchi. 2012. Role of chitotriosidase (chitinase 1) under normal and disease conditions. J Epithel Biol Pharmacol 5:1-9.

Kim, J. M., T. K. To, T. Nishioka, and M. Seki. 2010. Chromatin regulation functions in plant abiotic stress responses. Plant Cell Environ 33:604-611.

Letwin, K., L. Mizzen, B. Motro, Y. Ben-David, A. Bernstein, and T. Pawson. 1992. A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell cycle regulator and highly expressed in meiotic germ cells. EMBO J 11:3521-3531.

Liu, Z., L. Wang, Z. Zhou, Y. Sun, M. Wang, H. Wang, Z. Hou, D. Gao, Q. Gao, and L. Song. 2016. The simple neuroendocrine-immune regulatory network in oyster Crassostrea gigas mediates complex functions. Sci Rep 6:26396.

Muralidharan, S. and P. Mandrekar. 2013. Cellular stress response and innate immune signaling: integrating pathways in host defense and inflammation. J Leukoc Biol 94:1167-1184.

Nuzzo, R. 2014. Scientific method: statistical errors. Nature 506:150-152.

Priest, H. D., S. E. Fox, E. R. Rowley, J. R. Murray, T. P. Michael, and T. C. Mockler. 2014. Analysis of global gene expression in Brachypodium distachyon reveals extensive network plasticity in response to abiotic stress. PLoS ONE 9:e87499.

Radwan, F. F. and J. S. Ramsdell. 2008. Brevetoxin forms covalent DNA adducts in rat lung following intratracheal exposure. Environ Health Perspect 116:930-936.

Ren, Y., Y. Cui, X. Li, B. Wang, L. Na, J. Shi, L. Wang, L. Qiu, K. Zhang, G. Liu, and Y. Xu. 2015. A co-expression network analysis reveals lncRNA abnormalities in peripheral blood in early-onset schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 63:1-5.

Rivera-Casas, C., R. Gonzalez-Romero, M. S. Cheema, J. Ausio, and J. M. Eirin-Lopez. 2016a. The characterization of macroH2A beyond vertebrates supports an ancestral origin and conserved role for histone variants in chromatin. Epigenetics 11:415-425.

Rivera-Casas, C., R. Gonzalez-Romero, A. Vizoso-Vazquez, M. S. Cheema, M. E. Cerdan, J. Mendez, J. Ausio, and J. M. Eirin-Lopez. 2016b. Characterization of mussel H2A.Z.2: a new H2A.Z variant preferentially expressed in germinal tissues from *Mytilus*. Biochem Cell Biol 94:480-490.

Safari-Alighiarloo, N., M. Rezaei-Tavirani, M. Taghizadeh, S. M. Tabatabaei, and S. Namaki. 2016. Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis. PeerJ 4:e2775.

Schneider, R. F., Y. Li, A. Meyer, and H. M. Gunter. 2014. Regulatory gene networks that shape the development of adaptive phenotypic plasticity in a cichlid fish. Mol Ecol 23:4511-4526.

Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498-2504.

Stohs, S. J. and D. Bagchi. 1995. Oxidative mechanisms in the toxicity of metal ions. Free Radic Biol Med 18:321-336.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, M. V. Prego-Faraldo, F. Florez-Barros, A. Sexto-Iglesias, J. Mendez, and J. M. Eirin-Lopez. 2015a. Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid. PeerJ 3.

Suarez-Ulloa, V., R. Gonzalez-Romero, and J. M. Eirin-Lopez. 2015b. Environmental epigenetics: A promising venue for developing next-generation pollution biomonitoring tools in marine invertebrates. Mar Pollut Bull.

Svensson, S., A. Sarngren, and L. Forlin. 2003. Mussel blood cells, resistant to the cytotoxic effects of okadaic acid, do not express cell membrane p-glycoprotein activity (multixenobiotic resistance). Aquat Toxicol 65:27-37.

T. Cickovski, V. A.-P., W. Huang, S. Mahmoud and G. Narasimhan. 2016. Lightweight Microbiome Analysis Pipelines. Proceedings of 4th International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO16).

Tang, J., F. Wang, Z. Wang, Z. Huang, A. Xiong, and X. Hou. 2013. Characterization and co-expression analysis of WRKY orthologs involved in responses to multiple abiotic stresses in Pak-choi (Brassica campestris ssp. chinensis). BMC Plant Biol 13:188.

Tantai, J. C., X. F. Pan, and H. Zhao. 2015. Network analysis of differentially expressed genes reveals key genes in small cell lung cancer. Eur Rev Med Pharmacol Sci 19:1364-1372.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. Nucl. Acids Res. 22:4673-4680.

Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105-1111.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562-578.

Ucar, D., I. Neuhaus, P. Ross-MacDonald, C. Tilford, S. Parthasarathy, N. Siemers, and R. R. Ji. 2007. Construction of a reference gene association network from multiple profiling data: application to data analysis. Bioinformatics 23:2716-2724.

Vilella, A. J., J. Severin, A. Ureta-Vidal, L. Heng, R. Durbin, and E. Birney. 2009. EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Res 19:327-335.

Zhang, B. and S. Horvath. 2005. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 4:Article17.

Zhang, G., X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que, Y. Xie, P. W. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, J. Meng, L. Yang, J. Liu, B. Wen, N. Zhang, Z. Huang, Q. Zhu, Y. Feng, A. Mount, D. Hedgecock, Z. Xu, Y. Liu, T. Domazet-Loso, Y. Du, X. Sun, S. Zhang, B. Liu, P. Cheng, X. Jiang, J. Li, D. Fan, W. Wang, W. Fu, T. Wang, B. Wang, J. Zhang, Z. Peng, Y. Li, N. Li, M. Chen, Y. He, F. Tan, X. Song, Q. Zheng, R. Huang, H. Yang, X. Du, L. Chen, M. Yang, P. M. Gaffney, S. Wang, L. Luo, Z. She, Y. Ming, W. Huang, B. Huang, Y. Zhang, T. Qu, P. Ni, G. Miao, Q. Wang, C. E. Steinberg, H. Wang, L. Qian, X. Liu, and Y. Yin. 2012a. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49-54.

Zhang, L., R. Hou, H. Su, X. Hu, S. Wang, and Z. Bao. 2012b. Network analysis of oyster transcriptome revealed a cascade of cellular responses during recovery after heat shock. PLoS ONE 7:e35484.

Figure 4: Venn diagram showing the overlap between sets of transcripts differentially expressed under different environmental stressors (i.e., salinity, temperature, pathogenic infection by *Vibrio* spp. and exposure to the metal Zn).

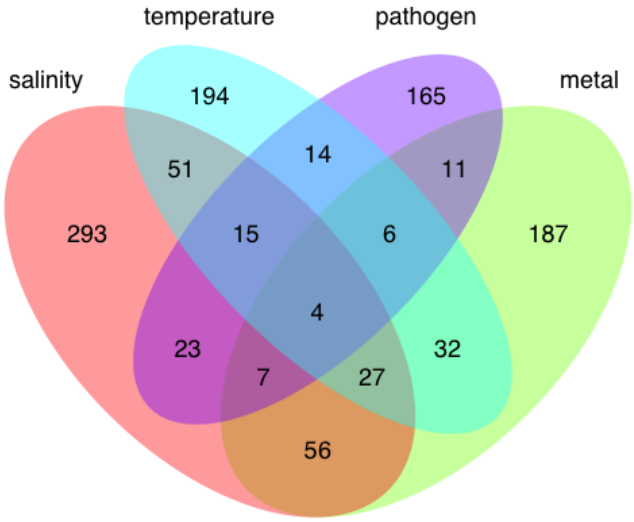


Figure 2: The distribution of general (low-level) GO terms for all three main categories (BP: Biological Process, MF: Molecular Function and CC: Cellular Component) are shown for the four different stressors; A: Salinity, B: Temperature, C: Metal and D: Pathogen. The distributions are calculated based on the number of differentially expressed transcripts annotated with GO terms descendant of the represented low-level categories.

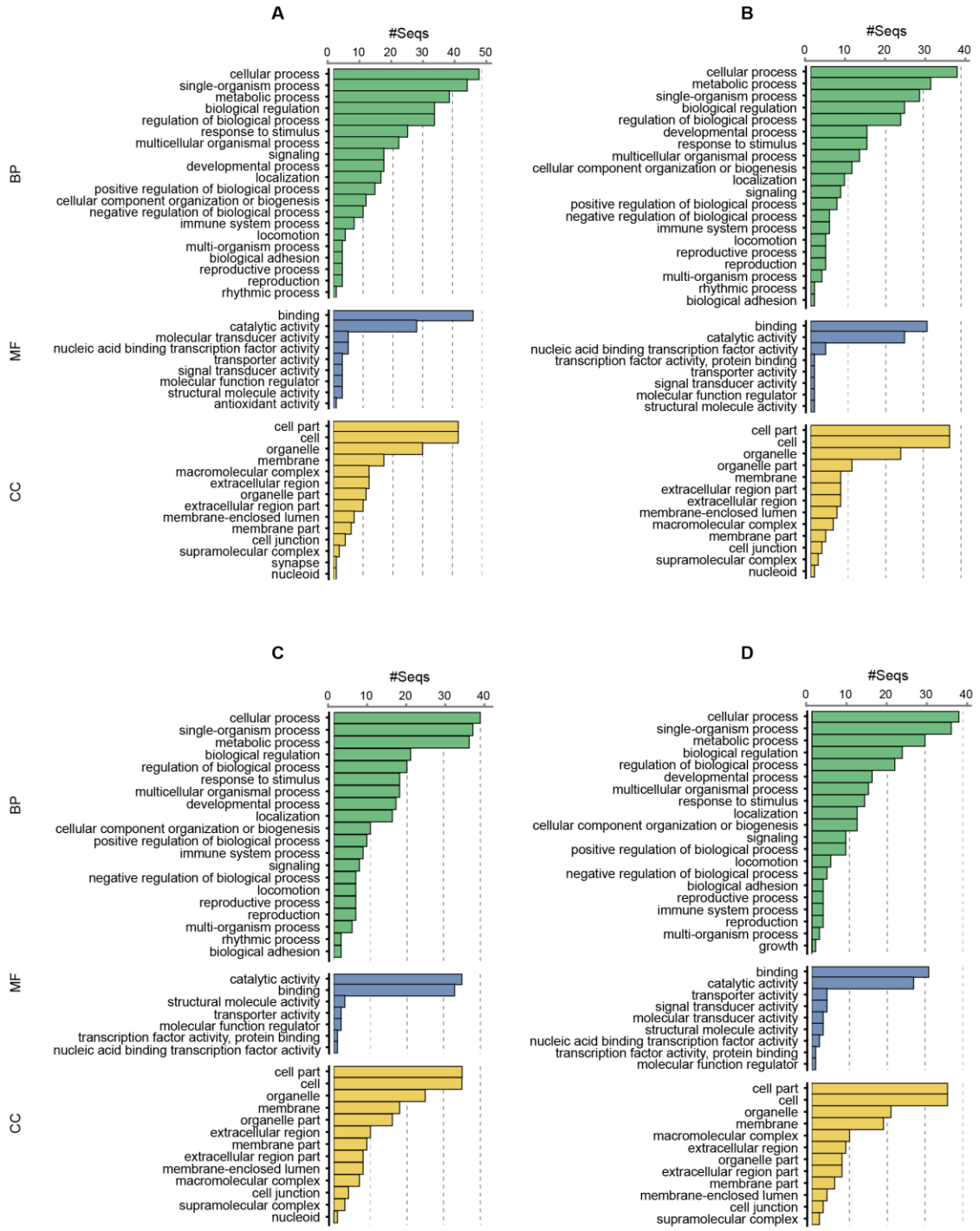


Figure 3: Co-expression gene network of differentially expressed genes under pathogen challenge. Each node represents a transcript and each edge a significant correlation between two nodes. The most interconnected clusters within the network (colored) represent modules with distinct expression profiles subjected to functional analysis. Selected modules show distinct patterns demonstrating the dynamics of the transcriptomic response to pathogen challenge, which allow a chronological organization of the cascade of biological processes that compose the response of oysters to changing environmental conditions. Letters A, B, C, D and E indicate correspondence between clusters and the ordered expression profiles plotted on the right.

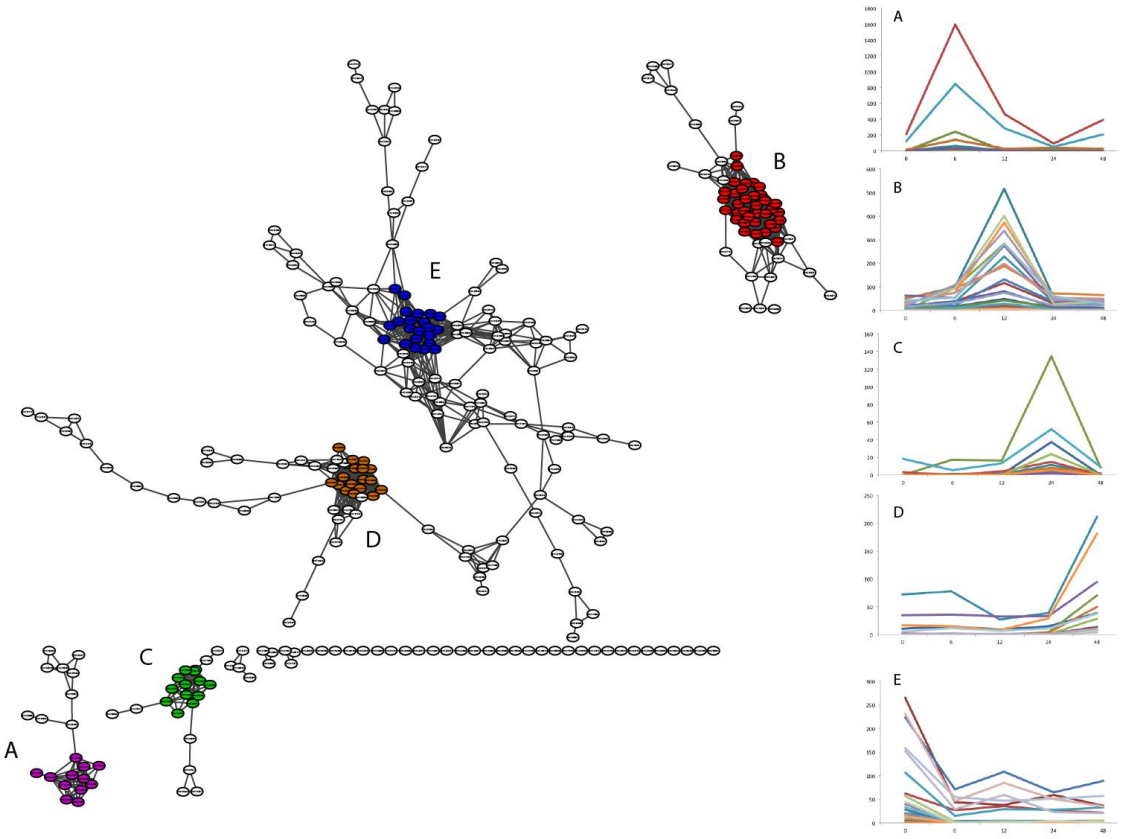


Figure 4: The distribution of general (low-level) GO terms for all three main categories (BP: Biological Process, MF: Molecular Function and CC: Cellular Component) are shown for (A) gills and (B) digestive gland tissues exposed to Zn (1 mg/L). The distributions are calculated based on the number of differentially expressed transcripts annotated with GO terms descendant of the represented low-level categories.

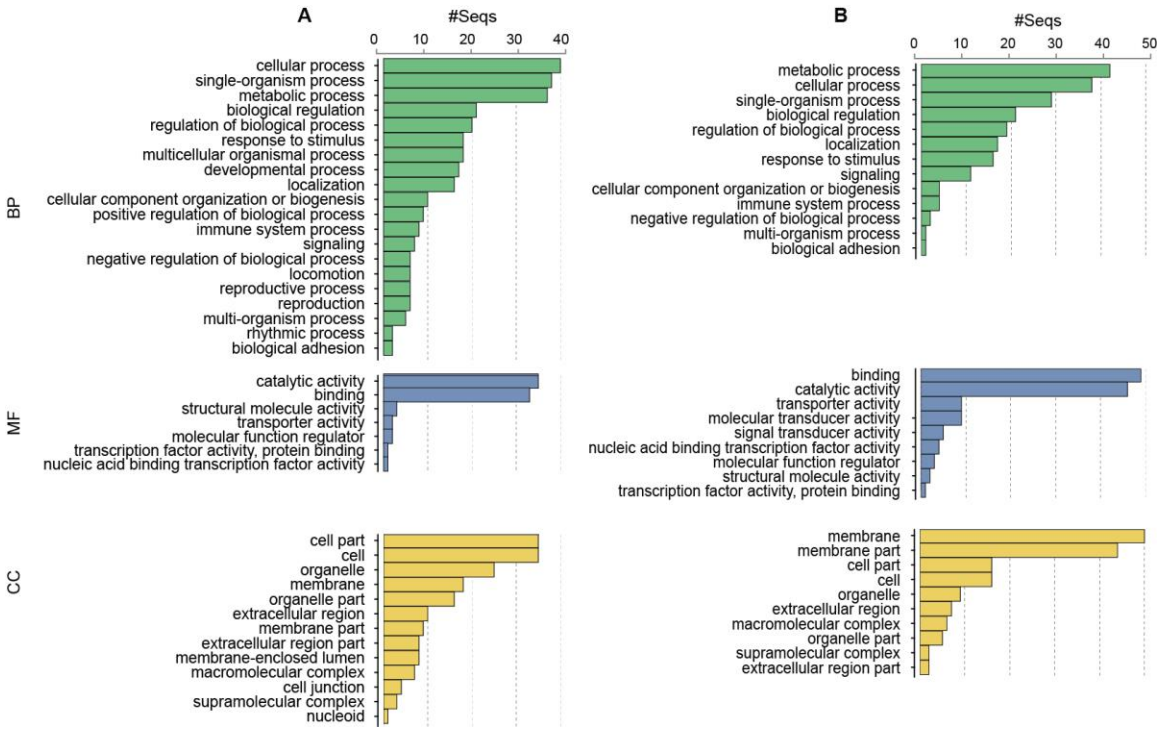


Figure 5: Co-expression network of targeted chromatin-associated genes in *C. gigas* under changing temperature conditions. Nodes represent transcripts with a size proportional to their level of expression, and the edges represent significant correlations in expression between pairs of transcripts, with green indicating a positive correlation and red indicating a negative correlation.

