1 **Transcriptomic Evaluation of the American oyster, *Crassostrea virginica*, Deployed**
2 **During the Deepwater Horizon Oil Spill: Evidence of an Active Hydrocarbon Response**
3 **Pathway.**

4
5 Matthew J. Jenny^{a,*}, William C. Walton^b, Samantha L. Payton^a, John M. Powers^a, Robert H.
6 Findlay^a, Britton O'Shields^a, Kirsten Diggins^a, Mark Pinkerton^a, Danielle Porter^a, Daniel M.
7 Crane^a, Jeffrey Tapley^a, Charles Cunningham^c

8
9 From the ^aDepartment of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487,
10 ^bSchool of Fisheries, Aquaculture and Aquatic Sciences, Auburn University, Dauphin Island, AL
11 36528, and ^cDepartment of Biology, University of New Mexico, Albuquerque, NM 87131.

12
13 Running title: Evidence of an active hydrocarbon response pathway in *Crassostrea virginica*

14
15 * Address correspondence: Matthew J. Jenny, Department of Biological Sciences, Box 870344,
16 University of Alabama, Tuscaloosa, AL 35487, USA. Tel: 205-348-8225, Fax: 205-348-1786,
17 Email: mjjenny@ua.edu

18
19
20 **Keywords:**

21 Oyster, biomarker, transcriptomics, aryl hydrocarbon receptor pathway, stress response,
22 polycyclic aromatic hydrocarbons

Abstract

Estuarine organisms were impacted by the Deepwater Horizon oil spill which released ~5 million barrels of crude oil into the Gulf of Mexico in the spring and summer of 2010. *Crassostrea virginica*, the American oyster, is a keystone species in these coastal estuaries and is routinely used for environmental monitoring purposes. However, very little is known about their cellular and molecular responses to hydrocarbon exposure. In response to the spill, a monitoring program was initiated by deploying hatchery-reared oysters at three sites along the Alabama and Mississippi coast (Grand Bay, MS, Fort Morgan, AL, and Orange Beach, AL). Oysters were deployed for 2-month periods at five different time points from May 2010 to May 2011. Gill and digestive gland tissues were harvested for gene expression analysis and determination of aliphatic and polycyclic aromatic hydrocarbon (PAH) concentrations. To facilitate identification of stress response genes that may be involved in the hydrocarbon response, a nearly complete transcriptome was assembled using Roche 454 and Illumina high-throughput sequencing from RNA samples obtained from the gill and digestive gland tissues of deployed oysters. This effort resulted in the assembly and annotation of 27,227 transcripts comprised of a large assortment of stress response genes, including members of the aryl hydrocarbon receptor (AHR) pathway, Phase I and II biotransformation enzymes, antioxidant enzymes and xenobiotic transporters. From this assembly several potential biomarkers of hydrocarbon exposure were chosen for expression profiling, including the AHR, two cytochrome P450 1A genes (CYP1A-like 1 and CYP1A-like 2), Cu/Zn superoxide dismutase (CuZnSOD), glutathione S-transferase theta (GST theta) and multidrug resistance protein 3 (MRP3). Higher expression levels of GST theta and MRP3 were observed in gill tissues from all three sites during the summer to early fall 2010 deployments. Linear regression analysis indicated a statistically significant relationship between total PAH levels in digestive gland tissue samples with CYP1A-like 2, CuZnSOD, GST *theta* and MRP3 induction. These observations provide evidence of a potentially conserved AHR pathway in invertebrates and yield new insight into the development of novel biomarkers for use in environmental monitoring activities.

Introduction

As a result of the Deepwater Horizon (DH) oil spill, estuarine organisms were exposed to fresh, weathered and dispersed crude oil which is a complex mixture of aromatic and non-aromatic hydrocarbons that vary in length and complexity (Hunt 1995). Polycyclic aromatic hydrocarbons (PAHs) are a significant constituent of crude oil and their toxicity has been well characterized in vertebrate models. PAHs pose a significant risk to humans and other animals inducing tissue damage from oxidative stress, cancer, developmental defects and immune suppression (Hahn 1998; Nebert et al., 2000; Wassenberg and Giulio 2004; Wassenberg et al., 2005; Maertens et al., 2008). As much of this research has focused on vertebrates, very little is known about the molecular and physiological responses to sublethal oil exposures in benthic estuarine invertebrates.

The aryl hydrocarbon receptor (AHR) pathway is the major regulatory pathway in vertebrates responsible for mediating the molecular response to hydrocarbon exposure (Hahn et al., 1997; Boutros et al., 2004). AHR is a ligand-activated, inducible transcription factor that pairs with its dimerization partner, AHR nuclear translocator (ARNT), to regulate expression of several Phase I and II metabolic enzymes. The primary AHR target genes and classic biomarkers of hydrocarbon exposure include cytochrome P450 1A (CYP1A), a Phase I enzyme, as well as Phase II conjugating enzymes, such as glutathione S-transferases and UDP-glucuronosyltransferases (Nebert *et al.*, 2000; Boutros *et al.*, 2004). These enzymes are involved in the oxidative and conjugative biotransformation of hydrocarbons which facilitates their excretion. In vertebrates, activation of AHR and induction of its target genes, most notably CYP1A, serve as biomarkers of hydrocarbon exposure (Montie et al., 2008; Zanette et al., 2009; Esler et al., 2011). Although previous studies have identified AHRs in molluscan species (Butler et al., 2001; Liu et al., 2010) and other marine invertebrates (Goldstone et al., 2006; Reitzel et al., 2014), the evolutionarily conserved responsiveness and regulatory aspect of this pathway is not well characterized in these organisms. Furthermore, although the structure of P450

enzymes is evolutionarily conserved among animals (Nelson et al., 2013), the validity of a classic CYP1 biomarker gene in invertebrate model organisms remains a major gap in our knowledge.

Sequencing of the Pacific oyster (*Crassostrea gigas*) genome provided some unique insight into the expansion of multiple gene families involved in the cellular stress response (Zhang et al., 2012), particularly CYP, Families 1 thru 4 which play a major role in xenobiotic biotransformation. Furthermore, a recent high-throughput sequencing approach in the mangrove oyster, *Crassostrea brasiliana*, also demonstrated the presence of multiple CYPs within each of these four families, as well as other Phase II conjugating enzymes, xenobiotic transporters and genes involved in the response to oxidative stress (Luchmann et al., 2015). *Crassostrea virginica*, the American oyster, is a keystone species in estuarine systems and an ideal sentinel for monitoring the health and recovery of the Gulf of Mexico ecosystem. Further characterization of the oyster AHR pathway and validation of its regulation of hydrocarbon-responsive gene expression would serve to establish a much needed battery of invertebrate biomarker genes indicative of hydrocarbon exposure that could be useful in monitoring the health of these oyster populations.

Similar to recent studies that have had success in identifying *C. virginica* genes involved in osmoregulation and innate immunity (Eierman and Hare 2014; McDowell et al., 2014; Zhang et al., 2014), we adopted a next-generation sequencing approach to identify stress response genes that play a classic role in mediating the molecular response to PAH exposure, such as those regulated by the AHR pathway, as well as the master regulator of the oxidative stress response, nuclear factor (erythroid-derived 2)-like 2 (NRF2). An initial *C. virginica* transcriptome was assembled from both gill and digestive gland using reads generated from the Roche 454 platform. The final assembly was enhanced by the generation of additional Illumina short reads that were used to increase the overall transcriptome coverage. Here we present a summary of the results of the annotation of the *C. virginica* transcriptome, with a specific focus on xenobiotic

stress regulatory pathways and classic biomarkers of hydrocarbon exposure. Furthermore, we provide evidence of an active hydrocarbon response pathway based on gene expression data of several AHR-related biomarker genes in hatchery-reared oysters that were deployed along the Mississippi-Alabama coast during and after the Deepwater Horizon oil spill.

