A peer-reviewed version of this preprint was published in PeerJ on 19 November 2015.

View the peer-reviewed version (peerj.com/articles/1429), which is the preferred citable publication unless you specifically need to cite this preprint.

https://doi.org/10.7717/peerj.1429
Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid

Victoria Suarez-Ulloa, Juan Fernandez-Tajes, Vanessa Aguiar-Pulido, M. Veronica Prego-Faraldo, Fernanda Florez-Barros, Alexia Sexto-Iglesias, Josefina Mendez, Jose M. Eirin-Lopez

**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 main responsible 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). Modifications in the expression of environmentally relevant transcripts were initially assessed using qPCR. 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 expression 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 OA biomarkers.
Unbiased high-throughput characterization of mussel transcriptomic responses to sublethal concentrations of the biotoxin okadaic acid

Victoria Suarez-Ulloa¹, Juan Fernandez-Tajes², Vanessa Aguiar-Pulido³, M. Veronica Preg-ofalado¹,⁴, Fernanda Florez-Barros⁵, Alexia Sexto-Iglesias⁴, Josefina Mendez⁴, Jose M. Eirin-Lopez¹,*

¹ Chromatin Structure and Evolution Group (CHROMEVOL), Department of Biological Sciences, Florida International University, North Miami, Florida 33181 USA.
² McCarthy Group, Wellcome Trust Center for Human Genetics, University of Oxford, OX3 7BN Oxford, UK.
³ Bioinformatics Research Group (BioRG), School of Computing & Information Sciences, Florida International University, Miami, Florida 33199 USA.
⁴ XENOMAR Group, Department of Cellular and Molecular Biology, University of A Coruña, A Coruña 15071, Spain.
⁵ Centre for Nephrology, Royal Free Hospital, University College London, WC1E 6BT London, UK.

*Address for correspondence: Florida International University, Marine Sciences Program, 3000 NE 151st St., suite MSB-360, North Miami, FL 33181 USA. Email: jeirinlo@fiu.edu; Phone: 305-919-4000.

Key words: Harmful Algal Blooms; Diarrhetic Shellfish Poisoning; Okadaic Acid;
Transcriptomics; Microarray; Biomonitoring; Marine Pollution
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 main responsible 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). Modifications in the expression of environmentally relevant transcripts were initially assessed using qPCR. 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 expression 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 OA biomarkers.
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 main responsible 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 transform the extraordinary amount of molecular data resulting from omic experiments into sensible and rapid biomarkers of marine pollution.

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 (de Jesus Romero-Geraldo and Yolanda Hernandez-Saavedra 2014, Romero-Geraldo et al. 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 biomarkers of
marine pollution during DSP blooms. 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. Sampled individuals 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 $1 \times 10^{-6}$ (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).

*Gene expression analysis using quantitative PCR*

Four biological replicates per condition (exposed and control) were defined for gene expression analyses using quantitative PCR (qPCR), with two technical replicates analyzed per biological replicate. cDNA was synthesized from 1 µg of total RNA from digestive gland, using First Strand cDNA Synthesis kit following the manufacturer’s instructions (Roche Diagnostics, Indianapolis, IN). Specific PCR primers were designed according to qPCR restrictions based on DNA sequences of target genes defined in Supplementary Materials S1 using the program Primer3 (Rozen and Skaletsky 2000). Primer specificity was verified through PCR reactions, yielding a single DNA band with the expected length after electrophoresis in agarose gels. qPCR amplifications were subsequently carried out using the FastStart Essential DNA Green Master kit on a LightCycler 96 instrument (Roche Diagnostics, Indianapolis, IN) following the manufacturer’s instructions with the following modifications: 20 µl final volume of master mix containing 6.4 µl H$_2$O, 0.8 µl of each primer (10 µM), 10 µl of the SYBR Green Mix (kit) and 2 µl of each reverse transcribed RNA. The specificity of the obtained qPCR products was analyzed by melting curve analysis. Relative expression levels were normalized using RPS4 and GAPDH genes as references. Quantitation cycle (Cq) values were obtained using the qPCR instrument.
software LightCycler Software 1.5.0 (Roche Diagnostics, Indianapolis, IN) and analyzed using qBasePlus software (Biogazelle), assuming 100% amplification efficiency.

**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 (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).

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

<table>
<thead>
<tr>
<th></th>
<th>Normalized Libraries</th>
<th>SSH Libraries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>exposed</td>
<td>control</td>
</tr>
<tr>
<td>Reads</td>
<td>514,276</td>
<td>404,901</td>
</tr>
<tr>
<td>Contigs</td>
<td>24,624</td>
<td>16,395</td>
</tr>
<tr>
<td>Annotated Contigs</td>
<td>10,617 (43%)</td>
<td>7,335 (45%)</td>
</tr>
</tbody>
</table>
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.