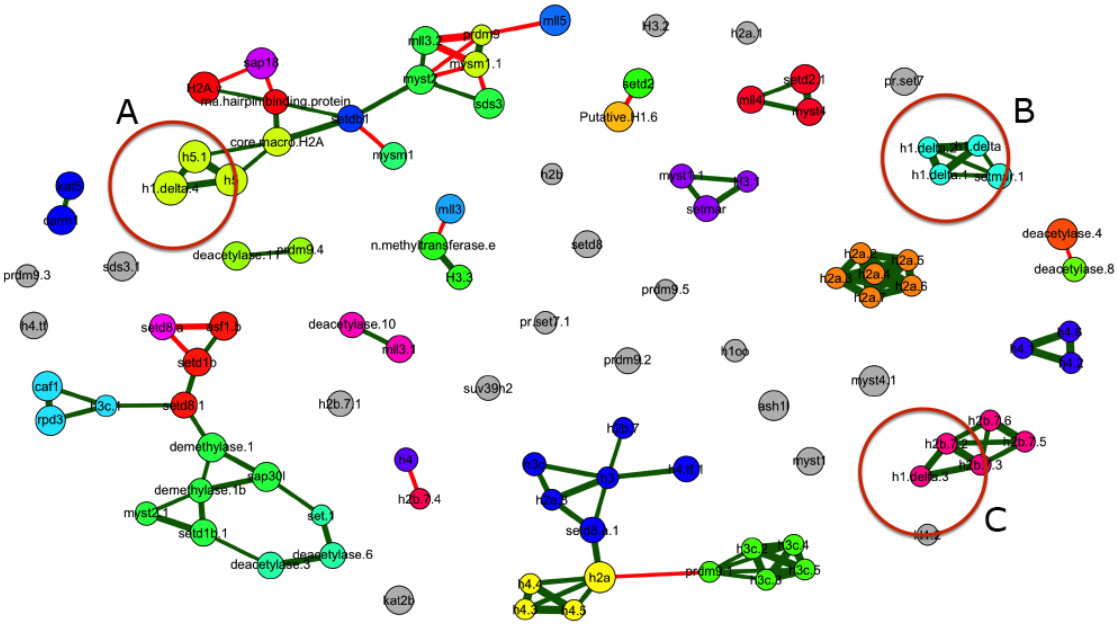


Figure 6: Gene expression profiles of histone, histone variants and histone-modifying enzymes included in network analysis for each one of the different challenge experiments: A) Includes data series with increasing levels of stressor intensities; Temperature challenge and salinity challenge, B) Time series including exposure to pathogen of the *Vibrio* spp. and heavy metal exposure, Zn (1 mg/L), with gill samples and digestive gland samples. Across all these different experiments, the H1-delta EKC17653 shows a higher level of expression than other histones showing substantial variations with changing conditions.

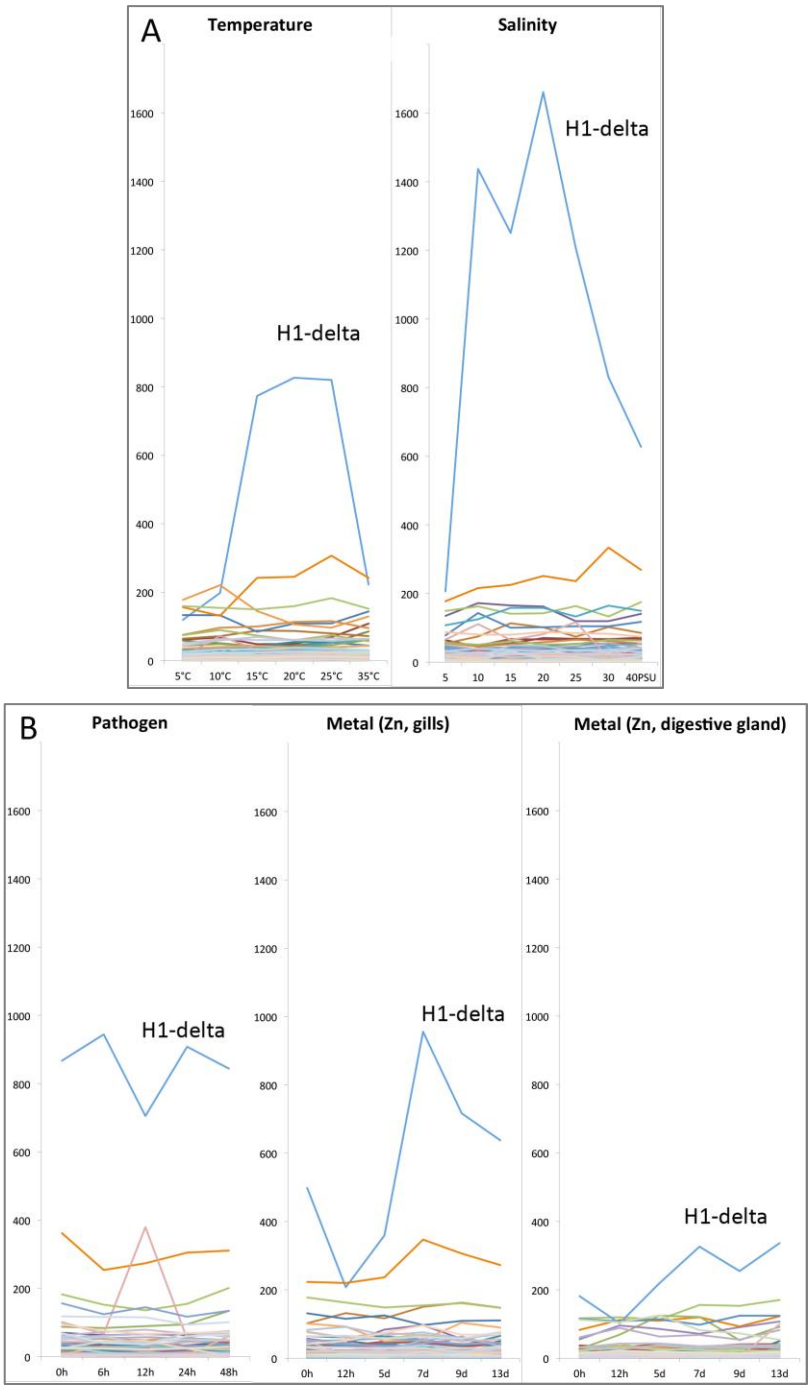
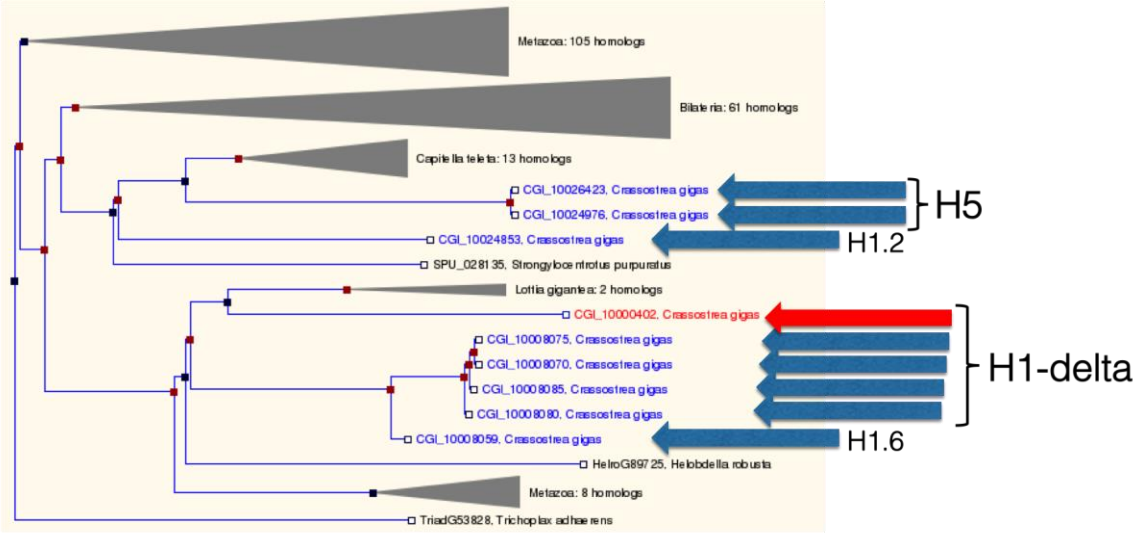


Figure 7: A maximum likelihood tree constructed with TreeBeST on the ENSEMBL web portal, showing homologs for the H1-delta, gene: CGI_10000402 and transcript: EKC17653 (marked in red), which appears divergent from other assumed paralogs H1-delta in the same species *C. gigas*. The protein alignment between these paralogs is also included showing a large number of amino-acid substitutions and gaps.




```
EKC17653 MSDAEVETPVEQPEAAEKDEKATKPAKSPKKKTAKAKNKTPLNHPHYRDMIKDALTNLKERGGSSRQAILKY
EKC28024.A.TATT...KK--------.V...-------------.V.AA..K.V...RA..ES..............
EKC28029.A.TATT...KK--------.V...-------------.V.AA..K.V...RA..ES..............
EKC28034.A.TATT...KK--------.V...-------------.V.AA..K.V...RA..ES..............
EKC28039.A.TATT...KK--------.V...-------------.V.AA..K.V...RA..ES..............

EKC17653 IVKNFKVE-DEKSANNHLKMALRAGVKNGSLKQSKGTGASGSFRLGE-SKPPSKPKAK---KEKAAKPK---
EKC28024 MA....GN.VNPI.V......KN...K.A...A.....T...K..DKP.TEK.....KVT.P.......KAT
EKC28029 MA....GN.VNPI.V......KN...K.A...A.....T...K..DKP.TEK.....KVT.P.......KAT
EKC28034 MA....GN.VNPI.V......KN...K.A...A.....T...K..DKP.TEK.....KVT.P.......KAT
EKC28039 MA....GN.VNPI.V......KN...K.A...A.....T...K..DKP.TEK.....KVT.P.......KAT

EKC17653 ---------AEKKAKKKPKVKAKAEKKPKEKAEKK---AKSPKKAAAK--KASKPKK----------KSPKKAAAKKX
EKC28024 AVKPRKAAGEK.T.E..K.SPK..AGPK.V.TP..KTA..........PK..KT...-AAAAKPKKA.T...T.....
EKC28029 AVKPRKAAGEK.T.E..K.SPK..AGPK.V.TP..KTA..........PK..KT...-AAAAKPKKA.T...T.....
EKC28034 AVKPRKAAGEK.T.E..K.SPK..AGPK.V.TP..KTA..........PK..KT...-AAAAKPKKA.T...T.....
EKC28039 AVKPRKAAGEK.T.E..K.SPK..AGPK.V.TP..KTA..........PK..KT...AAAAAKPKKA.T...T.....
```

Table 2: The datasets included in this meta-analysis are listed including their accession number and the description of the experimental sample or treatment they represent. Those samples marked with an asterisk correspond with the control samples for each series.

Salinity		**Temperature**		**Metal (gills)**		**Metal (digestive gland)**		**Pathogen**	
SRA accession	Sample	SRA accession	Sample	SRA accession	Sample	SRA accession	Sample	SRA accession	Sample
SRR334269	5 PSU	SRR334262	5 °C	SRR334312	0 h*	SRR334313	0 h*	SRR796582	0 h*
SRR334270	10 PSU	SRR334263	10 °C	SRR334291	12 h	SRR334324	12 h	SRR796583	6 h
SRR334271	15 PSU	SRR334264	15 °C	SRR334319	5 d	SRR334315	5 d	SRR796584	12 h
SRR334272	20 PSU	SRR334265	20 °C*	SRR334320	7 d	SRR334316	7 d	SRR796585	24 h
SRR334273	25 PSU	SRR334266	25 °C	SRR334321	9 d	SRR334317	9 d	SRR796586	48 h
SRR334274	30 PSU*	SRR334268	35 °C	SRR334322	13 d	SRR334318	13 d		
SRR334275	40 PSU								

Table 3: The 20 most significant GO terms for each set of differentially expressed transcripts show similarities across stressors, including references to mechanisms of metabolic processes, oxidative stress, protein folding and degradation as well as immune responses.

Salinity	Temperature	Metal	Pathogen
Superoxide metabolic process	Alpha-amino acid biosynthetic process	Oxidation-reduction process	Anion transport
Alpha-amino acid biosynthetic process	Cellular modified amino acid metabolic process	Protein ubiquitination	Cell-matrix adhesion
Protein ubiquitination	Sulfur amino acid metabolic process	Dicarboxylic acid transport	Oxidation-reduction process
Peptide biosynthetic process	Aspartate family amino acid metabolic process	Inorganic anion transmembrane transport	Lipoprotein metabolic process
Oxidation-reduction process	Immune response	Proteolysis	Immune response
Dicarboxylic acid transport	Transition metal ion transport	Transition metal ion transport	Lipid transport
DNA integration	Tryptophan catabolic process	Peptide biosynthetic process	Dicarboxylic acid transport
Cell-matrix adhesion	Response to ATP	Metal ion homeostasis	Cell development
Tryptophan catabolic process	Purinergic nucleotide receptor signaling pathway	Defense response to virus	Peptidyl-tyrosine phosphorylation
Benzene-containing compound metabolic process	Benzene-containing compound metabolic process	Pyrimidine nucleotide metabolic process	Reproduction
DNA recombination	Metal ion homeostasis	Response to oxidative stress	Aspartate family amino acid biosynthetic...
Positive regulation of macromolecule biosynthetic process	Iron-sulfur cluster assembly	Peptidyl-tyrosine phosphorylation	Glutathione metabolic process
Positive regulation of cellular biosynthetic process	Peptidyl-tyrosine phosphorylation	Cellular chemical homeostasis	Amide biosynthetic process

Regulation of cell proliferation	Cellular chemical homeostasis	Protein folding	Phosphatidylinositol phosphorylation
Positive regulation of nucleobase-containing compound metabolic process	Cellular ketone metabolic process	Chitin catabolic process	Superoxide metabolic process
Defense response to virus	Amide biosynthetic process	Receptor-mediated endocytosis	Glucose metabolic process
Protein phosphorylation	Oxidation-reduction process	Negative regulation of endopeptidase activity	Carbohydrate biosynthetic process
Pyrimidine nucleotide metabolic process	Superoxide metabolic process	Carbohydrate metabolic process	Methylation
Cellular ketone metabolic process	Nucleosome assembly	Protein peptidyl-prolyl isomerization	Carbohydrate metabolic process
Regulation of developmental process	Carbohydrate transport	Response to biotic stimulus	Wnt signaling pathway

Table 4: The number of up-regulated and down-regulated transcripts considering any time point or stressor level is added and the totals are compared. There is no clear trend towards up- or down-regulation across stressors.

	Salinity	**Temperature**	**Metal (gills)**	**Metal (d. g.)**	**Pathogen**
Up-regulated	427	208	220	219	216
Down-regulated	350	237	271	230	203
Difference	**77**	**-29**	**-51**	**-11**	**13**

Table 5: Gene Ontology terms in the Biological Process category enriched in the ordered modules of differentially expressed transcripts under stress by metal exposure (Zn, 1mg/L).

12 hours after exposure to metal	**5 days after exposure to metal**
Peptide biosynthetic process Protein peptidyl-prolyl isomerization Alpha-amino acid biosynthetic process Protein folding Negative regulation of endopeptidase activity	Negative regulation of endopeptidase activity Regulation of transcription, DNA-templated Nucleobase-containing compound metabolic process Transcription, DNA-templated Proteolysis
7 days after exposure	**13 days after exposure**
Carbohydrate metabolic process Dicarboxylic acid transport Inorganic anion transmembrane transport Response to biotic stimulus Regulation of transcription, DNA-templated	Chitin metabolic process Inorganic anion transmembrane transport Oxidation-reduction process Carbohydrate metabolic process Aminoglycan metabolic process

Table 6: Gene Ontology terms in the Biological Process category enriched in the ordered modules of differentially expressed transcripts under stress by pathogen exposure by *Vibrio* spp.

6 hours after exposure to pathogen	**12 hours after exposure to pathogen**
Proteolysis	Inorganic anion transport
Biological process	Anion transmembrane transport
Metabolic process	Carbohydrate metabolic process
Protein metabolic process	Immune response
Macromolecule metabolic process	Response to biotic stimulus
24 hours after exposure to pathogen	**48 hours after exposure to pathogen**
Protein phosphorylation	Proteolysis
Translational initiation	Biological_process
Carbohydrate metabolic process	Metabolic process
Translation	Protein metabolic process
Cellular protein modification process	Macromolecule metabolic process

CHAPTER VI

CONCLUSIONS

Global change poses unprecedented challenges for the oceans, which are directly impacted by the effects of growing human populations, increasing temperatures and pollution. Consequently, such changes in the environment will elicit molecular responses in marine organisms in order to maintain homeostasis and survival, and these responses may set the basis for adaptive mechanisms that may preserve the future of the species. By improving our knowledge about these molecular responses and mechanisms in marine organisms, we will gain a better position to control and possibly mitigate the consequences of environmental stress.

Despite of the critical ecological and economical value of bivalve molluscs and their extensive exploit as bioindicators, further efforts are required to understand the molecular strategies they can display to cope with environmental changes and stress. Generally speaking, this taxonomical group has been traditionally underrepresented in molecular biology research, illustrated with the first complete genome of a bivalve organism released only as recently as 2012 (Zhang et al. 2012). In this sense, oysters and mussels represent the vanguard of recent efforts to fill this gap in knowledge, constituting the target of the majority of existing genomic and epigenetic studies (Suarez-Ulloa et al. 2013b). The present dissertation makes a relevant contribution to this goal, pioneering the transcriptomic and epigenetic characterization of the response to Harmful Algal Blooms (HABs) in oysters and mussels.