Methods and Materials

2.1 Oyster Field Deployments

Hatchery-reared oysters from the Auburn University Shellfish Laboratory on Dauphin Island, Alabama were transplanted at three sites along the Mississippi-Alabama coastline (Grand Bay, Fort Morgan, Orange Beach). A subset of pre-deployed oysters was measured to determine average shell length, width and height, as well as condition index prior to deployment of the remaining oysters. Oysters were deployed for two-month intervals at five different timepoints over a 14 month period spanning from May 2010 to July 2011, and the first deployment began two weeks after the start of the DH oil spill (Table 1). Upon collection of oysters after each two-month deployment, growth parameters (changes in length, width and height) and condition index were measured on a second set of post-deployment oysters (May 2010 deployment sample size of $n = 6$; remaining four deployments sample size $n = 12$). Subadult oysters (mean average shell height = 47 mm) were transplanted in mesh bags measuring 50 x 50 x 10 cm deep with 3.8 cm mesh ($n = 4$ bags per site; 15 oysters per bag). The bags were placed in anchored cages and floated in 1.0 m of water. After the two-month deployment, oysters were collected, randomly assigned to one of several analyses and immediately dissected upon return to the Auburn University Shellfish Laboratory (see Supplemental Table 1 for full details of experimental design and resulting tissue assignments). Local salinity and temperature readings were recorded at the time of collection (Table 2). Gill and digestive gland tissue samples were flash frozen and stored at -80 °C. Tissue samples were shipped to the University of Alabama (Tuscaloosa, AL) on dry ice for RNA isolations and hydrocarbon tissue analyses. All raw data

for the growth and condition index assessments are available in the Gulf of Mexico Research Initiative Information & Data Cooperative (DOI: 10.7266/N7CR5R80).

2.2 Determination of Growth and Condition Index in Deployed Oysters

After measuring shell height, length and width, oysters were weighed individually on a Metler Toledo AL204 balance to the nearest 0.001 g as a measure of whole wet weight. Oysters were then dissected and the tissues separated from the shells. Shells were allowed to air dry for 48 h (± 2 h) and then weighed as a measure of dry shell weight. Tissues were dried in a Fisher Scientific Isotemp oven at 80 °C for 48 h (± 2 h), and then weighed as a measure of dry tissue weight. Using Abbe and Albright's (2003) formula, condition index (CI) was calculated as: $[(\text{dry tissue weight}) / (\text{whole wet weight} - \text{dry shell weight})] * 100$. The change in growth parameters and condition index between pre-deployed oysters and post-deployed oysters was calculated as: $[(\text{post-deployment parameter of individual oysters} - \text{pre-deployment average oyster parameter}) / (\text{total number of deployment days})]$.

2.3 Library Preparation and Next-Generation Sequencing

Total RNA was extracted and purified from gill and digestive gland from each of six oysters, two randomly chosen from each of the May 2010 deployment sites, using the TRIzol® reagent (Life Technologies, Invitrogen) protocol followed by DNase treatment and additional cleanup using the RNeasy Mini columns (QIAGEN Inc, Valencia, CA). Total RNA concentration was determined with a Nanodrop 2000C (Thermo Scientific, Wilmington, DE) and equal amounts from each individual oyster were pooled together to create one sample from each tissue for library preparation. Quality of the total RNA was confirmed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) prior to library preparation.

For 454 pyrosequencing, mRNA was purified from total RNA using the MPG mRNA Purification Kit (PureBiotech, USA) and cDNA synthesized using the cDNA Synthesis System Kit with random primers (Roche, USA). A cDNA Rapid library was constructed with the GS FLX

Titanium Rapid Library Preparation Kit and emulsion-based PCR amplification of the library was performed using a GS FLX Titanium LVemPCR Lib-L Kit. Pyrosequencing was conducted by the Center for Evolutionary and Theoretical Immunology at the University of New Mexico using a GS FLX Titanium Sequencing Kit XLR70 in a GS FLX+ System. All reagents and protocols used were from Roche 454 Life Sciences (USA).

For Illumina sequencing, library preparation and sequencing was performed by the Genomics Service Laboratory at the HudsonAlpha Institute for Biotechnology using the HiSeq 2500 platform.

2.4 Transcriptomic Assembly

The raw 454 reads from the individual tissues were trimmed and assembled using SeqMan NGen 4.0 (DNASTAR Lasergene Genomics Suite) and the following parameters: minimum average low quality score of 30 (99.9% accuracy), minimum of a 25 bp overlap between assembled reads, minimum match percentage of 92%, and contigs required a minimum of 3 assembled reads with an overall final assembled length of >250 bp. To remove redundant 454 assembled contigs between the two tissues, the contigs from both tissues were further assembled using SeqMan NGen 4.0 with the following parameters: minimum of a 25 bp overlap between assembled reads, minimum match percentage of 92%, and non-redundant contigs required a minimum of 1 assembled contig from each tissue with an overall final assembled length of >250 bp.

Oyster FASTQ files generated from the Illumina sequencing were assembled *de novo* using the Trinity alignment software for paired-end reads (Grabherr et al., 2011). The final alignment with the following parameters, 350-bp minimum contig length and 75-bp minimum overlap length at 99% identity, was chosen after optimization of the alignment protocol using multiple variations of minimum contig length, minimum overlap length and percent identity. To remove redundant contigs between the individual 454 assemblies and the Trinity assembly, all three sets of sequences were further assembled using SeqMan NGen 4.0 (DNASTAR Lasergene

Genomics Suite) with the following parameters: minimum of a 25 bp overlap between assembled reads, minimum match percentage of 92%, and non-redundant contigs required a minimum of 1 assembled contig or singleton from two of the previous assemblies with an overall final assembled length of >250 bp.

2.5 Gene Annotation and Identification of Genes Related to the Hydrocarbon Stress Response

Annotation of assembled contigs and singletons longer than 250 bp were searched in the NCBI non-redundant (nr) protein database using BLASTX using a cut-off *E* value of 1×10^{-6} . Data was collected through an automated program that pulled information from NCBI BLASTX results. The assembled contigs and singletons were then run through the BLAST search at AmiGO with a maximum number of alignments set at 50 to retrieve associated Gene Ontology (GO) terms describing biological processes, molecular functions and cellular components (Carbon et al., 2009).

2.6 Gene Expression Assessment by Real-Time RT- PCR

To determine if any of the newly identified transcripts had any potential to serve as molecular biomarkers of hydrocarbon exposure in oysters, several genes were selected for expression profiling from the deployed oysters. The AHR, two Cyp1A-like isoforms and CuZnSOD were chosen because of their classic role as biomarkers of hydrocarbon exposure in vertebrates. About a dozen other glutathione S-transferase (GST) and multidrug resistance (MRP) genes were initially screened from a subset of samples for any evidence of significant changes in gene expression between the different deployment sites (data not shown) and from that screening two transcripts with strong homology to MRP3 and GST *theta* were chosen for a full assessment from all five deployments.