**Figure 1.** 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.

*Expression modifications in response to low OA concentrations*

Expression analyses using qPCR confirmed the expression trends (upregulation or downregulation) of five genes chosen among those showing the most dramatic fold changes in a
preliminary comparison of normalized read counts. These genes are associated with protein repair and cell cycle regulation, and thus they are likely to be mediators of the response to environmental stress according to the general model proposed by Anderson et al. (Anderson et al. 2015). More specifically, this set of genes comprises (Figure 2): Heat shock protein 70, associated with the repair or degradation of misfolded proteins under a variety of environmental conditions (Kiang and Tsokus 1998); Spectrin beta, skeletal protein associated with the endoplasmic reticulum and involved in the regulation of apoptosis (Bennett and Baines 2001); Ras-related protein rab-8 and Ral-a-like isoform X1, small GTPases regulating membrane transport and also involved in the regulation of cell growth and proliferation (Folsch 2005, Loubery and Coudrier 2008). Consistently, Parkin co-regulated gene protein, a suppressor of cell death, was found downregulated (Imai et al. 2003). These results confirm that the OA exposure conditions set throughout DSP HAB experimental simulations is enough to elicit genetic responses related to cell cycle regulation and apoptosis in the digestive gland of mussels. Altogether, these results suggest that early stage DSP blooms trigger genetic responses that can be detected using molecular approaches, supporting the potential of this model for biomonitoring applications.
Rab-8, Ras-related protein rab-8; Ral-X1, Ral-a-like isoform X1; Spectrin, Spectrin beta; PACRG, Parkin co-regulated gene protein. Each bar represents the CNRQ (four biological replicates per treatment) and the error bars represent the standard error. Differences are significant in all cases ($p$-value < 0.05).

**Microarray-based analysis of transcriptomic responses to OA**

Building on the expression differences across individual genes, the present work expanded the analysis of the mussel genetic responses to OA exposure using an omic approach. 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 3). 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 4). The detailed description of the transcripts displaying the highest 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.
Figure 3. V-plots showing gene expression differences in digestive gland (left) and gill (right) tissues, 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.

The present microarray analysis identified a set of transcripts displaying sharp expression differences between exposed and control treatments (Supplementary Materials S2 and S3), expanding the list of transcripts potentially involved in the response to OA (Manfrin et al. 2010, Suarez-Ulloa et al. 2013a). In doing so, this work sets a framework for the future validation of these sequences as potential OA pollution biomarkers. 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 S2 and S3) include heat shock proteins, proteases like cathepsins and several proteasomal subunits, commonly associated with an accumulation of misfolded or oxidized proteins observed under different types of environmental stress (Gotze et al. 2014).

Interestingly, our results corroborate previous analyses describing the responses of *Mytilus galloprovincialis* to OA stress (Manfrin et al. 2010), particularly the strong upregulation of cAMP-responsive element binding protein (CREB), vdg3 and elongation factor 2.

On the other hand, 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 across 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).

**Figure 4.** 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 ($R^2 \cong 0.6$).

In addition to transcripts previously linked to OA responses, our results found an upregulation of
two antimicrobial peptides (mytimacin and mytilin) as well as an antifungal peptide (mytimycin)
specific from mussels (Sonthi et al. 2011, Gerdol et al. 2012). Interestingly, mytimacin (partial)
was identified as one of the most upregulated transcripts in both gill and digestive gland. The
upregulation of a member of the mytilin protein family, mytilin B, was previously observed under conditions of environmental stress different from bacterial challenge such as heat shock and induced physical stress by shell filing (Mitta et al. 2000). Additionally, 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 C1q has not been found in invertebrates yet, a large number of proteins containing this domain have been described in the mussel Mytilus (Gerdol et al. 2011). 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.

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 encoding genes are involved. A comparison of the functional profile for the two tissues studied is shown in Figure 5. 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). Additionally, the overall top ten of most represented GO terms for both upregulated and downregulated sets of transcripts can be found in Supplementary Materials S4 and S5.
Figure 5. Graphical representation of the GO terms (general sub-categories in Biological Process ontology) most represented in transcripts differentially expressed for each mussel tissue. The length of the bars is proportional to the number of sequences annotated for each specific GO term.

Although absolute differences are evident between digestive gland and gill (Figure 3), no major functional differences were found when comparing both tissues (Figures 4 and 5). Nonetheless, such comparison may be hampered by sample size differences (e.g., subtle tissue-specific differences could remain undetected). 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). 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 expression 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 time-resolved expression signatures rather than qualitative assessment of individual biomarkers for biomonitoring purposes.

The present work was completed by investigating the metabolic pathways associated with those enzymes identified as differentially expressed under OA exposure conditions (Supplementary Materials S6). 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 (Astuaya et al. 2015). Indeed, our results show an upregulation in important chaperones (Hsp70) and proteases (cathepsins b and d) (Supplementary Materials S2 and S3) 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 (Supplementary Materials S4 and S5), 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 transcriptomic changes constituting a general environmental response as well as on the specific transcriptomic signature of DSP toxicity stress.
ACKNOWLEDGMENTS

The present work was supported by grants from the Biomolecular Sciences Institute (800005997) and start-up funds from the College of Arts and Sciences at Florida International University (J.M.E.-L.), and from the Spanish Ministry of Economy and Competitivity (AGL2012-30897, J.M.). V.S.-U. was supported by a Graduate Assistantship from the Department of Biological Sciences at FIU. V.A.-P. was supported by the College of Engineering and Computing at Florida International University, under the supervision of Dr. Giri Narasimhan.
LITERATURE CITED