HABs represent an environmental issue of growing concern given their increasing trends in frequency and severity and their effects on marine bivalve populations can be critical for the strongly dependent aquaculture industry and the health of human consumers (Anderson 2009). Given their extraordinary ability to accumulate biotoxins in

their tissues, bivalves have been commonly used as sentinel organisms for HAB monitoring purposes, traditionally using physiological or chemical techniques. However, innovative molecular tools based on transcriptomic or epigenetic analyses have the potential to improve some of the of shortcomings of these traditional methods, while providing relevant information about how these organisms respond to the environmental stress associated with HABs (Prego-Faraldo et al. 2013, Suarez-Ulloa et al. 2013b).

With this goal in mind, Chapter II and Chapter III have demonstrated that there are sharp changes in gene expression even with low concentration of dinoflagellates in the water, corresponding to an early stage of a HAB, and even with amounts of accumulated toxin below legal limits for consumption (Suarez-Ulloa et al. 2013a, Suarez-Ulloa et al. 2015).

Interestingly, many of the genes showing a significant change in expression were related with biological functions that are consistent with a previously reported consensus model for stress responses in oysters (Anderson et al. 2015). Other authors have also identified such concept as a universal core response to stress, coining the term of 'minimal cellular stress response' (van Oppen et al. 2017), showing that many of the molecular processes triggered in many organisms as part of the response to environmental stress are actually unspecific to the stressor. Accordingly, our results showed little specificity in the functional profile of the response of mussels to HABs, while supporting a relevant role of the innate immune system as part of the response of bivalves to stress. Immune-related genes had been previously reported as responsive to abiotic stress in other marine invertebrates like corals (Barshis et al. 2013, Pinzon et al. 2015), as discussed by Van Oppen et al. (van Oppen et al. 2017). Naturally, this leads to

further questions on the potential of genomic tools in monitoring efforts and the particularities of the minimal cellular stress response in bivalves.

In order to better understand what is common and what is specific in the transcriptomic response of bivalves when dealing with different environmental stressors, publicly available RNA-Seq data series from the oyster *Crassostrea gigas* were analyzed in Chapter V. Results showed that gene expression is highly dynamic, displaying a great variability with time, the intensity level of the stressor and the type of tissue. Specific sets of differentially expressed genes showed a relatively small overlap when datasets from experiments with different stressors were compared. However, those sets of genes yield a similar functional profile independent of the stressor. This suggests that although different stressors can specifically regulate different genes, those sets of genes fulfill similar functions and biological processes consistent with the previous consensus model for a core response to stress (Anderson et al. 2015).

These results taken together further support the concept of a minimal cellular stress response unspecific of the environmental conditions, including stress associated with HABs toxicity. The search for sensitive and specific biomarkers based on gene expression patterns must have these observations into account, ruling out those molecular pathways that are less informative and focusing in those that are specific and selective. Similarly, the observed variability with time and levels of intensity of stress suggests that these are relevant factors to consider. Based on these results, time-resolved multimarker approaches seem a promising avenue for proper biomarker development. The potential of such methods and their predictive power will improve as the amount of available transcriptomic data increases in the upcoming years.

Although for bivalves, transcriptomic studies are massively abundant when compared to epigenetic and epigenomic studies, the latter have the potential of providing valuable insights into the heritability of acquired environmental information. Indeed, Environmental Epigenetics represents an emergent field with great potential to better understand not only how organisms respond to changing environmental conditions but also how those events may have consequences for later stages in life or even further generations.

In the present dissertation, Chapter II characterized the set of chromatin-associated genes in the mussel *Mytilus* together with their corresponding expression levels under exposure to HABs. This information was organized and catalogued in a publicly available database designed as a specialized resource of functional genomic data. The use of this database has led to the discovery of specific histone variants identified and characterized for the first time in marine invertebrates (Rivera-Casas et al. 2016a, Rivera-Casas et al. 2016b).

Based on this work, which allowed the development of specific primers and antibodies, Chapter IV was able to dig deeper into the role of histone variants and other epigenetic mechanisms such as DNA methylation in the response of bivalves to HABs (i.e., Florida Red Tides). In particular, this work showed that the phosphorylation of H2A.X is triggered by exposure to increasing concentrations of toxic Karenia brevis, supporting the genotoxic potential of brevetoxins and the role of γH2A.X in mechanisms of DNA repair in a mollusc. Additionally, DNA methylation patterns showed significant changes within hours of exposure, potentially participating in the epigenetic memory of bivalves (Gonzalez-Romero et al. 2017).

On the other hand, histone variants H2A.Z, H2A.X and macroH2A failed to show significant changes in expression during exposure. Consistently, the work carried out in Chapter V confirmed that for most histone variants and other chromatin-associated proteins, the changes in expression are generally negligible under various sources of environmental stress. Interestingly, one specific histone variant defined by the transcript EKC17653 and annotated as H1-delta shows potential as biomarker of environmental stress. Further research will be required to clarify the nature and function of this gene in the environmental responses of oysters.

Overall, the present dissertation has broken new ground in our knowledge of epigenetic mechanisms in bivalves, leading to a better characterization of chromatin-associated genes in these organisms as well as the response of different epigenetic marks to environmental stress. Together with the transcriptomic characterization of environmental responses in bivalves, this work represents a significant contribution towards our understanding of molecular acclimatory processes and adaptation. Furthermore, this knowledge provides a springboard for the development of innovative tools for oceans biomonitoring.

Further research will be required to clarify the interaction between epigenetic marks and expression changes in these organisms. The generation of different types of data from the same experiment and subsequent analysis using integrative approaches is a critical requirement in order to better understand the fundamental regulatory role of epigenetics in bivalve molluscs. Since our current knowledge of epigenetics is primarily drawn from the study of vertebrate species, such integrative studies using alternative taxa have the exciting potential of challenging established paradigms in this field.

References

Anderson, D. M. 2009. Approaches to monitoring, control and management of harmful algal blooms (HABs). Ocean Coast Manag 52:342-347.

Anderson, K., D. A. Taylor, E. L. Thompson, A. R. Melwani, S. V. Nair, and D. A. Raftos. 2015. Meta-analysis of studies using suppression subtractive hybridization and microarrays to investigate the effects of environmental stress on gene transcription in oysters. PLoS ONE 10:e0118839.

Barshis, D. J., J. T. Ladner, T. A. Oliver, F. O. Seneca, N. Traylor-Knowles, and S. R. Palumbi. 2013. Genomic basis for coral resilience to climate change. Proc Natl Acad Sci U S A 110:1387-1392.

Gonzalez-Romero, R., V. Suarez-Ulloa, J. Rodriguez-Casariego, D. Garcia-Souto, G. Diaz, A. Smith, J. J. Pasantes, G. Rand, and J. M. Eirin-Lopez. 2017. Effects of Florida Red Tides on histone variant expression and DNA methylation in the Eastern oyster Crassostrea virginica. Aquat Toxicol 186:196-204.

Pinzon, J. H., B. Kamel, C. A. Burge, C. D. Harvell, M. Medina, E. Weil, and L. D. Mydlarz. 2015. Whole transcriptome analysis reveals changes in expression of immune-related genes during and after bleaching in a reef-building coral. R Soc Open Sci 2:140214.

Prego-Faraldo, M. V., V. Valdiglesias, J. Mendez, and J. M. Eirin-Lopez. 2013. Okadaic Acid meet and greet: an insight into detection methods, response strategies and genotoxic effects in marine invertebrates. Mar Drugs 11:2829-2845.

Rivera-Casas, C., R. Gonzalez-Romero, M. S. Cheema, J. Ausio, and J. M. Eirin-Lopez. 2016a. The characterization of macroH2A beyond vertebrates supports an ancestral origin and conserved role for histone variants in chromatin. Epigenetics 11:415-425.

Rivera-Casas, C., R. Gonzalez-Romero, A. Vizoso-Vazquez, M. S. Cheema, M. E. Cerdan, J. Mendez, J. Ausio, and J. M. Eirin-Lopez. 2016b. Characterization of mussel H2A.Z.2: a new H2A.Z variant preferentially expressed in germinal tissues from *Mytilus*. Biochem Cell Biol 94:480-490.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, M. V. Prego-Faraldo, F. Florez-Barros, A. Sexto-Iglesias, J. Mendez, and J. M. Eirin-Lopez. 2015. Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid. PeerJ 3.

Suarez-Ulloa, V., J. Fernandez-Tajes, V. Aguiar-Pulido, C. Rivera-Casas, R. Gonzalez-Romero, J. Ausio, J. Mendez, J. Dorado, and J. M. Eirin-Lopez. 2013a. The CHROMEVALOA database: a resource for the evaluation of okadaic acid contamination in the marine environment based on the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis*. Mar Drugs 11:830-841.

Suarez-Ulloa, V., J. Fernandez-Tajes, C. Manfrin, M. Gerdol, P. Venier, and J. M. Eirin-Lopez. 2013b. Bivalve omics: state of the art and potential applications for the biomonitoring of harmful marine compounds. Mar Drugs 11:4370-4389.

van Oppen, M. J., R. D. Gates, L. L. Blackall, N. Cantin, L. J. Chakravarti, W. Y. Chan, C. Cormick, A. Crean, K. Damjanovic, H. Epstein, P. L. Harrison, T. A. Jones, M. Miller, R. J. Pears, L. M. Peplow, D. A. Raftos, B. Schaffelke, K. Stewart, G. Torda, D. Wachenfeld, A. R. Weeks, and H. M. Putnam. 2017. Shifting paradigms in restoration of the world's coral reefs. Glob Chang Biol.

Zhang, G., X. Fang, X. Guo, L. Li, R. Luo, F. Xu, P. Yang, L. Zhang, X. Wang, H. Qi, Z. Xiong, H. Que, Y. Xie, P. W. Holland, J. Paps, Y. Zhu, F. Wu, Y. Chen, J. Wang, C. Peng, J. Meng, L. Yang, J. Liu, B. Wen, N. Zhang, Z. Huang, Q. Zhu, Y. Feng, A. Mount, D. Hedgecock, Z. Xu, Y. Liu, T. Domazet-Loso, Y. Du, X. Sun, S. Zhang, B. Liu, P. Cheng, X. Jiang, J. Li, D. Fan, W. Wang, W. Fu, T. Wang, B. Wang, J. Zhang, Z. Peng, Y. Li, N. Li, M. Chen, Y. He, F. Tan, X. Song, Q. Zheng, R. Huang, H. Yang, X. Du, L. Chen, M. Yang, P. M. Gaffney, S. Wang, L. Luo, Z. She, Y. Ming, W. Huang, B. Huang, Y. Zhang, T. Qu, P. Ni, G. Miao, Q. Wang, C. E. Steinberg, H. Wang, L. Qian, X. Liu, and Y. Yin. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49-54.

Appendix A: Supplementary materials for Chapter II.

Supplementary Material S1. List of keywords used to identify chromatin-associated transcripts in the assembled OA-specific transcriptome from *M. galloprovincialis*.

Acf
Centromere
Chd
Chrac
Chromatin
Dna binding
Dna damage
Dna repair
Double strand break
Epigenetic
High mobility group
Histone
Hmg-a,b,n
Ino80
Mi-2
Nucleosome
Nurf
Post translational modification
Protamine
Rsc
Rsf
Saga
Swi/snf

Supplementary Material S2. Script used to implement a keyword-based routine for the identification of chromatin-associated transcripts among sequence descriptions and related ontology terms.

```
#! /usr/bin/perl -w

#Victoria Suárez-Ulloa
#Selection of annotated sequences by keyword match

use strict;

print "Insert input file name:\n";
my $inputfile = <STDIN>;
chomp $inputfile;

print "Insert name for the output file:\n";
my $outputfile = <STDIN>;
chomp $outputfile;

open(DATA, $inputfile) || die "ERROR opening input file\n";
my @todo = <DATA>;
close(DATA);

my @selected;
my $i;

foreach my $k (@todo){
	if ($k=~ m/protamine|histone|chromatin|swi\Wsnf|ino80|chd\w|mi-
	2|centromere|dna\Wdamage|dna\Wrepair|dna\Wbinding|double\Wstrand\Wbreak|
	high\Wmobility\Wgroup|hmg[a,b,n]|nucleosom|epigenetic|post\Wtranslational\W
	modification|saga\W|rsc\W|nurf\W|chrac\W|acf\W|rsf\W/i)  {
		push (@selected,$k);
	}
}

open(DATAOUT, ">$outputfile") || die "ERROR generating outfile\n";
print DATAOUT @selected;
close DATAOUT;


exit;
```

Supplementary Material S3. Differential expression analysis results displaying unigenes with False Discovery Rate (FDR) < 0.1.

unigenes	logFC	logCPM	LR	P-Value	FDR
XP_002587564	-10.25158	11.62690	19.27328	0.00001	0.03494
CAD79340	-10.10697	11.48231	18.91439	0.00001	0.03494
XP_786549	-9.97335	11.34872	18.58287	0.00002	0.03494
CBN81883	-9.92006	11.29544	18.45067	0.00002	0.03494
XP_615721	-9.91270	11.28809	18.43243	0.00002	0.03494
XP_002738701	-9.89490	11.27028	18.38826	0.00002	0.03494
XP_002115324	-9.52408	10.89956	17.46895	0.00003	0.03649
XP_003460288	-9.47698	10.85247	17.35226	0.00003	0.03649
XP_002737341	-9.43652	10.81202	17.25203	0.00003	0.03649
XP_003417531	-9.31457	10.69011	16.95000	0.00004	0.03649
XP_003373959	-9.28753	10.66308	16.88304	0.00004	0.03649
EFN87950	-9.24599	10.62155	16.78021	0.00004	0.03649
XP_001973709	-9.21523	10.59080	16.70407	0.00004	0.03649
XP_002601470	-9.19839	10.57397	16.66239	0.00004	0.03649
XP_002732922	9.22415	10.86852	16.56410	0.00005	0.03649
XP_002605693	-8.97367	10.34933	16.10643	0.00006	0.03807
XP_002741302	-8.95660	10.33227	16.06424	0.00006	0.03807
XP_002434442	-8.95374	10.32941	16.05715	0.00006	0.03807
XP_002609450	-8.94511	10.32078	16.03583	0.00006	0.03807
XP_002599067	-8.83425	10.20997	15.76179	0.00007	0.04131
XP_002585781	-8.80594	10.18167	15.69185	0.00007	0.04131
XP_003742476	-8.72762	10.10339	15.49839	0.00008	0.04368
XP_785416	8.74077	10.38605	15.37011	0.00009	0.04471
XP_782273	-8.55317	9.92904	15.06775	0.00010	0.04538
YP_001474807	-8.51870	9.89459	14.98272	0.00011	0.04538
AAI61151	-8.47543	9.85133	14.87597	0.00011	0.04538
XP_003727072	-8.45936	9.83528	14.83636	0.00012	0.04538
XP_003447001	8.51848	10.16428	14.82215	0.00012	0.04538
BAC39029	-8.43081	9.80674	14.76596	0.00012	0.04538
AES07023	-8.42255	9.79849	14.74559	0.00012	0.04538
XP_003208890	-8.40588	9.78183	14.70450	0.00013	0.04538
XP_002603759	-8.38051	9.75648	14.64197	0.00013	0.04538
XP_782949	-8.31952	9.69553	14.49170	0.00014	0.04538
ABO61331	-8.29257	9.66859	14.42531	0.00015	0.04538

XP_002613292	8.35630	10.00253	14.42289	0.00015	0.04538
EFW45850	8.31568	9.96203	14.32298	0.00015	0.04538
XP_001631596	-7.91962	11.04169	14.30354	0.00016	0.04538
H2A0L8	-8.23240	9.60846	14.27713	0.00016	0.04538
NP_001106507	8.29194	9.93836	14.26461	0.00016	0.04538
XP_002597426	8.29194	9.93836	14.26461	0.00016	0.04538
XP_003126347	-8.20857	9.58465	14.21847	0.00016	0.04538
XP_426008	-8.20376	9.57984	14.20662	0.00016	0.04538
XP_001748929	-8.14972	9.52584	14.07364	0.00018	0.04548
XP_002600133	8.20565	9.85234	14.05251	0.00018	0.04548
XP_003222256	-8.11425	9.49040	13.98638	0.00018	0.04548
YP_721699	-8.09878	9.47494	13.94833	0.00019	0.04548
XP_003728402	8.13737	9.78428	13.88479	0.00019	0.04548
XP_002739711	-8.06201	9.43820	13.85793	0.00020	0.04548
BAL27710	-8.05134	9.42753	13.83168	0.00020	0.04548
EGI61398	-8.04597	9.42217	13.81848	0.00020	0.04548
XP_003741028	-8.04597	9.42217	13.81848	0.00020	0.04548
XP_001625301	-8.02974	9.40595	13.77859	0.00021	0.04548
XP_002606494	-8.02429	9.40051	13.76520	0.00021	0.04548
XP_003701050	-7.99673	9.37297	13.69746	0.00021	0.04558
XP_003690478	8.05868	9.70585	13.69163	0.00022	0.04558
XP_002741848	8.04455	9.69177	13.65695	0.00022	0.04559
XP_002736149	8.00495	9.65231	13.55982	0.00023	0.04559
BAB23185	-7.93417	9.31046	13.54379	0.00023	0.04559
NP_502235	7.99763	9.64502	13.54187	0.00023	0.04559
EFX84936	-7.91663	9.29294	13.50071	0.00024	0.04559
XP_002740733	7.97917	9.62662	13.49661	0.00024	0.04559
XP_003727690	7.91458	9.56228	13.33831	0.00026	0.04694
XP_002130810	-7.85042	9.22678	13.33818	0.00026	0.04694
XP_002602259	-7.85042	9.22678	13.33818	0.00026	0.04694
XP_002427821	-7.84424	9.22062	13.32303	0.00026	0.04694
GAA33255	7.89107	9.53886	13.28073	0.00027	0.04728
XP_002613584	7.85913	9.50704	13.20249	0.00028	0.04792
ABO33165	-7.79389	9.17031	13.19950	0.00028	0.04792
CAJ82869	-7.76803	9.14448	13.13611	0.00029	0.04885
ZP_01262383	-7.74834	9.12480	13.08782	0.00030	0.04906
XP_002599303	7.79303	9.44121	13.04070	0.00030	0.04906
XP_002590998	-7.72836	9.10485	13.03886	0.00031	0.04906
AFI39909	-7.72165	9.09814	13.02240	0.00031	0.04906