Total RNA was prepared from ~50 mg of digestive gland or gill tissue from post-deployment oysters using the TRIzol® reagent (Life Technologies, Invitrogen) protocol. Total RNA was prepared for real-time RT- PCR by DNase treatment with additional cleanup using the

224 RNeasy Mini columns (QIAGEN Inc, Valencia, CA). Total RNA concentration was determined
 225 with a Nanodrop 2000C (Thermo Scientific, Wilmington, DE), and the integrity of the RNA was
 226 verified by denaturing gel electrophoresis. Biological replicates were generated by pooling
 227 equal amounts of total RNA from four individual oysters to create a single biological replicate.
 228 Thus 24 oysters were used to generate six biological replicates from each site per time point
 229 unless otherwise noted (See Supplemental Table #1 for full details on oyster tissue samples).
 230 cDNA was synthesized from 1 µg total RNA using a mix of random hexamers and oligo-dT
 231 primers with the qScript™ cDNA Synthesis Kit (Quanta Biosciences, Gaithersburg, MD).
 232 Quantitative PCR was performed using the PerfeCTa® SYBR® Green Supermix for iQ™
 233 (Quanta Biosciences, Gaithersburg, MD) in a MyiQ2 Two-Color Real-Time PCR Detection
 234 system (Bio-Rad, Hercules, CA). The PCR conditions used were: 95°C for 3 min, 95°C for 15
 235 s/60°C for 30 sec (40 cycles). At the end of each PCR run, a melt curve analysis was
 236 performed to ensure that only a single product was amplified. Three technical replicates were
 237 used for each sample. For quantitation of gene expression for all oyster genes, relative
 238 expression was calculated according to following equation (Livak and Schmittgen 2001):
 239 $\text{relative mRNA expression} = 2^{-\Delta\Delta C_t}$; where $\Delta\Delta C_t = [C_{t(\text{target gene})} - C_{t(\text{actin})}]_{\text{treatment}} - [C_{t(\text{target transcript})} -$
 240 $C_{t(\text{actin})}]_{\text{control}}$. Actin values were consistent across all sites for gill (mean Ct = 17.6, 95% CI [17.5,
 241 17.7]) and digestive gland (mean Ct = 20.0, 95% CI [19.7, 20.3]) tissues. Changes in
 242 expression are reported as changes in fold induction by normalizing $\Delta\Delta C_t$ values to the May
 243 2011 samples for each geographic location. There were no samples available for May 2010
 244 from the Grand Bay site, so May 2011 was chosen as the time point for normalizing gene
 245 expression to provide consistency across all sites. Primer sequences are provided in
 246 Supplementary Table 2. All raw data from the real-time RT-PCR assessments are available in
 247 the Gulf of Mexico Research Initiative Information & Data Cooperative (DOI:
 248 10.7266/N7CR5R80).

2.7 Assessment of Polycyclic Aromatic and Aliphatic Hydrocarbon Levels in Digestive Gland Tissues from Deployed Oysters

PAH and aliphatic hydrocarbon (AHC) concentrations were determined using digestive gland tissues after homogenization and subsampling for molecular analysis. For most deployment sites and time points, samples were generated by pooling 50 mg of wet weight tissue from each of four individual oysters (the same individuals pooled for the real-time RT-PCR measurements). In those cases where only two oysters were available for pooling, 100 mg of wet weight tissue from each oyster was pooled together. Limitations in tissue samples from the May/June 2010 samples, at times, prevented generating 200 mg samples and in these cases pooled samples ranged from 83 and 170 mg wet weight. In five cases, May/June 2010 digestive gland hydrocarbon samples contained tissues from a single oyster (See Supplemental Table #1 for full details on oyster tissue samples).

Aliphatic and polycyclic aromatic hydrocarbons were extracted, purified, identified and quantified following the small sample procedure detailed in Fang and Findlay (1996). Briefly, tissues were extracted in a one-phase dichloromethane, methanol, phosphate buffer (50 mM, pH = 7.4) solution (2.5, 5.0 and 2.0 ml, respectively) in an ashed 15-ml test tube fitted with a PTFE screw cap. The extraction mixture was well mixed and allowed to stand overnight in darkness at 4°C. Lipids and lipophilic pollutants were partitioned from other biological components by adding dichloromethane and distilled-deionized water such that the final ratio of dichloromethane-methanol-buffer was 1:1:0.9. The separated phases were allowed to stand for 24 h. The organic phase was quantitatively recovered and filtered through a glass column filled with anhydrous Na₂SO₄. Solvents were removed to one drop under a stream of N₂ and the samples stored under fresh solvent at -70°C.

Lipophilic hydrocarbons were separated from biogenic lipids by silicic acid column chromatography. Samples were loaded onto silicic acid columns in a minimum volume of

hexane and the lipophilic pollutants selectively eluted from the column with 5 column volumes of hexane. Lipophilic pollutants were quantified and identified using gas chromatography – mass spectrometry. An Agilent 6890N gas chromatograph fitted with 30 m X 0.32 mm, DB-5 MS fused-silica capillary column (J & W Scientific, Folsom, CA) and an Agilent 5793 Inert mass selective detector operating in scan (identification) and selective ion (quantification) modes were used. Compounds were identified by co-elution with known standards and their ion impact spectra. Daily standard curves of known concentration standards were constructed for quantification. All raw data from the PAH And AHC assessments are available in the Gulf of Mexico Research Initiative Information & Data Cooperative (DOI: 10.7266/N7J1014X).

2.8 Statistical Analysis

All statistical analyses were performed with the Prism 5 software package (GraphPad Software Inc., San Diego, CA). Statistical analysis of the change in condition index and additional growth parameters was performed using a two-way ANOVA with Bonferroni post-test comparisons of the different timepoints between geographic locations. The real-time RT-PCR datasets were $\log_{10}$ transformed and checked for normality using the Kolmogorov-Smirnov test. After transformation most of the datasets were normally distributed and differences in relative fold induction was determined by a parametric one-way analysis of variance (ANOVA, $p\text{-value} \leq 0.05$) with Bonferroni multiple comparisons between individual time points. For those few datasets that were not normally distributed even after transformation, relative fold induction was evaluated using the non-parametric Kruskal-Wallis ANOVA ($p\text{-value} \leq 0.05$), with Dunn's correction for multiple comparisons. Statistical analysis of the percent change in total PAH and AHC levels in digestive gland tissues was performed using a two-way ANOVA with Bonferroni post-test comparisons of the different timepoints between geographic locations. To assess the relationship between the degree of gene induction and the total PAH levels, linear regression analyses were performed ($p\text{-value} \leq 0.05$) by comparing fold change in transcript level to the

Log₁₀ normalized total PAH levels in digestive gland tissues of the samples from all sites and timepoints. R-squared and p-values are included on the graphs, as well as the 95% confidence bands of the best-fit line.

Results

3.1 *De novo* Transcriptome Assembly

A total of 783,963 and 678,822 Roche 454 reads were generated from digestive gland and gill pooled total RNA, respectively, collected from multiple field-deployed oysters from the Fort Morgan and Orange Beach May 2010 sample sites to capture the greatest diversity of expressed stress response genes representing the eastern oyster transcriptome (summarized in Table 3). The assembly resulted in a total of 31,772 and 34,963 contigs for digestive gland and gill, respectively. A comparison of the individual 454 assemblies identified 14,743 redundant contigs between the two tissues. The remaining 29,756 unassembled contigs represented transcripts unique to either tissue, suggesting that the two tissues had an average shared coverage of 42-46%. Previous observations of gene expression for both *C. virginica* (Jenny *et al.*, 2002; Jenny *et al.*, 2007) and *C. gigas* (Zhang *et al.*, 2012) highlight the presence of tissue-specific gene expression signatures. The combined assembly also produced an N50 contig length of 1,577bp, an average increase of 13.6% compared to the N50 contig lengths of the individual tissues (Table 3).

To increase the coverage of the sequenced eastern oyster transcriptome, a total of 95,743,790 Illumina 100-bp paired-end reads were generated from pooling the gill and digestive gland total RNA samples used in the 454 sequencing. The Trinity assembler generated 64,120 “singleton” contigs which included all possible assembled splice variants from the same transcripts (Grabherr *et al.*, 2011). Identification and grouping of related splice variants into one component produced a total of 39,182 unique contigs with an N50 length of 1,766 bp (Table 3). A final reduction in redundancy was performed by assembling a total of 130,845 sequences

consisting of the individual 454 contig sets (31,772 and 34,953 from digestive gland and gill, respectively) with the 64,120 “singleton” contigs from the Trinity assembly. Of these combined sequences, 27,277 remained as singletons while the remaining 95,835 contigs were reduced down to 27,277 contigs with an N50 length of 2,000 bp, a 13.2% and 26.8% increase compared to the N50 lengths of the Trinity and combined 454 assemblies, respectively (Table 3). A publicly accessible database to allow researchers to access the assembled contigs is currently under construction (www.ecological_genomics.ua.edu). Files containing the raw reads have been submitted to the National Center for Biotechnology Information Short Read Archive [SRP076299].