XP_002608213	7.75447	9.40282	12.94639	0.00032	0.04993
XP_003223296	-7.68757	9.06410	12.93892	0.00032	0.04993
XP_002613213	7.73700	9.38542	12.90366	0.00033	0.05021
AEV89778	-7.63848	9.01506	12.81868	0.00034	0.05155
XP_003696069	-7.62414	9.00073	12.78357	0.00035	0.05155
XP_002604184	-7.61692	8.99352	12.76589	0.00035	0.05155
GAA31306	7.66952	9.31824	12.73876	0.00036	0.05155
XP_313650	7.66491	9.31365	12.72749	0.00036	0.05155
XP_002613081	-7.59503	8.97165	12.71231	0.00036	0.05155
XP_001843141	-7.58024	8.95689	12.67615	0.00037	0.05178
XP_421979	7.63692	9.28579	12.65913	0.00037	0.05178
EFN85665	-7.53497	8.91167	12.56541	0.00039	0.05356
EFA84418	-7.51956	8.89627	12.52773	0.00040	0.05356
XP_002603588	-7.51179	8.88852	12.50873	0.00041	0.05356
XP_003728822	-7.51179	8.88852	12.50873	0.00041	0.05356
EFA11608	-7.49613	8.87287	12.47045	0.00041	0.05398
XP_002732097	-7.48824	8.86499	12.45115	0.00042	0.05398
XP_786504	-7.48030	8.85706	12.43175	0.00042	0.05398
XP_002731879	-7.43993	8.81674	12.33313	0.00044	0.05566
XP_001631483	-7.42346	8.80029	12.29292	0.00045	0.05566
XP_002731704	-7.42346	8.80029	12.29292	0.00045	0.05566
XP_002731924	7.47245	9.12213	12.25795	0.00046	0.05566
XP_002596309	-7.40681	8.78366	12.25224	0.00046	0.05566
XP_001362648	7.46185	9.11158	12.23213	0.00047	0.05566
XP_001663053	-7.39841	8.77527	12.23173	0.00047	0.05566
XP_003751729	7.44581	9.09563	12.19305	0.00048	0.05566
GAA29877	-7.00087	10.68109	12.18212	0.00048	0.05566
Q05025	7.44042	9.09027	12.17993	0.00048	0.05566
AAF09840	-7.36430	8.74121	12.14850	0.00049	0.05583
XP_781514	-7.35565	8.73257	12.12739	0.00050	0.05583
XP_002600118	7.40213	9.05218	12.08672	0.00051	0.05583
XP_002742050	-7.32937	8.70633	12.06329	0.00051	0.05583
EFX76362	7.39101	9.04112	12.05964	0.00052	0.05583
XP_003480922	7.39101	9.04112	12.05964	0.00052	0.05583
XP_002414307	-7.32051	8.69747	12.04167	0.00052	0.05583
YP_131523	-7.31159	8.68856	12.01992	0.00053	0.05583
XP_001023282	-7.30261	8.67960	11.99803	0.00053	0.05583
XP_003401618	-7.30261	8.67960	11.99803	0.00053	0.05583
XP_002609208	7.34561	8.99598	11.94922	0.00055	0.05630

XP_003646937	7.34561	8.99598	11.94922	0.00055	0.05630
NP_001232024	-7.26613	8.64318	11.90914	0.00056	0.05702
XP_001635351	-7.22871	8.60581	11.81798	0.00059	0.05936
XP_002601497	-7.02918	9.24389	11.78751	0.00060	0.05982
XP_002591663	7.25030	8.90122	11.71764	0.00062	0.06110
AEE62746	7.24413	8.89509	11.70266	0.00062	0.06110
XP_002708436	-7.18053	8.55770	11.70068	0.00062	0.06110
ZP_08909002	-7.14079	8.51802	11.60401	0.00066	0.06356
XP_002939441	7.20018	8.85141	11.59600	0.00066	0.06356
XP_002607896	7.19379	8.84506	11.58050	0.00067	0.06357
CAB82366	-7.12050	8.49776	11.55467	0.00068	0.06394
XP_002609379	7.17445	8.82584	11.53360	0.00068	0.06414
XP_002435427	-7.09992	8.47722	11.50466	0.00069	0.06421
XP_002740460	7.16141	8.81288	11.50199	0.00070	0.06421
XP_002606415	6.72334	10.72772	11.47483	0.00071	0.06426
XP_002668154	7.14825	8.79981	11.47009	0.00071	0.06426
XP_001641836	-7.07905	8.45637	11.45393	0.00071	0.06426
XP_002416575	7.13497	8.78661	11.43792	0.00072	0.06426
XP_003725380	-7.06850	8.44584	11.42830	0.00072	0.06426
ADI24338	7.11482	8.76659	11.38910	0.00074	0.06513
XP_003388665	7.10122	8.75309	11.35618	0.00075	0.06545
XP_002160973	-7.02550	8.40291	11.32389	0.00077	0.06545
XP_003705881	-7.02550	8.40291	11.32389	0.00077	0.06545
XP_002594308	7.08059	8.73260	11.30624	0.00077	0.06545
XP_002943965	-7.01455	8.39198	11.29731	0.00078	0.06545
XP_003212697	-7.01455	8.39198	11.29731	0.00078	0.06545
XP_790859	-6.56744	11.21975	11.27431	0.00079	0.06546
XP_635792	-7.00351	8.38096	11.27053	0.00079	0.06546
XP_002738133	7.05262	8.70481	11.23854	0.00080	0.06580
CBJ25740	7.04554	8.69778	11.22142	0.00081	0.06580
XP_003202487	7.04554	8.69778	11.22142	0.00081	0.06580
XP_001365438	-6.96988	8.34740	11.18897	0.00082	0.06650
XP_002602545	-6.95850	8.33603	11.16136	0.00084	0.06656
XP_002604134	-6.95850	8.33603	11.16136	0.00084	0.06656
ABY20700	-6.94702	8.32458	11.13354	0.00085	0.06656
XP_003747201	-6.93546	8.31303	11.10551	0.00086	0.06656
ZP_01262732	-6.93546	8.31303	11.10551	0.00086	0.06656
XP_002740538	6.99498	8.64758	11.09920	0.00086	0.06656
XP_798461	6.99498	8.64758	11.09920	0.00086	0.06656

XP_002590771	6.98020	8.63291	11.06350	0.00088	0.06681
NP_001082876	-6.91204	8.28966	11.04877	0.00089	0.06681
AAB66556	-6.90019	8.27783	11.02007	0.00090	0.06681
CAG09487	-6.90019	8.27783	11.02007	0.00090	0.06681
XP_002427858	-6.90019	8.27783	11.02007	0.00090	0.06681
XP_783551	-6.90019	8.27783	11.02007	0.00090	0.06681
XP_001603746	-6.88824	8.26590	10.99113	0.00092	0.06743
XP_003449362	-6.87619	8.25387	10.96195	0.00093	0.06807
EHB08796	-6.86403	8.24174	10.93254	0.00094	0.06813
XP_002732889	6.91954	8.57269	10.91706	0.00095	0.06813
XP_968505	6.91954	8.57269	10.91706	0.00095	0.06813
ACO51858	-6.85178	8.22951	10.90289	0.00096	0.06813
BAK09601	-6.85178	8.22951	10.90289	0.00096	0.06813
XP_002602771	6.90397	8.55724	10.87949	0.00097	0.06840
XP_798483	-6.83942	8.21718	10.87298	0.00098	0.06840
XP_003221019	-6.82695	8.20473	10.84283	0.00099	0.06911
ZP_01254585	-6.81437	8.19218	10.81242	0.00101	0.06984
EHJ68398	-6.80169	8.17952	10.78175	0.00103	0.07042
XP_001842802	6.85621	8.50985	10.76438	0.00103	0.07042
XP_002587109	6.85621	8.50985	10.76438	0.00103	0.07042
NP_001072851	-6.78889	8.16675	10.75082	0.00104	0.07052
XP_794070	-6.77598	8.15386	10.71962	0.00106	0.07131
AEK81538	-6.76295	8.14086	10.68815	0.00108	0.07170
XP_002433513	-6.76295	8.14086	10.68815	0.00108	0.07170
AFE71369	-6.73653	8.11450	10.62436	0.00112	0.07338
XP_002739967	-6.73653	8.11450	10.62436	0.00112	0.07338
XP_002598499	6.78997	8.44415	10.60493	0.00113	0.07374
BAE39186	-6.72313	8.10113	10.59204	0.00114	0.07375
AAB33368	-6.70962	8.08764	10.55942	0.00116	0.07375
AAL99291	-6.70962	8.08764	10.55942	0.00116	0.07375
XP_002596753	-6.70962	8.08764	10.55942	0.00116	0.07375
AAH72205	6.76433	8.41872	10.54326	0.00117	0.07375
XP_002603175	6.76433	8.41872	10.54326	0.00117	0.07375
XP_001637862	6.75568	8.41014	10.52247	0.00118	0.07395
ABO61332	-6.68219	8.06028	10.49328	0.00120	0.07395
EGT39498	-6.68219	8.06028	10.49328	0.00120	0.07395
XP_003700634	-6.68219	8.06028	10.49328	0.00120	0.07395
XP_002730788	6.73822	8.39283	10.48051	0.00121	0.07395
ACN91289	-6.66828	8.04640	10.45975	0.00122	0.07395

GAA48059	-6.66828	8.04640	10.45975	0.00122	0.07395
XP_002664535	-6.66828	8.04640	10.45975	0.00122	0.07395
CBL22496	-6.64006	8.01824	10.39174	0.00127	0.07593
XP_002598321	-6.64006	8.01824	10.39174	0.00127	0.07593
XP_001633780	-6.62573	8.00395	10.35724	0.00129	0.07696
BAE90615	-6.61126	7.98951	10.32241	0.00131	0.07721
EDL08213	-6.59665	7.97493	10.28724	0.00134	0.07721
XP_001850184	-6.59665	7.97493	10.28724	0.00134	0.07721
XP_002732920	-6.59665	7.97493	10.28724	0.00134	0.07721
XP_003241032	-6.59665	7.97493	10.28724	0.00134	0.07721
XP_003382138	6.63823	8.29372	10.24058	0.00137	0.07721
XP_002741129	6.62879	8.28436	10.21795	0.00139	0.07721
XP_002193200	-6.56697	7.94533	10.21584	0.00139	0.07721
XP_003724601	-6.56697	7.94533	10.21584	0.00139	0.07721
XP_001301486	6.61929	8.27495	10.19518	0.00141	0.07721
AAX26400	-6.55190	7.93029	10.17960	0.00142	0.07721
CAB70731	-6.55190	7.93029	10.17960	0.00142	0.07721
CAH90435	-6.55190	7.93029	10.17960	0.00142	0.07721
EFX69098	-6.55190	7.93029	10.17960	0.00142	0.07721
XP_003291621	-6.55190	7.93029	10.17960	0.00142	0.07721
XP_003707588	-6.55190	7.93029	10.17960	0.00142	0.07721
XP_003723792	-6.55190	7.93029	10.17960	0.00142	0.07721
EFN60512	6.60972	8.26547	10.17227	0.00143	0.07721
AAH89352	-6.53667	7.91510	10.14300	0.00145	0.07721
CAL49346	-6.53667	7.91510	10.14300	0.00145	0.07721
XP_002613975	-6.53667	7.91510	10.14300	0.00145	0.07721
XP_002738101	-6.53667	7.91510	10.14300	0.00145	0.07721
AAN08149	-6.52127	7.89974	10.10602	0.00148	0.07721
XP_003386162	-6.52127	7.89974	10.10602	0.00148	0.07721
EAY91070	6.58063	8.23665	10.10263	0.00148	0.07721
NP_001036021	6.58063	8.23665	10.10263	0.00148	0.07721
XP_002728784	6.58063	8.23665	10.10263	0.00148	0.07721
CAJ83444	6.57080	8.22691	10.07911	0.00150	0.07721
CBQ69465	-6.50571	7.88422	10.06865	0.00151	0.07721
XP_001623551	6.56090	8.21711	10.05544	0.00152	0.07721
XP_002612285	6.55094	8.20724	10.03161	0.00154	0.07721
ZP_09111650	6.55094	8.20724	10.03161	0.00154	0.07721
AAX84973	-6.48999	7.86853	10.03089	0.00154	0.07721
AEP17623	-6.48999	7.86853	10.03089	0.00154	0.07721

XP_001807104	-6.48999	7.86853	10.03089	0.00154	0.07721
XP_002122175	-6.48999	7.86853	10.03089	0.00154	0.07721
XP_002589575	-6.48999	7.86853	10.03089	0.00154	0.07721
XP_003387894	6.54090	8.19730	10.00762	0.00156	0.07749
NP_958889	-6.47408	7.85267	9.99273	0.00157	0.07749
XP_002404411	-6.47408	7.85267	9.99273	0.00157	0.07749
XP_780780	-6.47408	7.85267	9.99273	0.00157	0.07749
AAK95589	6.53080	8.18730	9.98347	0.00158	0.07756
XP_002608858	6.52062	8.17722	9.95916	0.00160	0.07826
BAE22050	-6.44174	7.82042	9.91517	0.00164	0.07915
XP_002737009	-6.44174	7.82042	9.91517	0.00164	0.07915
XP_002738835	-6.44174	7.82042	9.91517	0.00164	0.07915
CAX74044	-6.42530	7.80401	9.87575	0.00167	0.08020
XP_002732582	-6.42530	7.80401	9.87575	0.00167	0.08020
AAL37183	-6.40866	7.78742	9.83590	0.00171	0.08034
ADR79275	-6.40866	7.78742	9.83590	0.00171	0.08034
XP_002739336	-6.40866	7.78742	9.83590	0.00171	0.08034
YP_720839	-6.40866	7.78742	9.83590	0.00171	0.08034
XP_001625155	6.46864	8.12576	9.83506	0.00171	0.08034
CBY06820	6.45801	8.11524	9.80972	0.00174	0.08067
NP_001191540	-6.39183	7.77064	9.79560	0.00175	0.08067
XP_002600408	-6.39183	7.77064	9.79560	0.00175	0.08067
XP_001373115	6.44731	8.10465	9.78420	0.00176	0.08067
ABO26647	-6.37480	7.75366	9.75484	0.00179	0.08067
BAC34934	-6.37480	7.75366	9.75484	0.00179	0.08067
CAZ27720	-6.37480	7.75366	9.75484	0.00179	0.08067
CCD21012	-6.37480	7.75366	9.75484	0.00179	0.08067
XP_001514950	-6.37480	7.75366	9.75484	0.00179	0.08067
XP_697184	-6.37480	7.75366	9.75484	0.00179	0.08067
XP_002198814	6.42566	8.08322	9.73260	0.00181	0.08102
XP_002734374	6.42566	8.08322	9.73260	0.00181	0.08102
BAK40203	-6.35757	7.73648	9.71361	0.00183	0.08155
AEN04481	6.40368	8.06148	9.68025	0.00186	0.08155
ACE75736	-6.34013	7.71909	9.67191	0.00187	0.08155
BAC06836	-6.34013	7.71909	9.67191	0.00187	0.08155
CAH04106	-6.34013	7.71909	9.67191	0.00187	0.08155
CCA37660	-6.34013	7.71909	9.67191	0.00187	0.08155
XP_002425888	-6.34013	7.71909	9.67191	0.00187	0.08155
AAH08024	-6.32247	7.70148	9.62971	0.00191	0.08263

XP_780658	-6.32247	7.70148	9.62971	0.00191	0.08263
XP_002607850	6.38136	8.03940	9.62713	0.00192	0.08263
XP_422162	-6.30460	7.68366	9.58702	0.00196	0.08384
XP_800195	-6.30460	7.68366	9.58702	0.00196	0.08384
XP_002605359	6.35869	8.01697	9.57321	0.00197	0.08385
XP_002736555	6.35869	8.01697	9.57321	0.00197	0.08385
XP_002401365	-6.28650	7.66562	9.54381	0.00201	0.08459
XP_002425306	-6.28650	7.66562	9.54381	0.00201	0.08459
AEO32528	-6.26818	7.64735	9.50007	0.00205	0.08509
BAC32100	-6.26818	7.64735	9.50007	0.00205	0.08509
EFN88349	-6.26818	7.64735	9.50007	0.00205	0.08509
XP_782331	-6.26818	7.64735	9.50007	0.00205	0.08509
ZP_05502520	-6.26818	7.64735	9.50007	0.00205	0.08509
XP_002735987	6.32400	7.98267	9.49079	0.00207	0.08522
XP_002594508	-6.24961	7.62885	9.45580	0.00210	0.08576
EFX83224	6.30040	7.95934	9.43476	0.00213	0.08576
XP_002121660	6.30040	7.95934	9.43476	0.00213	0.08576
NP_001191543	-6.23081	7.61010	9.41097	0.00216	0.08576
XP_002609620	-6.23081	7.61010	9.41097	0.00216	0.08576
YP_001999719	-6.23081	7.61010	9.41097	0.00216	0.08576
XP_002942405	5.84851	10.07905	9.40674	0.00216	0.08576
XP_001519063	6.27641	7.93562	9.37786	0.00220	0.08576
XP_002600359	6.27641	7.93562	9.37786	0.00220	0.08576
XP_003726941	6.27641	7.93562	9.37786	0.00220	0.08576
AAH90296	-6.21176	7.59111	9.36558	0.00221	0.08576
EFA01267	-6.21176	7.59111	9.36558	0.00221	0.08576
XP_001516235	-6.21176	7.59111	9.36558	0.00221	0.08576
XP_002599429	-6.21176	7.59111	9.36558	0.00221	0.08576
XP_002601662	-6.21176	7.59111	9.36558	0.00221	0.08576
XP_002741470	-6.21176	7.59111	9.36558	0.00221	0.08576
XP_003229251	-6.21176	7.59111	9.36558	0.00221	0.08576
ZP_04921972	-6.21176	7.59111	9.36558	0.00221	0.08576
EFR23189	-6.19245	7.57186	9.31961	0.00227	0.08622
XP_001516447	-6.19245	7.57186	9.31961	0.00227	0.08622
XP_002592002	-6.19245	7.57186	9.31961	0.00227	0.08622
XP_002600258	-6.19245	7.57186	9.31961	0.00227	0.08622
ZP_01066272	-6.19245	7.57186	9.31961	0.00227	0.08622
ZP_06300713	-6.19245	7.57186	9.31961	0.00227	0.08622
XP_001600879	-6.17288	7.55236	9.27304	0.00233	0.08699