3.2 Functional Annotation of the Combined Illumina and 454 Assembly

A total of 80.2% of the 27,277 non-redundant contigs from the combined Illumina and 454 assembly were putatively annotated based on strong matches (E value threshold of 1×10^{-6}) to the NCBI nr protein database. The removal of “hypothetical proteins” from the blast analysis results in a reduction to a total 62.6% putatively annotated contigs (Summary blast results provided in Supplementary Table 3). In comparison, 61.3% of the 35,010 unique singletons from the combined Illumina and 454 assembly were putatively annotated based on strong matches (E value threshold of 1×10^{-6}) to the NCBI nr protein database, with a reduction to 46.6% with the removal of “hypothetical proteins” (Summary blast results provided in Supplementary Table 4). The non-redundant contigs and “singletons” from the combined Illumina and 454 assembly were also subjected to Gene Ontology (GO) annotation. 4,623 GO terms matched to one or more contigs and/or singletons and a total of 527,500 GO assignments were obtained. All GO assignments fell into broad categories from the three major GO functional domains (Biological Processes, Cellular Components and Molecular Function) as presented in Figure 1. Catalytic activity (32.4%) and metabolism (17.1%) were the two most prominent GO terms associated with molecular function and biological processes, respectively (Figure 1; complete summary of the GO annotation is provided in Supplementary Table 5).

Annotation of the transcripts yielded a large assortment of stress response genes, including many associated with PAH toxicity, such as members of the AHR pathway, Phase I and II biotransformation enzymes, xenobiotic transporters and antioxidants. A large complement of cytochrome P450 (CYP), Family 1 thru 4 transcripts, 49 total contigs and 65 total singletons representing transcripts from all four CYP families, were identified from the final contig and singleton gene sets (Table 4). All four of these families of CYPs are known to play important roles in xenobiotic biotransformation, but CYP, Family 1 members (CYP1A) are considered to be standard molecular biomarkers for PAH exposure. The most abundant CYP family was CYP2, followed by CYP4, CYP3 and CYP1 respectively (Table 4). We also identified 5 contigs and 2 singletons with a strong match to flavin-containing monooxygenases, which share a similar activity profile with CYPs. Twenty two contigs and 16 singletons matching the major Phase II conjugating enzyme glutathione S-transferase (GST) gene family were also identified from the final assembly. Lastly, we identified 64 contigs and 69 singletons that matched to various classes within the ATP-binding cassette transporter or MRP gene family that is well known for the efflux transport of xenobiotics (Table 4).

A summary of some specific examples of the various CYPs, GSTs and MRPs identified in the annotation are provided in Table 5. In addition, some of the relevant antioxidant stress response genes, such as superoxide dismutase and catalase, are also provided in Table 5. Finally, we were also able to identify several transcription factors that are often considered to be master regulators of the stress response. This group includes the AHR and its binding partner ARNT, as well as the hypoxia-inducible factor 1 α (HIF-1 α), heat shock transcription factor 1 (HSF-1), nuclear factor erythroid 2-related factor 2 (NRF2) and the metal transcription factor 2 (MTF-2) (Table 5).

3.3 Changes in Growth and Condition Index of Oysters Deployed during the Deepwater Horizon Oil Spill

To monitor the impacts of the Deepwater Horizon oil spill, which began on 20 April 2010 and continued until the capping of the well on 15 July 2010, three sites along the MS-AL coastal region were selected for monitoring (May 2010 to May 2011) with hatchery-reared deployed oysters: Fort Morgan, AL, Grand Bay, MS, and Orange Beach, AL. Grand Bay was added as the third site in Aug 2010 because initial sediment PAH analysis indicated very low concentrations compared to the Fort Morgan and Orange Beach sites (data not shown). The Grand Bay oysters were the only group to experience a positive change in condition index during all five deployments (Figure 2A). In contrast, Fort Morgan and Orange Beach oysters had a statistically significant reduction in condition index during the May 2010 and May 2011 deployments, and Orange Beach oysters also had a reduction in condition index during the Aug 2010 deployment (Figure 2A).

Although the change in condition index in Fort Morgan and Orange Beach oysters increased during the Nov 2010 and Feb 2011 deployments, the Orange Beach oyster condition index was significantly lower than the Grand Bay oyster at both of those time points while the Fort Morgan oysters were only significantly lower than Grand Bay oysters during the Feb 2011 deployment. The changes in growth parameters (shell height, length and width) had similar trends to the change in condition index, particularly for the Aug 2010 and May 2011 deployments (Figure 2B-D).

3.4 Stress Response Gene Expression in Oysters Deployed during the Deepwater Horizon Oil Spill

Although AHR expression had a higher variation in Grand Bay oysters from the Aug 2010 deployment, there was not a strong difference in expression levels throughout the remaining deployments in either digestive gland or gill tissue (Figures 3 and 4). AHR expression in the digestive gland of Fort Morgan oysters were the highest during the May 2010 deployment and decreased with each subsequent deployment (Figure 3), whereas AHR expression in the gills was only moderately induced during the Nov 2010 deployment (Figure 4). Although gene

expression analysis was performed on all five timepoints collected from the Orange Beach deployment, subsequent observations called into question the integrity of the purified total RNA from the Aug 2010 digestive gland samples. The total RNA degraded very quickly and the real-time PCR results for the housekeeping and candidate target genes suggested the presence of some inhibitory compounds in the samples that could potentially be interfering with the cDNA synthesis or the PCR reactions. Unfortunately, since these oyster tissue samples were subdivided for different analyses, such as the PAH and THC analysis presented in this paper and the earlier report on isotopic and trace element analyses by Carmichael et al (2012), there was not enough remaining tissue samples to attempt a second purification using alternative methods. It should be noted that there were no issues with the integrity of the total RNA purified from gill tissues. Therefore, the digestive gland samples from this particular timepoint were not included in the presented data. No significant changes in AHR expression in digestive gland were observed between the remaining four deployments (Figure 3). However, there was a slight, yet statistically significant difference in AHR expression in the gill of Orange Beach oysters between the Aug 2010 and May 2011 deployments (Figure 4).

CYP1A-like expression in digestive gland tissue was highest in Grand Bay and Orange Beach oysters during the Nov 2010 deployment with comparable expression profiles observed in gill tissue (Figures 3 and 4). Fort Morgan oysters displayed similar expression profiles, but had the highest levels of induction for both tissues in the Nov 2010 and Feb 2011 deployments (Figures 3 and 4). Since CYP enzymes are capable of generating reactive oxygen species, co-induction of SOD is another common biomarker associated with CYP1A-related hydrocarbon exposure. Modest changes in CuZnSOD expression were observed in the digestive gland from Grand Bay and Fort Morgan oysters at all of the timepoints (Figure 5). However, CuZnSOD expression was significantly induced in the digestive gland of Orange Beach oysters during the Feb 2011 deployment (Figure 5). Although there were some statistically significant differences in CuZnSOD expression were observed in the gill tissue of oysters deployed at all three sites,

the overall changes in induction were quite modest (1.5 to 2-fold) between the different timepoints (Figure 6).

MRP3 expression in digestive gland was higher in Fort Morgan and Orange Beach oysters during the latter deployments compared to May 2010 expression levels, with variable expression in Grand Bay oysters between the last four deployments (Figure 5). However, the overall changes in expression levels never exceeded more than a 2-fold difference at any of the sites during any of the deployments. In contrast, all three sites had the most significant increase in expression of MRP3 in gill tissues only during the Aug 2010 deployment (Figure 6). The highest levels of GST *theta* expression in the digestive gland were observed in oysters during the last three deployments compared to May 2010 at all three sites, with the exception of Orange Beach oysters which were only significantly higher during the Nov 2010 and Feb 2011 deployments (Figure 5). There were no significant changes in GST *theta* expression in the gill tissue of Grand Bay oysters from any of the deployments. In contrast, there was a statistically significant increase in GST *theta* expression in gill tissue from Fort Morgan and Orange Beach oysters during the Aug 2010 deployment, followed by a general trend of decreasing expression levels at the later deployments with one exceptional increase in expression in Orange Beach oysters during the Feb 2011 deployment (Figure 6).