XP_002611407	-6.17288	7.55236	9.27304	0.00233	0.08699
XP_002738916	-6.17288	7.55236	9.27304	0.00233	0.08699
XP_002741986	6.22720	7.88699	9.26126	0.00234	0.08699
ABY75514	5.75227	10.33727	9.23712	0.00237	0.08699
CCF33422	6.21463	7.87457	9.23151	0.00238	0.08699
XP_002592165	6.21463	7.87457	9.23151	0.00238	0.08699
CAL49314	-6.15304	7.53258	9.22586	0.00239	0.08699
XP_001633753	-6.15304	7.53258	9.22586	0.00239	0.08699
XP_002738829	-6.15304	7.53258	9.22586	0.00239	0.08699
ADK38674	-6.13292	7.51254	9.17805	0.00245	0.08699
XP_002732836	-6.13292	7.51254	9.17805	0.00245	0.08699
XP_002740852	-6.13292	7.51254	9.17805	0.00245	0.08699
AAH59873	-6.11252	7.49221	9.12960	0.00252	0.08699
AAH95783	-6.11252	7.49221	9.12960	0.00252	0.08699
EHJ24574	-6.11252	7.49221	9.12960	0.00252	0.08699
XP_001632062	-6.11252	7.49221	9.12960	0.00252	0.08699
XP_002405655	-6.11252	7.49221	9.12960	0.00252	0.08699
XP_002605874	-6.11252	7.49221	9.12960	0.00252	0.08699
XP_002607042	-6.11252	7.49221	9.12960	0.00252	0.08699
XP_002608924	-6.11252	7.49221	9.12960	0.00252	0.08699
XP_002643989	-6.11252	7.49221	9.12960	0.00252	0.08699
EFZ12036	-6.09183	7.47159	9.08049	0.00258	0.08699
EGI63691	-6.09183	7.47159	9.08049	0.00258	0.08699
XP_002735038	-6.09183	7.47159	9.08049	0.00258	0.08699
YP_003387842	-6.09183	7.47159	9.08049	0.00258	0.08699
XP_002613350	6.15009	7.81082	9.07895	0.00259	0.08699
XP_003206645	6.15009	7.81082	9.07895	0.00259	0.08699
XP_001344731	6.13683	7.79773	9.04764	0.00263	0.08699
XP_003505948	6.13683	7.79773	9.04764	0.00263	0.08699
XP_002739316	-5.65210	10.05969	9.03614	0.00265	0.08699
AAN40846	-6.07084	7.45067	9.03069	0.00265	0.08699
AEF33410	-6.07084	7.45067	9.03069	0.00265	0.08699
AEO33510	-6.07084	7.45067	9.03069	0.00265	0.08699
EHJ71735	-6.07084	7.45067	9.03069	0.00265	0.08699
XP_003724812	-6.07084	7.45067	9.03069	0.00265	0.08699
XP_647665	-6.07084	7.45067	9.03069	0.00265	0.08699
XP_003452834	5.68862	9.92200	9.02525	0.00266	0.08699
BAE22065	6.12344	7.78451	9.01606	0.00268	0.08699
ABX82529	-6.04953	7.42944	8.98020	0.00273	0.08699

XP_001946467	-6.04953	7.42944	8.98020	0.00273	0.08699
XP_002607725	-6.04953	7.42944	8.98020	0.00273	0.08699
XP_002611765	-6.04953	7.42944	8.98020	0.00273	0.08699
XP_002738623	-6.04953	7.42944	8.98020	0.00273	0.08699
XP_003262676	-6.04953	7.42944	8.98020	0.00273	0.08699
XP_003445121	-6.04953	7.42944	8.98020	0.00273	0.08699
XP_001603629	6.09629	7.75771	8.95206	0.00277	0.08699
XP_003391145	5.72988	9.18767	8.93806	0.00279	0.08699
XP_003413538	-5.56270	11.15289	8.92922	0.00281	0.08699
EFN68572	-6.02791	7.40790	8.92898	0.00281	0.08699
NP_083025	-6.02791	7.40790	8.92898	0.00281	0.08699
XP_002408245	-6.02791	7.40790	8.92898	0.00281	0.08699
XP_002591304	-6.02791	7.40790	8.92898	0.00281	0.08699
XP_002824001	-6.02791	7.40790	8.92898	0.00281	0.08699
XP_003445054	-6.02791	7.40790	8.92898	0.00281	0.08699
AAM22012	6.08252	7.74413	8.91962	0.00282	0.08699
YP_002840971	6.08252	7.74413	8.91962	0.00282	0.08699
BAE94191	6.06862	7.73041	8.88689	0.00287	0.08699
EGI62872	6.06862	7.73041	8.88689	0.00287	0.08699
XP_001634991	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_002594881	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_002604367	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_002610657	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_002736079	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_003149509	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_003443302	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_003704298	-6.00596	7.38603	8.87703	0.00289	0.08699
XP_793640	-6.00596	7.38603	8.87703	0.00289	0.08699
1713406B	-5.98367	7.36382	8.82432	0.00297	0.08699
EFX90440	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_001607321	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_002424869	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_002430886	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_002590067	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_002608259	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_003379741	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_003495220	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_003706807	-5.98367	7.36382	8.82432	0.00297	0.08699
XP_003724102	-5.98367	7.36382	8.82432	0.00297	0.08699

XP_003728083	-5.98367	7.36382	8.82432	0.00297	0.08699
EFA75352	6.04041	7.70257	8.82052	0.00298	0.08699
CAF87168	-5.96103	7.34127	8.77082	0.00306	0.08699
CCD82066	-5.96103	7.34127	8.77082	0.00306	0.08699
EHJ70170	-5.96103	7.34127	8.77082	0.00306	0.08699
P07201	-5.96103	7.34127	8.77082	0.00306	0.08699
XP_003223024	-5.96103	7.34127	8.77082	0.00306	0.08699
XP_003276168	-5.96103	7.34127	8.77082	0.00306	0.08699
XP_002735518	6.01163	7.67419	8.75289	0.00309	0.08699
XP_003706805	6.01163	7.67419	8.75289	0.00309	0.08699
XP_003727906	5.99703	7.65979	8.71860	0.00315	0.08699
AEO35054	-5.93802	7.31836	8.71651	0.00315	0.08699
CAI46182	-5.93802	7.31836	8.71651	0.00315	0.08699
CAJ30045	-5.93802	7.31836	8.71651	0.00315	0.08699
EFA04949	-5.93802	7.31836	8.71651	0.00315	0.08699
EFX88629	-5.93802	7.31836	8.71651	0.00315	0.08699
EFX89033	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_002124863	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_002597968	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_002735223	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_003255664	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_003448039	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_003448299	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_003693223	-5.93802	7.31836	8.71651	0.00315	0.08699
XP_001637561	5.98227	7.64524	8.68398	0.00321	0.08699
XP_002939897	5.98227	7.64524	8.68398	0.00321	0.08699
AAK72969	-5.91465	7.29508	8.66137	0.00325	0.08699
AAV69062	-5.91465	7.29508	8.66137	0.00325	0.08699
CAI40491	-5.91465	7.29508	8.66137	0.00325	0.08699
EFW20245	-5.91465	7.29508	8.66137	0.00325	0.08699
XP_002602361	-5.91465	7.29508	8.66137	0.00325	0.08699
XP_003728064	-5.91465	7.29508	8.66137	0.00325	0.08699
NP_001170929	5.96737	7.63054	8.64901	0.00327	0.08699
AAK12359	5.95230	7.61569	8.61371	0.00334	0.08699
XP_002606412	5.95230	7.61569	8.61371	0.00334	0.08699
XP_003209223	5.95230	7.61569	8.61371	0.00334	0.08699
XP_003364926	5.95230	7.61569	8.61371	0.00334	0.08699
AEV53960	-5.89089	7.27142	8.60537	0.00335	0.08699
CAJ83480	-5.89089	7.27142	8.60537	0.00335	0.08699

EFX69526	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_001017101	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_001508218	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_002262216	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_002595411	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_002735625	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_002815001	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_002940311	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_003428681	-5.89089	7.27142	8.60537	0.00335	0.08699
XP_789620	-5.89089	7.27142	8.60537	0.00335	0.08699
NP_001191667	5.93708	7.60069	8.57805	0.00340	0.08699
XP_002046731	5.93708	7.60069	8.57805	0.00340	0.08699
XP_784034	5.93708	7.60069	8.57805	0.00340	0.08699
ACN58711	-5.86673	7.24736	8.54848	0.00346	0.08699
CAJ26346	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_001008643	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_001426760	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_002738904	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_002752766	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_003390327	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_786649	-5.86673	7.24736	8.54848	0.00346	0.08699
XP_975362	-5.86673	7.24736	8.54848	0.00346	0.08699
YP_004179856	-5.86673	7.24736	8.54848	0.00346	0.08699
ZP_09483044	-5.86673	7.24736	8.54848	0.00346	0.08699
AAR39412	5.92170	7.58553	8.54204	0.00347	0.08699
XP_003377920	5.92170	7.58553	8.54204	0.00347	0.08699
XP_003723764	5.92170	7.58553	8.54204	0.00347	0.08699
NP_001155276	5.90615	7.57021	8.50567	0.00354	0.08699
XP_002115163	5.90615	7.57021	8.50567	0.00354	0.08699
XP_002193920	5.90615	7.57021	8.50567	0.00354	0.08699
XP_002596572	5.90615	7.57021	8.50567	0.00354	0.08699
XP_003425652	5.90615	7.57021	8.50567	0.00354	0.08699
AAY86961	-5.84216	7.22290	8.49068	0.00357	0.08699
EHH55556	-5.84216	7.22290	8.49068	0.00357	0.08699
XP_002070310	-5.84216	7.22290	8.49068	0.00357	0.08699
XP_002169253	-5.84216	7.22290	8.49068	0.00357	0.08699
XP_002588513	-5.84216	7.22290	8.49068	0.00357	0.08699
XP_002611352	-5.84216	7.22290	8.49068	0.00357	0.08699
XP_002698598	-5.84216	7.22290	8.49068	0.00357	0.08699

XP_536929	-5.84216	7.22290	8.49068	0.00357	0.08699
XP_002400528	-5.65648	7.88110	8.48731	0.00358	0.08699
AAH24689	5.89043	7.55472	8.46892	0.00361	0.08699
AAI69208	5.89043	7.55472	8.46892	0.00361	0.08699
XP_002610105	5.89043	7.55472	8.46892	0.00361	0.08699
XP_003217401	5.89043	7.55472	8.46892	0.00361	0.08699
AAT09325	-5.81717	7.19802	8.43193	0.00369	0.08699
BAA78421	-5.81717	7.19802	8.43193	0.00369	0.08699
CAX33834	-5.81717	7.19802	8.43193	0.00369	0.08699
EFN66701	-5.81717	7.19802	8.43193	0.00369	0.08699
EFN68468	-5.81717	7.19802	8.43193	0.00369	0.08699
EGI59904	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_001623622	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_001637157	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_001992750	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_002399509	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_002735917	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_002742001	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_785443	-5.81717	7.19802	8.43193	0.00369	0.08699
XP_001628920	5.87453	7.53907	8.43179	0.00369	0.08699
XP_001640068	5.87453	7.53907	8.43179	0.00369	0.08699
XP_003199319	5.87453	7.53907	8.43179	0.00369	0.08699
XP_002742380	5.65760	8.18983	8.41661	0.00372	0.08699
CAF21863	5.85847	7.52324	8.39427	0.00376	0.08699
EGR50668	5.85847	7.52324	8.39427	0.00376	0.08699
XP_002608670	5.85847	7.52324	8.39427	0.00376	0.08699
XP_002612503	5.85847	7.52324	8.39427	0.00376	0.08699
XP_002732400	5.85847	7.52324	8.39427	0.00376	0.08699
XP_419581	5.85847	7.52324	8.39427	0.00376	0.08699
AEO34023	-5.79173	7.17269	8.37221	0.00381	0.08699
BAE39153	-5.79173	7.17269	8.37221	0.00381	0.08699
CBK21306	-5.79173	7.17269	8.37221	0.00381	0.08699
EAW94898	-5.79173	7.17269	8.37221	0.00381	0.08699
EFX67003	-5.79173	7.17269	8.37221	0.00381	0.08699
EGT56672	-5.79173	7.17269	8.37221	0.00381	0.08699
NP_001245290	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_001605649	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_001633418	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_002427447	-5.79173	7.17269	8.37221	0.00381	0.08699

XP_002606880	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_002608868	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_003471921	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_785456	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_969024	-5.79173	7.17269	8.37221	0.00381	0.08699
XP_002736461	-5.33422	10.49879	8.35754	0.00384	0.08699
BAD99027	5.84221	7.50724	8.35636	0.00384	0.08699
XP_001635963	5.84221	7.50724	8.35636	0.00384	0.08699
XP_001652929	5.84221	7.50724	8.35636	0.00384	0.08699
XP_001989824	5.84221	7.50724	8.35636	0.00384	0.08699
XP_002588348	5.84221	7.50724	8.35636	0.00384	0.08699
XP_002933745	-5.33820	10.00903	8.32145	0.00392	0.08699
ADL62715	5.82578	7.49106	8.31804	0.00393	0.08699
EFA84096	5.82578	7.49106	8.31804	0.00393	0.08699
XP_002741344	5.82578	7.49106	8.31804	0.00393	0.08699
AAI69185	-5.76584	7.14692	8.31149	0.00394	0.08699
AAV84265	-5.76584	7.14692	8.31149	0.00394	0.08699
NP_001073671	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002162663	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002592492	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002601381	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002605750	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002612895	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002736557	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_422260	-5.76584	7.14692	8.31149	0.00394	0.08699
XP_002739922	5.80915	7.47469	8.27931	0.00401	0.08783
XP_791927	-5.26332	11.75548	8.25921	0.00405	0.08783
AAV48595	-5.73947	7.12068	8.24971	0.00408	0.08783
ACH89433	-5.73947	7.12068	8.24971	0.00408	0.08783
AER98881	-5.73947	7.12068	8.24971	0.00408	0.08783
NP_957183	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002190212	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002433601	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002434360	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002597222	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002607260	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002940183	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_415238	-5.73947	7.12068	8.24971	0.00408	0.08783
XP_002593922	5.79233	7.45814	8.24016	0.00410	0.08797

XP_974007	-5.54130	7.76727	8.21621	0.00415	0.08797
EDL24736	5.77532	7.44140	8.20058	0.00419	0.08797
XP_002734689	5.77532	7.44140	8.20058	0.00419	0.08797
XP_967104	5.77532	7.44140	8.20058	0.00419	0.08797
EAW75624	-5.71262	7.09395	8.18686	0.00422	0.08797
EFX83489	-5.71262	7.09395	8.18686	0.00422	0.08797
XP_001632203	-5.71262	7.09395	8.18686	0.00422	0.08797
XP_002598814	-5.71262	7.09395	8.18686	0.00422	0.08797
XP_002742202	-5.71262	7.09395	8.18686	0.00422	0.08797
XP_002413853	5.75810	7.42445	8.16055	0.00428	0.08797
XP_002593669	5.75810	7.42445	8.16055	0.00428	0.08797
XP_003391309	5.75810	7.42445	8.16055	0.00428	0.08797
2XNF_A	-5.68525	7.06672	8.12289	0.00437	0.08797
AAM91821	-5.68525	7.06672	8.12289	0.00437	0.08797
ACG75958	-5.68525	7.06672	8.12289	0.00437	0.08797
AES02438	-5.68525	7.06672	8.12289	0.00437	0.08797
AES03553	-5.68525	7.06672	8.12289	0.00437	0.08797
CAA59198	-5.68525	7.06672	8.12289	0.00437	0.08797
EFN66218	-5.68525	7.06672	8.12289	0.00437	0.08797
GAA50512	-5.68525	7.06672	8.12289	0.00437	0.08797
NP_001135486	-5.68525	7.06672	8.12289	0.00437	0.08797
NP_001164310	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_001621329	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_001627192	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_002063027	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_002613173	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_002712574	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_002734698	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_002735688	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_002942697	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_003206052	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_640602	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_701169	-5.68525	7.06672	8.12289	0.00437	0.08797
XP_785654	-5.68525	7.06672	8.12289	0.00437	0.08797
CAA67544	5.74067	7.40731	8.12008	0.00438	0.08797
CAG04340	5.74067	7.40731	8.12008	0.00438	0.08797
NP_001079493	5.74067	7.40731	8.12008	0.00438	0.08797
AAH94402	5.72303	7.38996	8.07914	0.00448	0.08797
EFX81396	5.72303	7.38996	8.07914	0.00448	0.08797

XP_002077667	5.72303	7.38996	8.07914	0.00448	0.08797
XP_417386	5.72303	7.38996	8.07914	0.00448	0.08797
YP_006354273	5.72303	7.38996	8.07914	0.00448	0.08797
AAH94088	-5.65736	7.03897	8.05776	0.00453	0.08797
AEO34555	-5.65736	7.03897	8.05776	0.00453	0.08797
BAE35720	-5.65736	7.03897	8.05776	0.00453	0.08797
NP_998032	-5.65736	7.03897	8.05776	0.00453	0.08797
XP_001944340	-5.65736	7.03897	8.05776	0.00453	0.08797
XP_001999880	-5.65736	7.03897	8.05776	0.00453	0.08797
XP_002127989	-5.65736	7.03897	8.05776	0.00453	0.08797
XP_002197995	-5.65736	7.03897	8.05776	0.00453	0.08797
XP_002587935	-5.65736	7.03897	8.05776	0.00453	0.08797
ACQ91104	-5.47110	7.69795	8.05161	0.00455	0.08797
NP_001019499	-5.47110	7.69795	8.05161	0.00455	0.08797
NP_001133096	-5.47110	7.69795	8.05161	0.00455	0.08797
XP_001624678	-5.47110	7.69795	8.05161	0.00455	0.08797
XP_003427871	5.70517	7.37240	8.03774	0.00458	0.08797
XP_003459722	5.70517	7.37240	8.03774	0.00458	0.08797
XP_796434	5.70517	7.37240	8.03774	0.00458	0.08797
AEK10750	-5.20312	9.99028	8.01498	0.00464	0.08797
XP_002607858	5.68708	7.35462	7.99585	0.00469	0.08797
ACN91277	-5.62891	7.01067	7.99144	0.00470	0.08797
ADX31291	-5.62891	7.01067	7.99144	0.00470	0.08797
BAE39125	-5.62891	7.01067	7.99144	0.00470	0.08797
CAI21694	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_001121484	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_001184164	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_001188982	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_001201290	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_001863343	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_002595371	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_002610132	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_003220431	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_003223090	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_003247452	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_003293760	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_793995	-5.62891	7.01067	7.99144	0.00470	0.08797
XP_001639778	5.66877	7.33662	7.95346	0.00480	0.08797
XP_002610097	5.66877	7.33662	7.95346	0.00480	0.08797

ABS88697	-5.59990	6.98180	7.92386	0.00488	0.08797
ACO08993	-5.59990	6.98180	7.92386	0.00488	0.08797
AEO36020	-5.59990	6.98180	7.92386	0.00488	0.08797
EFX73186	-5.59990	6.98180	7.92386	0.00488	0.08797
EFX88195	-5.59990	6.98180	7.92386	0.00488	0.08797
NP_001082284	-5.59990	6.98180	7.92386	0.00488	0.08797
NP_001128702	-5.59990	6.98180	7.92386	0.00488	0.08797
Q964E3	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_001917051	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002431014	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002431121	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002606187	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002733477	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002733718	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002736186	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_002741112	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_003143113	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_003397232	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_003745138	-5.59990	6.98180	7.92386	0.00488	0.08797
XP_850799	-5.59990	6.98180	7.92386	0.00488	0.08797
YP_747305	-5.59990	6.98180	7.92386	0.00488	0.08797
AAH65041	5.65022	7.31839	7.91057	0.00491	0.08797
AAX27930	5.65022	7.31839	7.91057	0.00491	0.08797
AES05754	5.65022	7.31839	7.91057	0.00491	0.08797
BAA94854	5.65022	7.31839	7.91057	0.00491	0.08797
XP_002148768	5.65022	7.31839	7.91057	0.00491	0.08797
YP_002841142	5.65022	7.31839	7.91057	0.00491	0.08797
BAD99026	5.63143	7.29993	7.86717	0.00503	0.08797
NP_001139049	5.63143	7.29993	7.86717	0.00503	0.08797
XP_001313629	5.63143	7.29993	7.86717	0.00503	0.08797
XP_002590332	5.63143	7.29993	7.86717	0.00503	0.08797
XP_002680168	5.63143	7.29993	7.86717	0.00503	0.08797
XP_003443882	5.63143	7.29993	7.86717	0.00503	0.08797
3R2B_A	-5.57028	6.95235	7.85499	0.00507	0.08797
AAI61792	-5.57028	6.95235	7.85499	0.00507	0.08797
ACN91297	-5.57028	6.95235	7.85499	0.00507	0.08797
ADD18698	-5.57028	6.95235	7.85499	0.00507	0.08797
CAF33263	-5.57028	6.95235	7.85499	0.00507	0.08797
NP_001009986	-5.57028	6.95235	7.85499	0.00507	0.08797

XP_001636613	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_001640172	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_001647647	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_001895717	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_002131970	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_002591357	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_002603898	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_002732174	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_312429	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_320464	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_636792	-5.57028	6.95235	7.85499	0.00507	0.08797
XP_797423	-5.57028	6.95235	7.85499	0.00507	0.08797
EDL16829	5.61239	7.28123	7.82323	0.00516	0.08797
XP_002737958	5.61239	7.28123	7.82323	0.00516	0.08797
XP_003365356	5.61239	7.28123	7.82323	0.00516	0.08797
XP_629009	5.61239	7.28123	7.82323	0.00516	0.08797
XP_001019230	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_001296210	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_001329169	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002435466	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002593187	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002735710	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002736550	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002738667	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002739921	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_003213845	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_003498564	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_424694	-5.54005	6.92228	7.78478	0.00527	0.08797
XP_002606696	5.59310	7.26229	7.77875	0.00529	0.08797
XP_003198228	5.59310	7.26229	7.77875	0.00529	0.08797
AAP41214	5.57354	7.24309	7.73371	0.00542	0.08797
XP_001316844	5.57354	7.24309	7.73371	0.00542	0.08797
XP_002730929	5.57354	7.24309	7.73371	0.00542	0.08797
XP_002912748	5.57354	7.24309	7.73371	0.00542	0.08797
CAI15956	-5.50917	6.89158	7.71317	0.00548	0.08797
DAA16422	-5.50917	6.89158	7.71317	0.00548	0.08797
EGW05886	-5.50917	6.89158	7.71317	0.00548	0.08797
EHA98399	-5.50917	6.89158	7.71317	0.00548	0.08797
EHB09998	-5.50917	6.89158	7.71317	0.00548	0.08797

XP_001351550	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_001367489	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_001629536	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_001772366	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_001842282	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_001844033	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_001994475	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002588882	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002592263	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002599089	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002605575	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002730499	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002739059	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002741281	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002742061	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_002940872	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_003731438	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_646649	-5.50917	6.89158	7.71317	0.00548	0.08797
XP_788604	-5.50917	6.89158	7.71317	0.00548	0.08797
ZP_08205869	-5.50917	6.89158	7.71317	0.00548	0.08797
AAC00207	5.55372	7.22363	7.68810	0.00556	0.08797
AAI46846	5.55372	7.22363	7.68810	0.00556	0.08797
EFZ11915	5.55372	7.22363	7.68810	0.00556	0.08797
NP_001087022	5.55372	7.22363	7.68810	0.00556	0.08797
XP_002129966	5.55372	7.22363	7.68810	0.00556	0.08797
XP_002602847	5.55372	7.22363	7.68810	0.00556	0.08797
XP_002738789	5.55372	7.22363	7.68810	0.00556	0.08797
XP_003742500	5.55372	7.22363	7.68810	0.00556	0.08797
XP_973407	5.55372	7.22363	7.68810	0.00556	0.08797
YP_001958455	5.33893	7.87769	7.67378	0.00560	0.08797
AEO35323	5.53362	7.20391	7.64190	0.00570	0.08797
XP_002602373	5.53362	7.20391	7.64190	0.00570	0.08797
XP_002734510	5.53362	7.20391	7.64190	0.00570	0.08797
XP_002734543	5.53362	7.20391	7.64190	0.00570	0.08797
XP_002734661	5.53362	7.20391	7.64190	0.00570	0.08797
XP_002737120	5.53362	7.20391	7.64190	0.00570	0.08797
XP_003117370	5.53362	7.20391	7.64190	0.00570	0.08797
XP_003724058	5.53362	7.20391	7.64190	0.00570	0.08797
ADP08789	-5.47762	6.86021	7.64011	0.00571	0.08797

NP_001008648	-5.47762	6.86021	7.64011	0.00571	0.08797
NP_573311	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_001375787	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_001631865	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002115969	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002400122	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002596547	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002732306	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002735833	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002740846	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_003729354	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_783928	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_787226	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_970879	-5.47762	6.86021	7.64011	0.00571	0.08797
XP_002414754	-5.27902	7.50853	7.60378	0.00582	0.08797
1N0R_A	5.51324	7.18391	7.59510	0.00585	0.08797
AAF75839	5.51324	7.18391	7.59510	0.00585	0.08797
XP_002577332	5.51324	7.18391	7.59510	0.00585	0.08797
XP_002943235	5.51324	7.18391	7.59510	0.00585	0.08797
XP_395916	5.51324	7.18391	7.59510	0.00585	0.08797
AAH07747	-5.44536	6.82814	7.56554	0.00595	0.08797
ABF21059	-5.44536	6.82814	7.56554	0.00595	0.08797
AEO35704	-5.44536	6.82814	7.56554	0.00595	0.08797
CAB38180	-5.44536	6.82814	7.56554	0.00595	0.08797
EFX88686	-5.44536	6.82814	7.56554	0.00595	0.08797
EGI62678	-5.44536	6.82814	7.56554	0.00595	0.08797
EHB15871	-5.44536	6.82814	7.56554	0.00595	0.08797
NP_001072224	-5.44536	6.82814	7.56554	0.00595	0.08797
NP_001171814	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_001013132	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_001663322	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_001914755	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002078082	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002427810	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002431913	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002603723	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002609740	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002613805	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002614074	-5.44536	6.82814	7.56554	0.00595	0.08797

XP_002732563	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002734654	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002738314	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002739885	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_002831832	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_003389851	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_003437908	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_003490045	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_003494875	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_003727446	-5.44536	6.82814	7.56554	0.00595	0.08797
XP_417571	-5.44536	6.82814	7.56554	0.00595	0.08797
ZP_09965478	-5.44536	6.82814	7.56554	0.00595	0.08797
EFX86378	5.49256	7.16363	7.54769	0.00601	0.08806
XP_001379766	5.49256	7.16363	7.54769	0.00601	0.08806
XP_002736138	5.49256	7.16363	7.54769	0.00601	0.08806
XP_003445167	5.49256	7.16363	7.54769	0.00601	0.08806
XP_003731784	5.49256	7.16363	7.54769	0.00601	0.08806
XP_319948	5.49256	7.16363	7.54769	0.00601	0.08806
XP_675648	5.49256	7.16363	7.54769	0.00601	0.08806
XP_002731452	-5.23733	7.46747	7.50711	0.00615	0.08840
AAH42646	5.47159	7.14306	7.49964	0.00617	0.08840
XP_001629191	5.47159	7.14306	7.49964	0.00617	0.08840
XP_001657105	5.47159	7.14306	7.49964	0.00617	0.08840
XP_002732972	5.47159	7.14306	7.49964	0.00617	0.08840
XP_003216403	5.47159	7.14306	7.49964	0.00617	0.08840
XP_003503117	5.47159	7.14306	7.49964	0.00617	0.08840
AAH76415	-5.41236	6.79534	7.48938	0.00621	0.08840
BAE79743	-5.41236	6.79534	7.48938	0.00621	0.08840
EFZ21157	-5.41236	6.79534	7.48938	0.00621	0.08840
P86856	-5.41236	6.79534	7.48938	0.00621	0.08840
Q86MA7	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_001622679	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_001638741	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_001639688	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_002607271	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_003730972	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_003731756	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_312357	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_781600	-5.41236	6.79534	7.48938	0.00621	0.08840

XP_790269	-5.41236	6.79534	7.48938	0.00621	0.08840
XP_793314	-5.41236	6.79534	7.48938	0.00621	0.08840
YP_004894149	-5.41236	6.79534	7.48938	0.00621	0.08840
CAX13385	5.45030	7.12219	7.45094	0.00634	0.08911
CCD77454	5.45030	7.12219	7.45094	0.00634	0.08911
EFN63056	5.45030	7.12219	7.45094	0.00634	0.08911
NP_001158362	5.45030	7.12219	7.45094	0.00634	0.08911
XP_002603103	5.45030	7.12219	7.45094	0.00634	0.08911
XP_002608536	5.45030	7.12219	7.45094	0.00634	0.08911
XP_002751314	5.45030	7.12219	7.45094	0.00634	0.08911
XP_003240425	5.45030	7.12219	7.45094	0.00634	0.08911
XP_785648	5.45030	7.12219	7.45094	0.00634	0.08911
ZP_04532941	5.45030	7.12219	7.45094	0.00634	0.08911
ZP_08102074	5.45030	7.12219	7.45094	0.00634	0.08911
XP_002426989	-5.11568	7.87796	7.41962	0.00645	0.08935
BAB27697	-5.37858	6.76178	7.41158	0.00648	0.08935
BAH12781	-5.37858	6.76178	7.41158	0.00648	0.08935
BAJ24842	-5.37858	6.76178	7.41158	0.00648	0.08935
DAA34668	-5.37858	6.76178	7.41158	0.00648	0.08935
H2A0M0	-5.37858	6.76178	7.41158	0.00648	0.08935
NP_001121513	-5.37858	6.76178	7.41158	0.00648	0.08935
NP_724543	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_001491293	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_002416006	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_002586063	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_002605569	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_002739909	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_003724122	-5.37858	6.76178	7.41158	0.00648	0.08935
XP_796961	-5.37858	6.76178	7.41158	0.00648	0.08935
ABJ97377	5.42869	7.10102	7.40157	0.00652	0.08935
CAI11834	5.42869	7.10102	7.40157	0.00652	0.08935
NP_001128543	5.42869	7.10102	7.40157	0.00652	0.08935
XP_002604515	5.42869	7.10102	7.40157	0.00652	0.08935
XP_796030	5.42869	7.10102	7.40157	0.00652	0.08935
AFC98245	-5.05924	8.20967	7.40096	0.00652	0.08935
XP_002741846	5.40676	7.07953	7.35151	0.00670	0.09092
XP_003382559	5.40676	7.07953	7.35151	0.00670	0.09092
XP_003440100	5.40676	7.07953	7.35151	0.00670	0.09092
XP_003727967	5.40676	7.07953	7.35151	0.00670	0.09092

AAC46490	-5.34400	6.72743	7.33207	0.00677	0.09092
AEO34511	-5.34400	6.72743	7.33207	0.00677	0.09092
BAF63789	-5.34400	6.72743	7.33207	0.00677	0.09092
EGI70977	-5.34400	6.72743	7.33207	0.00677	0.09092
NP_001191639	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_002130339	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_002131378	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_002601230	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_002733462	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_002738807	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_002989592	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_003244885	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_414938	-5.34400	6.72743	7.33207	0.00677	0.09092
XP_800759	-5.34400	6.72743	7.33207	0.00677	0.09092
AAI61240	5.38449	7.05772	7.30075	0.00689	0.09177
BAH84829	5.38449	7.05772	7.30075	0.00689	0.09177
EFN81652	5.38449	7.05772	7.30075	0.00689	0.09177
XP_002128280	5.38449	7.05772	7.30075	0.00689	0.09177
XP_002319013	5.38449	7.05772	7.30075	0.00689	0.09177
XP_002427237	5.38449	7.05772	7.30075	0.00689	0.09177
XP_317010	5.38449	7.05772	7.30075	0.00689	0.09177
XP_002607353	4.94504	9.19652	7.27924	0.00698	0.09192
AAD12246	-5.30857	6.69223	7.25075	0.00709	0.09192
AAH08328	-5.30857	6.69223	7.25075	0.00709	0.09192
AFE74132	-5.30857	6.69223	7.25075	0.00709	0.09192
EFB18688	-5.30857	6.69223	7.25075	0.00709	0.09192
JC8022	-5.30857	6.69223	7.25075	0.00709	0.09192
NP_001139570	-5.30857	6.69223	7.25075	0.00709	0.09192
NP_001242994	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_001363658	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_002187407	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_002595765	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_002595800	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_002718217	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_002731212	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_003453111	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_787033	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_795044	-5.30857	6.69223	7.25075	0.00709	0.09192
XP_968982	-5.30857	6.69223	7.25075	0.00709	0.09192

XP_969067	-5.30857	6.69223	7.25075	0.00709	0.09192
YP_629985	-5.30857	6.69223	7.25075	0.00709	0.09192
ZP_01453154	-5.30857	6.69223	7.25075	0.00709	0.09192
CAA69658	5.36187	7.03557	7.24926	0.00709	0.09192
XP_002938053	5.36187	7.03557	7.24926	0.00709	0.09192
XP_003345830	5.36187	7.03557	7.24926	0.00709	0.09192
AAV85465	4.97062	8.75321	7.24129	0.00712	0.09222
AAC35953	-4.94694	8.40500	7.21406	0.00723	0.09309
AAH81274	5.33889	7.01307	7.19702	0.00730	0.09309
XP_002578551	5.33889	7.01307	7.19702	0.00730	0.09309
XP_002589335	5.33889	7.01307	7.19702	0.00730	0.09309
XP_002590403	5.33889	7.01307	7.19702	0.00730	0.09309
XP_002607686	5.33889	7.01307	7.19702	0.00730	0.09309
XP_002919141	5.33889	7.01307	7.19702	0.00730	0.09309
XP_003494521	5.33889	7.01307	7.19702	0.00730	0.09309
XP_003723431	5.33889	7.01307	7.19702	0.00730	0.09309
XP_003725724	5.33889	7.01307	7.19702	0.00730	0.09309
ACU83221	-5.27225	6.65616	7.16756	0.00742	0.09309
AEG78368	-5.27225	6.65616	7.16756	0.00742	0.09309
AEO32292	-5.27225	6.65616	7.16756	0.00742	0.09309
BAH14572	-5.27225	6.65616	7.16756	0.00742	0.09309
CAC82191	-5.27225	6.65616	7.16756	0.00742	0.09309
EAW68433	-5.27225	6.65616	7.16756	0.00742	0.09309
EFX84014	-5.27225	6.65616	7.16756	0.00742	0.09309
EFZ15737	-5.27225	6.65616	7.16756	0.00742	0.09309
NP_001134210	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_001627242	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_002109086	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_002131963	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_002590806	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_002735979	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_002738454	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_003729203	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_003742214	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_793154	-5.27225	6.65616	7.16756	0.00742	0.09309
ZP_09232839	-5.27225	6.65616	7.16756	0.00742	0.09309
XP_002609800	-5.08115	7.31380	7.14676	0.00751	0.09331
BAC34619	5.31554	6.99022	7.14402	0.00752	0.09331
XP_001021944	5.31554	6.99022	7.14402	0.00752	0.09331