3.5 Polycyclic Aromatic Hydrocarbons and Total Hydrocarbons in Oysters Deployed during the Deepwater Horizon Oil Spill

A total of 16 PAHs and 17 AHCs (C20 to C36) were identified from digestive gland tissue samples from oysters from all three sites and five deployments. There were no statistically significant differences in total PAH levels from any of the sites in oysters collected from the May 2010, Aug 2010, Nov 2010 and May 2011 deployments (Figure 7A). However, there was a strong general trend toward higher average total PAH levels in Orange Beach oysters during the May 2010 and Aug 2010 deployments. The highest total PAH levels were detected in Fort Morgan and Orange Beach oysters from the Feb 2011 deployment (Figure 7A). Of the 16 PAHs

analyzed, dibenzo(a,h)anthracene, 2-bromophthalene, acenophthalene and phenanthrene were not detected in any tissue samples. Acenaphthylene was only detected in two tissue samples in Fort Morgan oysters from the Aug 2010 deployment (Supplementary Table 6). Two other PAHs, indeno(1,2,3-cd)pyrene and benzo(g,h,i)perylene were only detected in Fort Morgan and Orange Beach oysters from the Aug 2010 deployment. However, only 1-2 pooled oyster samples from Fort Morgan had detectable levels of these two PAHs compared to the 5 out of 6 tissue samples from Orange Beach with detectable levels (Supplementary Table 6). Details on the average tissue levels of the remaining 9 PAHs is provided in Supplementary Table 6. There were no significant differences in AHC concentrations in the digestive gland tissue of oysters from Grand Bay and Fort Morgan for any of the deployments (Figure 7B). However, the Orange Beach oysters did have significantly higher AHC concentrations from the May 2010, Nov 2010 and Feb 2011 deployments compared to the other two sites, while the AHC concentrations from the Aug 2010 and May 2011 deployments were within the same lower range as the Grand Bay and Fort Morgan oysters at these timepoints (Figure 7B).

3.6 Linear Regression Analysis between Candidate Gene Expression and Polycyclic Aromatic Hydrocarbon levels

Based on visual observations, there did not appear to be a strong correlation between consistent gene expression profiles and PAH levels. To further investigate any potential relationship between the degree of gene induction and the total PAH levels, linear regression analyses were performed ($p\text{-value} \leq 0.05$) by comparing the Log_{10} normalized total PAH levels in digestive gland tissues to gene expression fold changes in digestive gland and gill. Due to limited financial resources, we were unable to measure PAH levels in gill tissues, but the digestive gland levels are assumed to be indicative of general levels of hydrocarbon exposure. Three of the six genes (CYP1A-like 2, SOD and GST *theta*) in both digestive gland and gill tissues demonstrated a statistically significant relationship between fold induction and higher PAH tissue levels (Figure 8). CYP1A-like 1 was not statistically significant in either tissue, while

AHR was only statistically significant in gill tissue and MRP3 was on statistically significant in digestive gland tissue (Figure 8).

Discussion

4.1 The *C. virginica* Transcriptome

The 454 assembly results for gill and digestive gland tissues (Table 3) presented in this study are consistent with a similar *C. virginica* 454 sequencing project which identified 28,939 contigs from a normalized cDNA library generated from RNA combined from multiple soft tissues (Eierman and Hare 2014). Similar comparisons can also be made between the Trinity assembly results presented here (39,182 contigs, minimum length >350 bp) with a similar *C. virginica* assembly that yielded 61,147 and 25,419 contigs with minimum length of >100 and >500 bp, respectively (Zhang *et al.*, 2014). The combination of 454 and Illumina sequencing appears to have provided a more effective assembly of the *C. virginica* transcriptome based on the significant increase in the N50 value (Table 3), as seen with a similar approach to sequencing the catfish genome (Jiang *et al.*, 2011). Furthermore, the high percentage of putatively annotated contigs (62.6%) from our combined 454 and Trinity assembly has a higher success rate compared to similar transcriptomic assemblies that resulted in the putative annotation of 47.5% of *C. brasiliensis* contigs (E value threshold of 1×10^{-6}) from 454 sequencing, and 23.1% (E value threshold of 1×10^{-5}) or 19.8% (E value threshold of 1×10^{-6}) of *C. virginica* contigs from Illumina sequencing (McDowell *et al.*, 2014; Zhang *et al.*, 2014; Luchmann *et al.*, 2015).

The transcriptomic data presented here, as well as in previous *C. virginica* studies (Eierman and Hare 2014; McDowell *et al.*, 2014; Zhang *et al.*, 2014), provide strong evidence for the same adaptive expansion of multiple gene families involved in the cellular response to both biotic and abiotic stressors as demonstrated by the sequencing and annotation of the genome from the Pacific oyster, *C. gigas* (Zhang *et al.*, 2012). The cytochrome P450 superfamily

represents an excellent example of the adaptive expansion within the *Crassostrea* family. Humans have a total of 57 CYP genes compared to the 136 genes identified from the *C. gigas* genome annotation (Zhang *et al.*, 2012). From our *C. virginica* transcriptome annotation, we found 49 contigs and 65 singletons that match to members of the CYP, Family 1 thru 4 genes that are known for their role in Phase I metabolism. CYP1 and CYP2 enzymes are well known for their role in xenobiotic biotransformation in vertebrates, while the CYP3 and CYP4 enzymes play a role in the metabolism of many pharmaceuticals. In addition to the diversity of Phase I enzymes, we also identified a large number of GST enzymes (Table 4), as well as some sulfotransferases (Supplementary Tables 2 and 3), that are well known for their roles as Phase II conjugative enzymes. Several studies have documented the utility of GSTs as biomarkers for pollutant exposure in multiple bivalve species (Vidal and Narbonne 2000; Park *et al.*, 2009; Vidal-Linan *et al.*, 2010; Wu *et al.*, 2013; Vidal-Linan *et al.*, 2014). Interestingly, we did not identify any UDP-glucuronosyltransferases, often considered to be the most important Phase II enzymes in many vertebrates. After oxidative and conjugative transformation of xenobiotics by the Phase I and II enzymes, transport of these modified chemicals across the plasma membrane is mediated by multidrug resistance (MRP) efflux proteins. While P-glycoprotein, also known as MRP1, has been partially characterized in *C. virginica* (Keppler and Ringwood 2001; Keppler and Ringwood 2001; Ivanina and Sokolova 2008), many of the other MRP proteins identified here likely play important roles in xenobiotic detoxification and have the potential to be excellent biomarker genes (Luckenbach and Epel 2008; Luchmann *et al.*, 2015).

The identification of several of the ‘master regulators’ of the stress response (Table 5) provides an opportunity to further expand the functional characterization of these pathways. Previous studies have partially characterized the HIF-1 α transcription factor in *C. virginica* (Piontkivska *et al.*, 2011), as well as HIF-1 α and HSF-1 in *C. gigas* (Kawabe and Yokoyama 2011; Kawabe and Yokoyama 2012). Although AHRs have been identified in molluscan species (Butler *et al.*, 2001; Liu *et al.*, 2010), there have been no studies on the expression or

characterization of ARNT, also known as HIF-1 β , that is required to complete the transcriptionally active heterodimeric complex with either AHR or HIF-1 α . Surprisingly, we did not find an MTF-1 isoform from our transcriptome annotation but did find a copy of the less widely studied MTF-2 isoform. Both MTF-1 and MTF-2 have been partially characterized in *C. gigas*, and their roles as ‘master regulators’ of metals homeostasis and metals detoxification appear to be conserved (Qui et al., 2013; Meng et al., 2015). Finally, there have been no bivalve studies on the identification or function of NRF2, the transcription factor which plays a master regulatory role in the response to oxidative stress. Thus, only a handful of studies exist on the characterization of these vital transcription factors in bivalve molluscs compared to the thousands of published papers characterizing their role in the vertebrate stress response. The assembly and annotation of the *C. virginica* genome, the sequencing of which is currently underway (Gomez-Chiarri et al., 2015), will provide important genomic resources necessary to support the future functional characterization of these transcription factors and investigate their evolutionarily conserved roles in the stress response.

4.2 Changes in Condition Index and Growth

Although the observed significant reduction in condition index in oysters deployed at the Orange Beach site during the May and August 2010 deployments may at first suggest that the oysters were experiencing significant stress due to the oil exposure (Figure 2), the significant reductions observed from both the Fort Morgan and Orange Beach deployments in May 2011 suggests that this may in fact be associated with seasonal abiotic stressors such as changes in temperature or salinity resulting in differences in physiological demand. The condition index of oysters can display moderate reductions during the spawning period due to the energy demands associated with gametogenesis and spawning, similar to the differences observed between the spring/summer periods and the fall/winter periods at the Grand Bay and Fort Morgan sites (Wilson et al., 2005; Gullian and Aguirre-Macedo, 2009). It is also not unusual for the condition index of oysters to display site-to-site variation at the same time of year, likely as a

result local differences in physic-chemical parameters (Volety, 2008). Furthermore, a higher condition index has been shown to occur at more moderate (~19 °C) temperatures compared to higher (~25 °C) temperatures (Clark et al., 2013), which may partially explain the higher condition indices during the Nov 2010 and Feb 2011 deployments for all three sites (See Table 2 for water temperatures at post-deployment collection dates). Thus, there is no clear indication as to whether the changes in condition index were related to hydrocarbon exposure or simply a result of season and site-specific variation.

4.3 Gene Expression Profiles of Deployed Oysters

The sites chosen for the oyster deployments were most heavily impacted by the Deepwater Horizon oil spill from mid-June 2010 to early August 2010, during which time significant amounts of crude oil washed up on beaches and into estuaries along the Alabama coast (Hayworth et al., 2011). However, stable isotope ratios and trace element analysis performed by collaborators on this project (Carmichael et al., 2012) were unable to confirm direct oil exposure in the oysters deployed during the early deployments (May 2010 to Oct 2010). Although these analyses suggest that the deployed oysters did not assimilate oil-derived carbon or nitrogen, this does not indicate a lack of oil exposure during the deployments and could suggest that the oysters did not consume enough oil-derived material to be detected (Carmichael et al., 2012). Since the summer of 2010, multiple studies have documented the persistence of crude oil constituents along the AL coast. For example, PAH levels remained easily detectable in sediment and sand samples taken from Orange Beach, AL in the fall of 2011 (Hayworth et al., 2011). Furthermore, tarballs and surface residual balls originating from the Deepwater Horizon oil spill have continued to wash up along beaches as late as 2014, especially after storm events, and are likely to persist for several years (Mulabagal et al., 2013; Hayworth et al., 2015). Thus, our deployment strategy provided an opportunity to monitor coastal habitats and further characterize the utility of potential molecular biomarkers for future biomonitoring activities.

Surprisingly, there was an observed difference in PAH levels between the periods when oysters were most likely to experience direct oiling from the spill (May and Aug 2010) and the highest levels of accumulated PAHs during the Feb 2011 deployments at Fort Morgan and Orange Beach (Figure 7). The higher total PAH levels observed in the Feb 2011 deployment (Figure 7) may be a consequence of PAH bioaccumulation combined with low elimination rates due to colder winter temperatures. There is also the possibility that oysters were also exposed to hydrocarbons from residual oil washing ashore later in the winter or from sources other than the Deepwater Horizon oil spill. Interestingly, there were no significant differences in AHC levels in oysters from any of the sites during the Aug 2010 deployment (Figure 7). This is in direct contrast to the significantly higher levels of AHC observed in Orange Beach oysters during the May 2010, Nov 2010 and Feb 2011 deployments. The lower levels of AHC in Orange Beach Aug 2010 oysters during this time raises some questions about the potential metabolic processes in oysters during these later summer months.

Bivalve species are frequently used in environmental monitoring programs, and although activation of the 'AHR gene battery' is well documented in vertebrates there is still a significant lack of dependable biomarkers indicative of hydrocarbon exposure in bivalves despite research efforts (Livingstone and Farrar 1984; Livingstone et al., 1989; Peters et al., 1998; Marigomez et al., 1999; Livingstone et al., 2000; Sole and Livingstone 2005). Previous studies have identified an AHR in two bivalve molluscs, *Mya arenaria* and *Ruditapes philippinarum*. Although the *M. arenaria* AHR was shown to not bind the potent AHR agonist TCDD (Butler et al., 2001), the *R. philippinarum* AHR was inducible by benzo[a]pyrene (Liu et al., 2010). In a more recent study utilizing the mangrove oyster, *Crassostrea brasiliana*, both Cyp1A and CuZnSOD transcript abundance were enriched in tissues from oysters exposed to either diesel fuel water-accommodated fraction (WAF) or phenanthrene (Luchmann et al., 2015). These previous studies support the results of our linear regression analyses in which increases in CYP1A-like 2 and CuZnSOD fold change in both digestive gland and gill tissue, and AHR in gill tissue,

positively correlated with increased PAH tissue levels (Figure 8). The increase in both Cyp1A and CuZnSOD is usually considered a strong indication of hydrocarbon exposure because previous studies have demonstrated the potential of Cyp1A to generate reactive oxygen species that can contribute to the toxicity associated with exposure to hydrocarbons or other Cyp1A substrates (Schlezingner et al., 1999; Schlezingner et al., 2006). As a result of the oxidative stress associated with hydrocarbon exposure, there is significant overlap between the AHR and NRF2 signaling pathways. For instance, NRF2 is a known transcriptional regulator of CuZnSOD, as well as multiple GST and MRP genes (Zhu et al., 2005; Maher et al., 2007). Luchmann et al (2015) also demonstrated an enrichment of several GST genes and MRP3 transcripts in oysters exposed to phenanthrene, which is consistent with our observation of a significant correlation between PAH levels and expression of MRP3 in digestive gland and GST *theta* in both tissues (Figure 8). These observations suggest a potential for tissue-specific responses to hydrocarbon exposure thru the AHR and/or NRF2 pathway, but there are still significant questions about the potential effectiveness of any of these genes as functional molecular biomarkers of hydrocarbon exposure in an invertebrate model. For example, while MRP3 fold change in digestive gland correlated with PAH levels, there was no significant correlation in gill tissues even though we observed significant induction of MRP3 in gill tissues during the Aug 2010 deployment at all three sites (Figure 6). Future studies such as controlled laboratory exposures with weathered crude oil and more definitive experiments, such as ligand-binding assays and functional promoter characterizations are required to definitively assess the transcriptional regulation of these genes in response to hydrocarbon exposure.

Conclusions

We have described a *de novo* transcriptome for *C. virginica* that is highly enriched in stress response genes related to xenobiotic exposure and detoxification. These data will significantly contribute to the development of novel genomic resources that are currently underway for this

economically and ecologically important species. Furthermore, we have presented strong evidence of a hydrocarbon response in oysters that may be related to classic signaling through the AHR and NRF2 pathways. These observations provide new insight into mechanisms that are important to understanding the impact of hydrocarbon toxicity on marine invertebrates. Future studies focusing on molecular signaling mechanisms and the endogenous functions of these proteins will provide an evolutionary context to these important stress response pathways.