XP_002196829	5.31554	6.99022	7.14402	0.00752	0.09331
XP_002595898	5.31554	6.99022	7.14402	0.00752	0.09331
XP_002738406	5.31554	6.99022	7.14402	0.00752	0.09331
XP_002741442	5.31554	6.99022	7.14402	0.00752	0.09331
XP_003515270	5.31554	6.99022	7.14402	0.00752	0.09331
XP_003726678	5.31554	6.99022	7.14402	0.00752	0.09331
XP_396524	5.31554	6.99022	7.14402	0.00752	0.09331
AAI05210	5.29180	6.96700	7.09022	0.00775	0.09358
EDL09901	5.29180	6.96700	7.09022	0.00775	0.09358
EHJ78503	5.29180	6.96700	7.09022	0.00775	0.09358
XP_002733315	5.29180	6.96700	7.09022	0.00775	0.09358
XP_002737235	5.29180	6.96700	7.09022	0.00775	0.09358
XP_002741523	5.29180	6.96700	7.09022	0.00775	0.09358
XP_003219674	5.29180	6.96700	7.09022	0.00775	0.09358
XP_003440255	5.29180	6.96700	7.09022	0.00775	0.09358
XP_623578	5.29180	6.96700	7.09022	0.00775	0.09358
XP_625002	5.29180	6.96700	7.09022	0.00775	0.09358
ADQ43243	-5.23499	6.61917	7.08240	0.00778	0.09358
BAH70506	-5.23499	6.61917	7.08240	0.00778	0.09358
BAM17961	-5.23499	6.61917	7.08240	0.00778	0.09358
C3YWU0	-5.23499	6.61917	7.08240	0.00778	0.09358
CAJ38815	-5.23499	6.61917	7.08240	0.00778	0.09358
EFX84334	-5.23499	6.61917	7.08240	0.00778	0.09358
NP_001079840	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002081583	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002167634	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002410016	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002611888	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002716385	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002732810	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002741428	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_002904935	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_003284143	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_003447492	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_003455212	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_644124	-5.23499	6.61917	7.08240	0.00778	0.09358
XP_728306	-5.23499	6.61917	7.08240	0.00778	0.09358
AAH23438	5.26766	6.94340	7.03560	0.00799	0.09459
CAD79439	5.26766	6.94340	7.03560	0.00799	0.09459

NP_001028276	5.26766	6.94340	7.03560	0.00799	0.09459
XP_002121377	5.26766	6.94340	7.03560	0.00799	0.09459
XP_002603744	5.26766	6.94340	7.03560	0.00799	0.09459
XP_002922184	5.26766	6.94340	7.03560	0.00799	0.09459
XP_003704775	5.26766	6.94340	7.03560	0.00799	0.09459
ACO09839	-5.19674	6.58120	6.99518	0.00817	0.09459
ADI78068	-5.19674	6.58120	6.99518	0.00817	0.09459
AES07133	-5.19674	6.58120	6.99518	0.00817	0.09459
BAK52806	-5.19674	6.58120	6.99518	0.00817	0.09459
BAK61936	-5.19674	6.58120	6.99518	0.00817	0.09459
BAM20344	-5.19674	6.58120	6.99518	0.00817	0.09459
EFN86147	-5.19674	6.58120	6.99518	0.00817	0.09459
EGT46744	-5.19674	6.58120	6.99518	0.00817	0.09459
EGU89445	-5.19674	6.58120	6.99518	0.00817	0.09459
NP_001120175	-5.19674	6.58120	6.99518	0.00817	0.09459
NP_001161644	-5.19674	6.58120	6.99518	0.00817	0.09459
NP_001187438	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_001634032	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002596805	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002605354	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002609809	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002610578	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002611948	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002736837	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002737332	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002742014	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_002918754	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_003383371	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_003706967	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_003728329	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_652421	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_788973	-5.19674	6.58120	6.99518	0.00817	0.09459
XP_797784	-5.19674	6.58120	6.99518	0.00817	0.09459
AAF08010	5.24312	6.91940	6.98014	0.00824	0.09459
AAI30064	5.24312	6.91940	6.98014	0.00824	0.09459
ABN58714	5.24312	6.91940	6.98014	0.00824	0.09459
EHB18107	5.24312	6.91940	6.98014	0.00824	0.09459
NP_001096464	5.24312	6.91940	6.98014	0.00824	0.09459
XP_001630804	5.24312	6.91940	6.98014	0.00824	0.09459

XP_001638183	5.24312	6.91940	6.98014	0.00824	0.09459
XP_002427605	5.24312	6.91940	6.98014	0.00824	0.09459
XP_002732982	5.24312	6.91940	6.98014	0.00824	0.09459
XP_002741356	5.24312	6.91940	6.98014	0.00824	0.09459
XP_003706024	5.24312	6.91940	6.98014	0.00824	0.09459
EFN77118	-4.90977	7.67650	6.94526	0.00840	0.09504
AEF33431	5.21815	6.89500	6.92381	0.00851	0.09504
EGZ11335	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002169156	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002423593	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002427334	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002603725	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002609157	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002612048	5.21815	6.89500	6.92381	0.00851	0.09504
XP_002737647	5.21815	6.89500	6.92381	0.00851	0.09504
XP_003199942	5.21815	6.89500	6.92381	0.00851	0.09504
XP_003201532	5.21815	6.89500	6.92381	0.00851	0.09504
XP_003451110	5.21815	6.89500	6.92381	0.00851	0.09504
AAB94002	-5.15745	6.54221	6.90580	0.00859	0.09504
AAX59986	-5.15745	6.54221	6.90580	0.00859	0.09504
CAA03852	-5.15745	6.54221	6.90580	0.00859	0.09504
CAH72156	-5.15745	6.54221	6.90580	0.00859	0.09504
CAM16369	-5.15745	6.54221	6.90580	0.00859	0.09504
CCC90229	-5.15745	6.54221	6.90580	0.00859	0.09504
EAW92073	-5.15745	6.54221	6.90580	0.00859	0.09504
EHB01690	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_001073612	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_001370030	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_001622790	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_001730439	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002121304	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002594115	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002594726	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002599337	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002606340	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002606484	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002606952	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002607276	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002613830	-5.15745	6.54221	6.90580	0.00859	0.09504

XP_002737170	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_002740120	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_003459606	-5.15745	6.54221	6.90580	0.00859	0.09504
XP_003743136	-5.15745	6.54221	6.90580	0.00859	0.09504
AAM12624	5.19274	6.87018	6.86659	0.00878	0.09607
ABF18435	5.19274	6.87018	6.86659	0.00878	0.09607
BAI78310	5.19274	6.87018	6.86659	0.00878	0.09607
XP_001655365	5.19274	6.87018	6.86659	0.00878	0.09607
XP_001917505	5.19274	6.87018	6.86659	0.00878	0.09607
XP_002597495	5.19274	6.87018	6.86659	0.00878	0.09607
XP_002603507	5.19274	6.87018	6.86659	0.00878	0.09607
XP_700884	5.19274	6.87018	6.86659	0.00878	0.09607
XP_002735254	-4.82453	7.98076	6.86034	0.00881	0.09607
XP_003704588	-4.82453	7.98076	6.86034	0.00881	0.09607
AAN71247	-5.11706	6.50214	6.81414	0.00904	0.09607
AAO16597	-5.11706	6.50214	6.81414	0.00904	0.09607
ABJ17046	-5.11706	6.50214	6.81414	0.00904	0.09607
ABZ04225	-5.11706	6.50214	6.81414	0.00904	0.09607
ACN25141	-5.11706	6.50214	6.81414	0.00904	0.09607
ACY92475	-5.11706	6.50214	6.81414	0.00904	0.09607
ADY43920	-5.11706	6.50214	6.81414	0.00904	0.09607
EFN60550	-5.11706	6.50214	6.81414	0.00904	0.09607
EIM88312	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_001623400	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_001624183	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002191916	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002599532	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002603910	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002613713	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002730816	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002735108	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002737296	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002739880	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002740327	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_002941352	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_003230829	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_003249185	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_003377489	-5.11706	6.50214	6.81414	0.00904	0.09607
XP_003416137	-5.11706	6.50214	6.81414	0.00904	0.09607

XP_792701	-5.11706	6.50214	6.81414	0.00904	0.09607
ABK63640	5.16687	6.84493	6.80844	0.00907	0.09607
AFA34361	5.16687	6.84493	6.80844	0.00907	0.09607
CAF96009	5.16687	6.84493	6.80844	0.00907	0.09607
NP_001177233	5.16687	6.84493	6.80844	0.00907	0.09607
XP_001638576	5.16687	6.84493	6.80844	0.00907	0.09607
XP_002118044	5.16687	6.84493	6.80844	0.00907	0.09607
XP_002120553	5.16687	6.84493	6.80844	0.00907	0.09607
XP_002419737	5.16687	6.84493	6.80844	0.00907	0.09607
XP_002731635	5.16687	6.84493	6.80844	0.00907	0.09607
XP_003412872	5.16687	6.84493	6.80844	0.00907	0.09607
XP_003443814	5.16687	6.84493	6.80844	0.00907	0.09607
BAE36201	5.14054	6.81922	6.74934	0.00938	0.09740
CAQ14975	5.14054	6.81922	6.74934	0.00938	0.09740
NP_001011952	5.14054	6.81922	6.74934	0.00938	0.09740
XP_002524490	5.14054	6.81922	6.74934	0.00938	0.09740
XP_002586461	5.14054	6.81922	6.74934	0.00938	0.09740
XP_002597866	5.14054	6.81922	6.74934	0.00938	0.09740
XP_002740084	5.14054	6.81922	6.74934	0.00938	0.09740
XP_003514888	5.14054	6.81922	6.74934	0.00938	0.09740
AAH20198	-5.07551	6.46092	6.72009	0.00953	0.09740
AAI42797	-5.07551	6.46092	6.72009	0.00953	0.09740
AAX26856	-5.07551	6.46092	6.72009	0.00953	0.09740
AES11373	-5.07551	6.46092	6.72009	0.00953	0.09740
BAE00688	-5.07551	6.46092	6.72009	0.00953	0.09740
CAG05170	-5.07551	6.46092	6.72009	0.00953	0.09740
CBX41740	-5.07551	6.46092	6.72009	0.00953	0.09740
DAA01766	-5.07551	6.46092	6.72009	0.00953	0.09740
EFX90172	-5.07551	6.46092	6.72009	0.00953	0.09740
P25324	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001630985	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001632595	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001648420	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001650130	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001747261	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001811118	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001901240	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002107929	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002589568	-5.07551	6.46092	6.72009	0.00953	0.09740

XP_002595454	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002599181	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002602987	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002608254	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002671735	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002708625	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002732024	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_002734349	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_003224930	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_003486382	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_003693496	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_392313	-5.07551	6.46092	6.72009	0.00953	0.09740
YP_501570	-5.07551	6.46092	6.72009	0.00953	0.09740
XP_001861636	-4.66810	8.95223	6.71439	0.00956	0.09763
AAH56009	5.11371	6.79305	6.68925	0.00970	0.09798
AAS86702	5.11371	6.79305	6.68925	0.00970	0.09798
ABR68007	5.11371	6.79305	6.68925	0.00970	0.09798
ACL68661	5.11371	6.79305	6.68925	0.00970	0.09798
ADU19852	5.11371	6.79305	6.68925	0.00970	0.09798
AEF33398	5.11371	6.79305	6.68925	0.00970	0.09798
AEO33549	5.11371	6.79305	6.68925	0.00970	0.09798
CAX14245	5.11371	6.79305	6.68925	0.00970	0.09798
XP_001098430	5.11371	6.79305	6.68925	0.00970	0.09798
XP_002190622	5.11371	6.79305	6.68925	0.00970	0.09798
XP_002585735	5.11371	6.79305	6.68925	0.00970	0.09798
XP_002741525	5.11371	6.79305	6.68925	0.00970	0.09798
AAL40415	-4.59598	10.01389	6.65657	0.00988	0.09868
3ZU7_B	5.08637	6.76639	6.62814	0.01004	0.09868
ACA53359	5.08637	6.76639	6.62814	0.01004	0.09868
AEO34496	5.08637	6.76639	6.62814	0.01004	0.09868
BAB24599	5.08637	6.76639	6.62814	0.01004	0.09868
NP_001003525	5.08637	6.76639	6.62814	0.01004	0.09868
NP_001246171	5.08637	6.76639	6.62814	0.01004	0.09868
XP_002733177	5.08637	6.76639	6.62814	0.01004	0.09868
XP_002735577	5.08637	6.76639	6.62814	0.01004	0.09868
XP_003218952	5.08637	6.76639	6.62814	0.01004	0.09868
XP_003444610	5.08637	6.76639	6.62814	0.01004	0.09868
CAF95840	-5.03272	6.41850	6.62353	0.01006	0.09868
DAA01286	-5.03272	6.41850	6.62353	0.01006	0.09868

EFR22584	-5.03272	6.41850	6.62353	0.01006	0.09868
NP_001073476	-5.03272	6.41850	6.62353	0.01006	0.09868
NP_997822	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_001381503	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_001600405	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_001623090	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002161139	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002422774	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002431934	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002593165	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002612124	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002730573	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002735319	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_002742026	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_003215978	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_003227281	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_003228880	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_003437657	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_003480569	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_623653	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_968186	-5.03272	6.41850	6.62353	0.01006	0.09868
XP_001509166	-4.85175	7.08866	6.62322	0.01007	0.09868
AAH48199	5.05851	6.73923	6.56597	0.01039	0.09978
EDL08506	5.05851	6.73923	6.56597	0.01039	0.09978
EFN87015	5.05851	6.73923	6.56597	0.01039	0.09978
EFX89628	5.05851	6.73923	6.56597	0.01039	0.09978
EHJ77866	5.05851	6.73923	6.56597	0.01039	0.09978
NP_001085695	5.05851	6.73923	6.56597	0.01039	0.09978
NP_001087707	5.05851	6.73923	6.56597	0.01039	0.09978
NP_758033	5.05851	6.73923	6.56597	0.01039	0.09978
XP_001371296	5.05851	6.73923	6.56597	0.01039	0.09978
XP_002076869	5.05851	6.73923	6.56597	0.01039	0.09978
XP_002586024	5.05851	6.73923	6.56597	0.01039	0.09978
XP_002592162	5.05851	6.73923	6.56597	0.01039	0.09978
XP_002592437	5.05851	6.73923	6.56597	0.01039	0.09978
XP_002600896	5.05851	6.73923	6.56597	0.01039	0.09978
XP_002607406	5.05851	6.73923	6.56597	0.01039	0.09978
XP_003443102	5.05851	6.73923	6.56597	0.01039	0.09978
XP_003726012	5.05851	6.73923	6.56597	0.01039	0.09978

XP_003746101	5.05851	6.73923	6.56597	0.01039	0.09978
XP_782883	5.05851	6.73923	6.56597	0.01039	0.09978
XP_787047	5.05851	6.73923	6.56597	0.01039	0.09978
XP_003389557	-4.82386	7.06133	6.56007	0.01043	0.09978
XP_003382892	-4.62174	8.55560	6.55695	0.01045	0.09978
AAG42824	-4.98863	6.37479	6.52430	0.01064	0.09978
AAI42789	-4.98863	6.37479	6.52430	0.01064	0.09978
AAK96227	-4.98863	6.37479	6.52430	0.01064	0.09978
AAM18869	-4.98863	6.37479	6.52430	0.01064	0.09978
AAT97079	-4.98863	6.37479	6.52430	0.01064	0.09978
AAW24718	-4.98863	6.37479	6.52430	0.01064	0.09978
AEO34520	-4.98863	6.37479	6.52430	0.01064	0.09978
CAP19429	-4.98863	6.37479	6.52430	0.01064	0.09978
EAW89598	-4.98863	6.37479	6.52430	0.01064	0.09978
EGD80777	-4.98863	6.37479	6.52430	0.01064	0.09978
NP_001103594	-4.98863	6.37479	6.52430	0.01064	0.09978
NP_001177260	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_001601734	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_001635714	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_001640621	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_001659775	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_001956625	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002105029	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002588214	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002589790	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002593069	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002601232	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002601620	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002732898	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002738217	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_002834412	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_003230802	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_003243638	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_003384634	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_003726964	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_785073	-4.98863	6.37479	6.52430	0.01064	0.09978
XP_790246	-4.98863	6.37479	6.52430	0.01064	0.09978
CCD78935	5.03010	6.71154	6.50271	0.01077	0.09995
EEE31154	5.03010	6.71154	6.50271	0.01077	0.09995

EFZ17875	5.03010	6.71154	6.50271	0.01077	0.09995
EGT39833	5.03010	6.71154	6.50271	0.01077	0.09995
XP_001301665	5.03010	6.71154	6.50271	0.01077	0.09995
XP_001438334	5.03010	6.71154	6.50271	0.01077	0.09995
XP_001497503	5.03010	6.71154	6.50271	0.01077	0.09995
XP_001622021	5.03010	6.71154	6.50271	0.01077	0.09995
XP_002032980	5.03010	6.71154	6.50271	0.01077	0.09995
XP_002608279	5.03010	6.71154	6.50271	0.01077	0.09995
XP_002730917	5.03010	6.71154	6.50271	0.01077	0.09995
XP_002734366	5.03010	6.71154	6.50271	0.01077	0.09995
XP_002736788	5.03010	6.71154	6.50271	0.01077	0.09995

Figure S4. Gene Ontology (GO) terms with highest significance levels (Fisher's exact test) in (A) upregulated and (B) downregulated unigenes.