Acknowledgements

This study was supported by a Gulf of Mexico Research Initiative Year-One Block Grant to the Alabama Marine Environmental Science Consortium (MJJ, WCW and RHF) and in part by the National Science Foundation under award number 1057152 (MJJ). Research reported in this publication was supported in part by the National Institute of General Medical Sciences of the National Institutes of Health under award number P30 GM110907 (CC). The authors would like to thank Dr. Ruth Carmichael (Dauphin Island Sea Lab) and Dr. Anne Boettcher (University of South Alabama) for their assistance with the oyster tissue dissections and sample shipping. In addition, the authors would like to thank the staff of the Molecular Biology Core Facility of the Center for Evolutionary and Theoretical Immunology, Department of Biology, University of New Mexico who performed the Roche 454 sequencing. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Figure Legends

Figure 1. Gene ontology annotation of the *C. virginica* transcriptome. Functional annotation of the 27,277 contigs from gill and digestive gland generated from the combined assembly of the individual 454 and Illumina assemblies. Each pie chart represents the distribution of the GO annotations into major functional categories for each division. Categories represented by less

than 1% of the annotations were combined into the “others” category for convenience. A full description of all GO categories can be found in Supplementary Table #5.

Figure 2. Change per day in condition index (A), and additional growth parameters including change in shell height (B), change in shell length (C), and change in shell width (D) from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. Equal letters indicate an absence of statistical difference between the deployment sites based on a 2-Way ANOVA (Bonferroni post-test comparisons, p -value < 0.05).

Figure 3. Relative fold change of *C. virginica* AHR, CYP1A-like 1 and CYP1A-like 2 in digestive gland tissues from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. Fold change was determined by normalizing all samples within a single geographic location to the May 2011 expression levels. The graphs are plotted as box and whiskers with a 5 to 95 percentile range. Equal letters indicate an absence of statistical difference between the groups based on 1-Way ANOVA (Bonferroni multiple comparisons, p -value < 0.05). Datasets that were not normally distributed even after transformation (designated with a KW in upper left corner of graph), relative fold induction was evaluated using the non-parametric Kruskal-Wallis ANOVA (Dunn’s correction for multiple comparisons, p -value < 0.05). $n = 6$ biological replicates.

Figure 4. Relative fold change of *C. virginica* AHR, CYP1A-like 1 and CYP1A-like 2 in gill tissues from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. Fold change was determined by normalizing all samples within a single geographic location to the May 2011 expression levels. The graphs are plotted as box and whiskers with a 5 to 95 percentile range. Equal letters indicate an absence of statistical difference between the groups based on 1-Way ANOVA (Bonferroni multiple comparisons, p -value < 0.05). Datasets that were not normally distributed even after transformation (designated with a KW in upper left corner of graph), relative fold induction was evaluated using the non-parametric Kruskal-Wallis ANOVA (Dunn’s correction for multiple comparisons, p -value < 0.05). $n = 6$ biological replicates.

Figure 5. Relative fold change of *C. virginica* SOD, MRP3 and GST *theta* in digestive gland tissues from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. Fold change was determined by normalizing all samples within a single geographic location to the May 2011 expression levels. The graphs are plotted as box and whiskers with a 5 to 95 percentile range. Equal letters indicate an absence of statistical difference between the groups

based on 1-Way ANOVA (Bonferroni multiple comparisons, p -value < 0.05). Datasets that were not normally distributed even after transformation (designated with a KW in upper left corner of graph), relative fold induction was evaluated using the non-parametric Kruskal-Wallis ANOVA (Dunn's correction for multiple comparisons, p -value < 0.05). $n = 6$ biological replicates.

Figure 6. Relative fold change of *C. virginica* SOD, MRP3 and GST *theta* in gill tissues from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. Fold change was determined by normalizing all samples within a single geographic location to the May 2011 expression levels. The graphs are plotted as box and whiskers with a 5 to 95 percentile range. Equal letters indicate an absence of statistical difference between the groups based on 1-Way ANOVA (Bonferroni multiple comparisons, p -value < 0.05). Datasets that were not normally distributed even after transformation (designated with a KW in upper left corner of graph), relative fold induction was evaluated using the non-parametric Kruskal-Wallis ANOVA (Dunn's correction for multiple comparisons, p -value < 0.05). $n = 6$ biological replicates.

Figure 7. Total PAH and AHC concentrations in *C. virginica* digestive gland tissue from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. Equal letters indicate an absence of statistical difference between the deployment sites based on 2-Way ANOVA (Bonferroni post-test comparisons, p -value < 0.05). In most case the sample size consists of 6 biological replicates generated from pooling tissue from four separate oysters.

Figure 8. Linear regression analysis between total PAH and changes in relative fold induction in *C. virginica* digestive gland and gill tissue from oysters deployed along the MS-AL coast during the Deepwater Horizon Oil Spill. The analysis was performed on Log_{10} normalized total PAH levels in digestive gland tissues of the samples from all sites and timepoints. 95% confidence bands (dashed lines) of the best-fit line (solid line) are included in the plots.

References

- Abbe, G.R., Alright, B.W., 2003. An improvement to the determination of meat condition index for the Eastern oyster, *Crassostrea virginica* (Gmelin, 1791). Journal of Shellfish Research 22, 747-752.
- Boutros, P.C., Moffat, I.D., Franc, M.A., Tijet, N., Tuomisto, J., Pohjanvirta, R., Okey, A.B., 2004. Dixon-responsive AHRE-II gene battery: identification by phylogenetic footprinting. Biochemical and Biophysical Research Communication 321, 707-715.

731 Butler, R.A., Kelley, M.L., Powell, W.H., Hahn, M.E., Beneden, R.J.V., 2001. An aryl
732 hydrocarbon receptor (AHR) homologue from the soft-shell clam, *Mya arenaria*:
733 evidence that invertebrate AHR homologues lack 2,3,7,8-tetrachlorodibenzo-p-dioxin
734 and beta-naphthoflavone binding. *Gene* 278, 223-234.

735 Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S., AmiGO Hub, Web
736 Presence Working Group, 2009. AmiGO: online access to ontology and annotation data.
737 *Bioinformatics* 25, 288-289.

738 Carmichael, R.H., Jones, A.L., Patterson, H.K., Walton, W.C., Perez-Huerta, A., Overton, E.B.,
739 Dailey, M., Willett, K.L., 2012. Assimilation of oil-derived elements by oysters due to the
740 Deepwater Horizon Oil Spill. *Environmental Science & Technology* 46, 12787-12795.

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

745 Eierman, L.E., Hare, M.P., 2014. Transcriptomic analysis of candidate osmoregulatory genes in
746 the Eastern oyster *Crassostrea virginica*. *BMC Genomics* 15, 503.

747 Esler, D., Ballachey, B.E., Trust, K.A., Iverson, S.A., Reed, J.A., Miles, A.K., Henderson, J.D.,
748 Woodin, B.R., Stegeman, J.J., McAdie, M., Mulcahy, D.M., Wilson, B.W., 2011.
749 Cytochrome P4501A1 biomarker indication of the timeline of chronic exposure of
750 Barrow's goldeneyes to residual Exxon Valdez oil. *Marine Pollution Bulletin* 62, 609-614.

751 Fang, J., Findlay, R.H., 1996. The use of a classic lipid extraction method for simultaneous
752 recovery of organic pollutants and microbial lipids from sediments. *Journal of*
753 *Microbiological Methods* 27, 63-71.

754 Goldstone, J.V., Hamdoun, A., Cole, B.J., Howard-Ashby, M., Nebert, D.W., Scally, M., Dean,
755 M., Epel, D., Hahn, M.E., Stegeman, J.J., 2006. The chemical defenseome: environmental
756 sensing and response genes in the *Strongylocentrotus purpuratus* genome.
757 *Developmental Biology* 300, 366-384.

758 Gomez-Chiarri, M., Warren, W.C., Guo, X., Proestou, D., 2015. Developing tools for the study of
759 molluscan immunity: the sequencing of the genome of the eastern oyster, *Crassostrea*
760 *virginica*. *Fish & Shellfish Immunology* 46, 2-4.