(A)

GO.ID	Term	Annotated	Signif.	Expected	Classic Fisher
GO:0071695	anatomical structure maturation	20	6	0.690000	0.00004
GO:0007165	signal transduction	1395	75	48.390000	0.00001
GO:0055085	transmembrane transport	337	27	11.690000	0.00003
GO:0006414	translational elongation	145	16	5.030000	0.00004
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	91	12	3.160000	0.00006
GO:0048738	cardiac muscle tissue development	49	8	1.700000	0.00024
GO:0047496	vesicle transport along microtubule	28	6	0.970000	0.00032
GO:0010927	cellular component assembly involved in morphogenesis	182	16	6.310000	0.00052
GO:0006415	translational termination	83	10	2.880000	0.00054
GO:0006413	translational initiation	156	14	5.410000	0.00095
GO:0009207	purine ribonucleoside triphosphate catabolic process	513	31	17.800000	0.00142
GO:0042274	ribosomal small subunit biogenesis	25	5	0.870000	0.00144
GO:0007368	determination of left/right symmetry	81	9	2.810000	0.00182
GO:0043270	positive regulation of ion transport	16	4	0.560000	0.00184
GO:0031032	actomyosin structure organization	83	9	2.880000	0.00216
GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	100	10	3.470000	0.00229
GO:0034470	ncRNA processing	101	10	3.500000	0.00247
GO:0019083	viral transcription	101	10	3.500000	0.00247
GO:0003002	regionalization	195	15	6.760000	0.00298
GO:0002378	immunoglobulin biosynthetic process	10	3	0.350000	0.00412

(B)

GO.ID	Term	Annotated	Signif.	Expected	Classic Fisher
GO:0006338	chromatin remodeling	78	13	5.83000	0.00470
GO:0006333	chromatin assembly or disassembly	85	14	6.35000	0.00380
GO:0051384	response to glucocorticoid stimulus	42	8	3.14000	0.01120
GO:0032320	positive regulation of Ras GTPase activity	28	6	2.09000	0.01540
GO:0043954	cellular component maintenance	14	4	1.05000	0.01680
GO:0070301	cellular response to hydrogen peroxide	14	4	1.05000	0.01680
GO:0042059	negative regulation of epidermal growth	37	7	2.76000	0.01790
GO:0045862	positive regulation of proteolysis	37	7	2.76000	0.01790
GO:0021915	neural tube development	83	12	6.20000	0.01940
GO:0009161	ribonucleoside monophosphate metabolic process	16	4	1.20000	0.02720
GO:0033333	fin development	16	4	1.20000	0.02720
GO:0032967	positive regulation of collagen biosynthetic process	10	3	0.75000	0.03340
GO:0035329	hippo signaling cascade	10	3	0.75000	0.03340
GO:0050686	negative regulation of mRNA processing	10	3	0.75000	0.03340
GO:0002064	epithelial cell development	17	4	1.27000	0.03350
GO:0032092	positive regulation of protein binding	17	4	1.27000	0.03350
GO:0060271	cilium morphogenesis	92	12	6.87000	0.03990
GO:0070192	chromosome organization involved in meiosis	18	4	1.34000	0.04060
GO:0045732	positive regulation of protein catabolic process	35	6	2.61000	0.04270
GO:0001736	establishment of planar polarity	61	8	4.56000	0.08210

Appendix B: Supplementary materials for Chapter III.

Supplementary Material S1. List of differentially expressed transcripts with annotation found in digestive gland tissue showing an expression change greater than 100-fold (|logFC| > 2) in the microarray analysis.

UPREGULATED

Description	logFC
nose resistant to fluoxetine protein 6	8.95
autocrine motility factor receptor	8.80
c1q domain containing protein 1q3	8.77
hypothetical protein CGI_10026086 [Crassostrea gigas]	7.86
cathepsin d	7.55
nacre protein	6.74
bcl2 adenovirus e1b 19-kd protein-interacting	6.60
proteasome beta 4 subunit	6.58
mytimacin-5 partial	6.42
endo-1,3-beta-xylanase	6.38
c1q domain containing protein 1q25	6.36
cathepsin b	5.83
mantle gene 8	5.32
collagen alpha-5 chain	5.11
interferon-inducible gtpase 5-like	5.10
acetylcholinesterase	4.96
hypothetical protein CGI_10012557 [Crassostrea gigas]	4.72
deleted in malignant brain tumors 1	4.64
atpase h+ transporting lysosomal 21 kda v0 subunit	4.50
superoxide dismutase	4.42

vitelline membrane outer layer protein 1 homolog	4.38
col protein	4.37
uncharacterized protein loc102449188	4.24
hypothetical protein CGI_10015342 [Crassostrea gigas]	4.24
proteasome subunit beta type-4	4.09
developmentally-regulated vdg3	3.98
cell adhesion molecule-related down-regulated by oncogenes-like	3.92
vdg3 [Mytilus edulis]	3.87
neurogenic locus notch homolog protein 2-like	3.85
transitional endoplasmic reticulum atpase	3.85
peptidylglycine alpha-amidating monooxygenase precursor	3.79
collagen alpha-4 chain	3.78
fibrinogen c domain-containing protein 1	3.76
c-binding protein	3.75
fibrinogen-like protein a	3.71
ankyrin repeat domain-containing protein 50	3.68
mytimycin precursor	3.67
gtpase imap family member 7-like	3.65
hypothetical protein CGI_10011963 [Crassostrea gigas]	3.58
cytoplasmic partial	3.54
uncharacterized protein loc101862413	3.53
kazal-like serine protease inhibitor domain-containing protein	3.53
type-2 ice-structuring	3.51
nidogen and egf-like domain-containing protein 1	3.49
hypothetical protein [Paramecium tetraurelia strain d4-2]	3.49
gtpase imap family member 4-like	3.48
hypothetical protein [Acinetobacter sp. ANC 3789]	3.42

peptidylglycine alpha-amidating monooxygenase	3.35
sec1 family domain-containing protein 2	3.35
fimbrial protein pilin	3.34
hypothetical protein CGI_10003274 [Crassostrea gigas]	3.34
polyubiquitin	3.31
synaptosomal-associated protein 25	3.25
proline iminopeptidase	3.19
heavy metal-binding protein hip	3.17
uncharacterized protein loc101850813	3.14
jagged protein	3.13
apical endosomal glyco	3.11
low-density lipoprotein receptor-related protein 6	3.02
nc domain-containing protein	3.02
component of the counting factor complex	3.00
hypothetical protein CGI_10020658 [Crassostrea gigas]	2.99
nfx1-type zinc finger-containing protein 1	2.99
apoptosis inhibitor iap	2.99
mammalian ependymin-related protein 1	2.98
si:ch211- protein	2.94
hypothetical protein BRAFLDRAFT_106560 [Branchiostoma floridae]	2.91
nattectin precursor	2.89
hypothetical protein	2.88
hypothetical protein CGI_10012644 [Crassostrea gigas]	2.86
endo-1,3-beta-d-glucanase	2.85
peroxisomal proliferator-activated receptor a-interacting complex 285 kda	2.83
fibrinogen-related protein	2.83
rab gdp dissociation inhibitor beta	2.79

nadh dehydrogenase subunit 6	2.79
pancreatic secretory granule membrane major glycoprotein gp2	2.78
peptidyl-glycine alpha-amidating monooxygenase-like	2.70
oncoprotein-induced transcript 3 protein	2.66
heat shock 70 kda protein 12b	2.66
daz interacting protein zinc finger-like	2.65
complement c1q tumor necrosis factor-related protein 3	2.60
neuronal nitric oxidse synthase protein	2.59
uncharacterized protein loc580197	2.59
hypothetical protein BRAFLDRAFT_79532 [Branchiostoma floridae]	2.59
predicted protein [Nematostella vectensis]	2.58
60 kda ss-a ro ribonucleoprotein	2.53
propionyl- carboxylase alpha mitochondrial	2.53
succinate dehydrogenase	2.51
hypothetical protein DAPPUDRAFT_255671 [Daphnia pulex]	2.50
elongation factor 2	2.47
sialic acid binding lectin	2.47
syringomycin biosynthesis enzyme	2.42
protocadherin fat 4-like	2.41
phenylalanine hydroxylase	2.40
PREDICTED: neuroglian-like [Acyrthosiphon pisum]	2.38
hypothetical protein CGI_10013901 [Crassostrea gigas]	2.36
low affinity immunoglobulin epsilon fc receptor	2.35
vitelline membrane outer layer protein 1	2.34
group xvi phospholipase a2	2.32
gtpase imap family member 4	2.28
hypothetical protein CGI_10014841 [Crassostrea gigas]	2.26

gtpase imap family member 8-like	2.26
uncharacterized protein loc585517 isoform 2	2.25
translational elongation factor-2	2.19
virion core protein (lumpy skin disease virus)	2.19
tartrate-resistant acid phosphatase type 5	2.17
complement c1q-like protein 4	2.12
endoglucanase [Mizuhopecten yessoensis]	2.11
interferon alpha-inducible protein 27-like protein 2-like	2.09
complement c1q tumor necrosis factor-related protein 8	2.09
microtubule-associated protein futsch	2.08
Titin [Crassostrea gigas]	2.05
ankyrin unc44	2.05
jagged 1-like	2.04
3-hydroxyisobutyrate mitochondrial-like	2.03
hypothetical protein CGI_10008425 [Crassostrea gigas]	2.01

DOWNREGULATED

Description	**logFC**
sodium-dependent neutral amino acid transporter b at2-like	-2.01
alpha-tubulin	-2.04
ribonucleoside-diphosphate reductase subunit m2-like	-2.05
ribonucleotide reductase m2 polypeptide	-2.06
alpha-1,3-mannosyl-glycoprotein 4-beta-n-acetylglucosaminyltransferase b	-2.10
kex2p	-2.13
gastrointestinal growth factor xp4-like	-2.14
vitelline envelope zona pellucida domain protein 14	-2.26
kif21a protein	-2.29

reticulon-like protein	-2.30
phosphatidylinositol-binding clathrin assembly	-2.31
low quality protein: solute carrier organic anion transporter family member 5a1	-2.39
zona pellucida domain-containing protein 1	-2.41
conserved protein	-2.47
cytokine induced apoptosis inhibitor 1	-2.62
hypothetical protein LOTGIDRAFT_155380 [Lottia gigantea]	-2.71
hypothetical protein [Vibrio anguillarum]	-2.78
schlafen family member 13	-2.88
hypothetical protein CGI_10008221 [Crassostrea gigas]	-3.04
microsomal glutathione s-transferase 3-like protein	-3.21
dna-directed rna polymerase	-3.46
starch-binding domain-containing protein 1	-3.48
hypothetical protein [Butyricimonas synergistica]	-3.77
peroxisomal acyl-coenzyme a oxidase 2	-4.18
probable small nuclear ribonucleoprotein sm d2-like	-5.05
achain reduced peptidylglycine alpha-hydroxylating monooxygenase	-5.76

Supplementary Material S2. List of differentially expressed transcripts with annotation found in gill tissue showing an expression change greater than 100-fold (|logFC| > 2) in the microarray analysis.

UPREGULATED

Description	logFC
actin type-1 partial [Ostrea edulis]	4.52
mytimacin- partial	3.47
hypothetical protein CGI_10003274 [Crassostrea gigas]	2.47
bcl2 adenovirus e1b 19-kd protein-interacting	2.43
hypothetical protein CGI_10026086 [Crassostrea gigas]	2.27
vitelline envelope zona pellucida domain 9	2.09
c-type lectin domain family 4 member g	2.03

DOWNREGULATED

Description	logFC
rna-binding protein	-2.08
probable small nuclear ribonucleoprotein sm d2-like	-2.25
histone 3-like	-2.25
apextrin-like protein	-2.72
kif21a protein	-2.87
vitelline membrane outer layer protein 1 homolog	-3.59

Supplementary Material S3. List of active metabolic pathways associated with upregulated enzymes in digestive gland identified by microarray analysis and mapped to KEGG database.

Pathway
Biosynthesis of antibiotics
Purine metabolism
Cysteine and methionine metabolism
Glutathione metabolism
Fatty acid degradation
Methane metabolism
Arginine and proline metabolism
Glycolysis / Gluconeogenesis
Drug metabolism - cytochrome P450
Drug metabolism - other enzymes
beta-Alanine metabolism
Metabolism of xenobiotics by cytochrome P450
alpha-Linolenic acid metabolism
Chloroalkane and chloroalkene degradation
Glycine, serine and threonine metabolism
One carbon pool by folate
Oxidative phosphorylation
Pyruvate metabolism
Valine, leucine and isoleucine degradation
Starch and sucrose metabolism
Tyrosine metabolism
Amino sugar and nucleotide sugar metabolism
Carbon fixation pathways in prokaryotes

Drug metabolism - cytochrome P451

Metabolism of xenobiotics by cytochrome P451

Drug metabolism - cytochrome P452

Glycerophospholipid metabolism

Monobactam biosynthesis

Mucin type O-Glycan biosynthesis

Nicotinate and nicotinamide metabolism

Pantothenate and CoA biosynthesis

Phenylalanine, tyrosine and tryptophan biosynthesis

Sulfur metabolism

Tetracycline biosynthesis

Arginine biosynthesis

Biosynthesis of ansamycins

C5-Branched dibasic acid metabolism

Caffeine metabolism

Dioxin degradation

Folate biosynthesis

Glycosaminoglycan degradation

Glycosphingolipid biosynthesis - globo series

Inositol phosphate metabolism

Isoquinoline alkaloid biosynthesis

N-Glycan biosynthesis

Nitrogen metabolism

Other glycan degradation

Phosphatidylinositol signaling system

Primary bile acid biosynthesis

Riboflavin metabolism

Steroid degradation
Synthesis and degradation of ketone bodies
Taurine and hypotaurine metabolism
Thiamine metabolism
Various types of N-glycan biosynthesis
Xylene degradation

VITA

MARIA VICTORIA SUAREZ ULLOA

Born, A Coruña, Spain

1998-2007	B.A., (Licenciatura) Chemistry University of A Coruña A Coruña, Spain
2005-2007	M.S., Fundamental and Environmental Chemistry University of A Coruña A Coruña, Spain
2008-2010	M.S., Bioinformatics International University of Andalucia Málaga and Sevilla, Spain
2010-2013	Research Assistant University of A Coruña A Coruña, Spain
2013-2017	Doctoral Candidate Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

2017 Victoria Suarez-Ulloa, Vanessa Aguiar-Pulido, Giri Narasimhan, Jose M. Eirin-Lopez. Transcriptional patterns of chromatin-associated genes in response to environmental stress revealed by network analysis. National Shellfisheries Association Meeting, Knoxville, TN (US). Type of contribution: Oral communication.

2017 Gonzalez-Romero, R.*, Suarez-Ulloa, V.*, Rodriguez-Casariego, J., Garcia-Souto, D., Diaz, G., Smith, A., Pasantes, J. J., Rand, G., Eirin-Lopez J.M., Effects of Florida Red Tides on histone variant expression and DNA methylation in the Eastern oyster *Crassostrea virginica*. *Aquatic Toxicology* 7;186:196-204. (*Equal contribution)

2016 Victoria Suarez-Ulloa, Vanessa Aguiar-Pulido, Giri Narasimhan, Jose M. Eirín-Lopez. Dynamics of the expression of chromatin-associated genes in response to environmental stress using network analysis. Intelligent Systems for Molecular Biology (ISMB). Orlando, FL (US). Type of contribution: Poster.

2016 Aguiar-Pulido, V., Suarez-Ulloa, V., Huang, W., Cickovski, T., Mathee, K., Narasimhan, G., Metagenomics, Metatranscriptomics and Metabolomics Approaches for Microbiome Analysis, *Evolutionary Bioinformatics*. 12(Suppl. 1): 5-16.

2015 Suarez-Ulloa, V., Fernandez-Tajes, J., Aguiar-Pulido, V., Prego-Faraldo, M.V., Florez-Barros, F., Sexto-Iglesias, A., Mendez, J., Eirin-Lopez, J.M., Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid. *PeerJ*. 3:e1429.

2015 Suarez-Ulloa, V., Gonzalez-Romero, R., Eirin-Lopez, J.M. Environmental epigenetics: A promising venue for developing next-generation pollution biomonitoring tools in marine invertebrates. *Marine Pollution Bulletin*. 98(1-2):5-13.

2015 Victoria Suarez-Ulloa, Jose M. Eirin-Lopez. Framing epigenetic signatures of the Pacific oyster under environmental stress using network analysis. Asilomar Chromatin Chromosomes and Epigenetics Conference, Pacific Grove, CA (US). Type of contribution: Oral communication.

2015 Victoria Suarez-Ulloa, Juan Fernandez Tajes, Vanessa Aguiar-Pulido, Veronica Prego-Faraldo, Jose M. Eirin-Lopez. Multiple transcriptomic approach for the analysis of the molecular response of mussels to the effects of the biotoxin okadaic acid. Ecological & Evolutionary Genomics Gordon Research Conference. Biddeford, ME (US). Type of contribution: Poster.

2015 Aguiar-Pulido, V.*, Suarez-Ulloa, V.*, Eirin-Lopez, J.M., Pereira, J., Narasimhan, G. Computational methods in epigenetics. Personalized Epigenetics. Elsevier. ISBN: 9780124201354. (*****Equal contribution)

2014 Victoria Suarez-Ulloa, Jose M. Eirin-Lopez. Environmental epigenetics in bivalves: applications for oceans biomonitoring. Asilomar Chromatin and Chromosomes Conference, Pacific Grove, CA (US). Type of contribution: Oral communication.

2013 Suarez-Ulloa, V., Fernandez-Tajes, J., Manfrin, C., Gerdol, M., Venier, P. and Eirin-Lopez, J.M. Bivalve omics: state of the art and potential applications for the biomonitoring of harmful marine compounds. *Marine Drugs*. 11(11):4370-89.

2013 Aguiar-Pulido, V.*, Suarez-Ulloa, V.*, Rivero, D., Eirin-Lopez, J.M., and Dorado, J. Clustering of gene expression profiles applied to marine research. *IWANN 2013, Part I, Lecture Notes on Computer Science (LNCS)* 7902: 453–462. (*****Equal contribution)

2013 Suarez-Ulloa, V., Fernandez-Tajes, J., Aguiar-Pulido, V., Rivera-Casas, C., Gonzalez-Romero, R., Ausio, J., Mendez, J., Dorado, J. and Eirin-Lopez, J.M. The CHROMEVALOA database: a resource for the evaluation of okadaic acid contamination in the marine environment based on the chromatin-associated transcriptome of the mussel *Mytilus galloprovincialis*. *Marine Drugs* 11, 830-841.

2013 Vanessa Aguiar-Pulido*, Victoria Suarez-Ulloa*, Daniel Rivero, Jose M. Eirin-Lopez and Julian Dorado. Clustering of gene expression profiles applied to marine research. International Work Conference in Artificial Neural Networks (IWANN), Tenerife (Spain). Type of contribution: Oral communication (*Equal contribution).