761 Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X.,
762 Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A.,
763 Rhind, N., Palma, F.d., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman,

764 N.,Regev, A., 2011. Full-length transcriptome assembly from RNA-seq without a
765 reference genome. *Nature Biotechnology* 29, 644-652.

766 Gullian M. and Aguirre-Macedo, L., 2009. Seasonal variation of physiological parameters in the
767 Eastern oyster *Crassostrea virginica* from a tropical region of the Gulf of Mexico. *Journal*
768 *of Shellfish Research* 28, 439-446.

769 Hahn, M.E., 1998. The aryl hydrocarbon receptor: a comparative prespective. *Comparative*
770 *Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology* 121,
771 23-53.

772 Hahn, M.E., Karchner, S.I., Shapiro, M.A.,Perera, S.A., 1997. Molecular evolution of two
773 vertebrate aryl hydrocarbon (dioxin) receptors (AHR1 and AHR1) and the PAS family.
774 *Proceedings of the National Academy of Sciences* 94, 13743-13748.

775 Hayworth, J.S., Clement, T.P., John, G.F.,Yin, F., 2015. Fate of Deepwater Horizon oil in
776 Alabama's beach system: Understanding physical evolution processes based on
777 observational data. *Marine Pollution Bulletin* 90, 95-105.

778 Hayworth, J.S., Clement, T.P.,Valentine, J.F., 2011. Deppwater Horizon oil spill impacts on
779 Alabama beaches. *Hydrology and Earth System Sciences* 15, 3639-3649.

780 Hunt, J.M. (1995). Petroleum geochemistry and geology. New York,, W. H. Freeman and
781 Company.

782 Ivanina, A.V.,Sokolova, I.M., 2008. Effects of cadmium exposure on expression and activity of
783 P-glycoprotein in eastern oysters, *Crassostrea virginica* Gmelin. *Aquatic Toxicology* 88,
784 19-28.

785 Jenny, M.J., Chapman, R.W., Mancia, A., Chen, Y.A., McKillen, D.J., Trent, H., Lang, P.,
786 Escoubas, J.M., Bachere, E., Boulo, V., Liu, Z.J., Gross, P.S., Cunningham, C., Cupit,
787 P.M., Tanguy, A., Guo, X., Moraga, D., Boutet, I., Huvet, A., De Guise, S., Almeida, J.S.,
788 Warr, G.W., 2007. A cDNA microarray for *Crassostrea virginica* and *C. gigas*. *Marine*
789 *Biotechnology* 9. 577-591.

790 Jenny, M.J., Ringwood, A.H., Lacy, E.R., Lewitus, A.J., Kempton, J.W., Gross, P.S., Warr,
791 G.W., Chapman, R.W., 2002. Potential indicators of stress response identified by
792 expressed sequence tag analysis of hemocytes and embryos from the American oyster,
793 *Crassostrea virginica*. *Marine Biotechnology* 4, 81-93.

794 Jiang, Y., Lu, J., Peatman, E., Kucuktas, H., Liu, S., Wang, S., Sun, F.,Liu, Z., 2011. A pilot
795 study for channel catfish whole genome sequencing and *de novo* assembly. *BMC*
796 *Genomics* 12, 629.

797 Kawabe, S., Yokoyama, Y., 2011. Novel isoforms of heat shock transcription factor 1 are
 798 induced by hypoxia in the Pacific oyster *Crassostrea gigas*. Journal of Experimental
 799 Zoology, Part A, Ecological Genetics and Physiology 315, 394-407.
 800 Kawabe, S., Yokoyama, Y., 2012. Role of hypoxia-inducible factor a in response to hypoxia and
 801 heat shock in the Pacific oyster *Crassostrea gigas*. Marine Biotechnology 14, 106-119.
 802 Keppler, C., Ringwood, A.H., 2001. Expression of P-glycoprotein in the gills of oysters,
 803 *Crassostrea virginica*: seasonal and pollutant related effects. Aquatic Toxicology 54,
 804 195-204.
 805 Keppler, C.J., Ringwood, A.H., 2001. Expression of P-glycoprotein in southeastern oysters,
 806 *Crassostrea virginica*. Marine Environmental Research 52, 81-96.
 807 Liu, N., Pan, L., Miao, J., Xu, C., Zhang, L., 2010. Molecular cloning and sequence analysis and
 808 the response of a aryl hydrocarbon receptor homologue gene in the clam *Ruditapes*
 809 *philippinarum* exposed to benzo(a)pyrene. Comparative Biochemistry and Physiology
 810 Part C: Toxicology & Pharmacology 152, 279-287.
 811 Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time
 812 quantitative PCR and the 2⁻(Delta Delta C(T)) Method. Methods 25, 402-408.
 813 Livingstone, D.R., Chipman, J.K., Lowe, D.M., Minier, C., Mitchelmore, C.L., Moore, M.N.,
 814 Peters, L.D., Pipe, R.K., 2000. Development of biomarkers to detect the effects of
 815 organic pollution on aquatic invertebrates: recent molecular, genotoxic, cellular and
 816 immunological studies on the common mussel (*Mytilus edulis* L.) and other mytilids.
 817 International Journal of Environment & Pollution 13, 56-91.
 818 Livingstone, D.R., Farrar, S.V., 1984. Tissue and subcellular distribution of enzyme activities of
 819 mixed-function oxygenase and benzo(a)pyrene metabolism in the common mussel
 820 *Mytilus edulis* L. Science of the Total Environment 39, 209-235.
 821 Livingstone, D.R., Kirchin, M.A., Wiseman, A., 1989. Cytochrome P-450 and oxidative
 822 metabolism in molluscs. Xenobiotica 19, 1041-1062.
 823 Luchmann, K.H., Clark, M.S., Bainy, A.C.D., Gilbert, J.A., Craft, J.A., Chipman, J.K., Thorne,
 824 M.A.S., Mattos, J.J., Siebert, M.N., Schroeder, D.C., 2015. Key metabolic pathways
 825 involved in xenobiotic biotransformation and stress responses revealed by
 826 transcriptomics of the mangrove oyster *Crassostrea brasiliana*. Aquatic Toxicology 166,
 827 10-20.
 828 Luckenbach, T., Epel, D., 2008. ABCB- and ABCC-type transporters confer multixenobiotic
 829 resistance and form an environment-tissue barrier in bivalve gills. American Journal of
 830 Physiology. Regulatory, Integrative and Comparative Physiology 294, R1919-1929.

831 Maertens, R.M., Yang, X., Zhu, J., Gagne, R.W., Douglas, G.R., White, P.A., 2008. Polycyclic
832 aromatic hydrocarbon content and excess lifetime cancer risk from preschool exposure.
833 Environmental Science and Technology 42, 1747-1753.

834 Maher, J.M., Dieter, M.Z., Aleksunes, L.M., Slitt, A.L., Guo, G., Tanaka, Y., Scheffer, G.L.,
835 Chan, J.Y., Manautou, J.E., Chen, Y., Dalton, T.P., Yamamoto, M., Klaassem, C.D.,
836 2007. Oxidative and electrophilic stress induces multidrug resistance-associated protein
837 transporters via the nuclear factor-E2-related factor-2 transcriptional Pathway.
838 Hepatology 46, 1597-1610.

839 Marigomez, I., Lekube, X., Cancio, I., 1999. Immunochemical localisation of proliferating cells in
840 mussel digestive gland tissue. The Histochemical Journal 31, 781-788.

841 McDowell, I.C., Nikapitiya, C., Aguiar, D., Lane, C.E., Istrail, S., Gomez-Chiarri, M., 2014.
842 Transcriptome of American oysters, *Crassostrea virginica*, in response to bacterial
843 challenge: insights into potential mechanisms of disease resistance. PloS One 9,
844 e105097.

845 Meng, J., Zhang, L., Li, L., Li, C., Wang, T., Zhang, G., 2015. Transcription factor CgMTF-1
846 regulates CgZnT1 and CgMT expression Pacific oyster (*Crassostrea gigas*) under zinc
847 stress. Aquatic Toxicology 165, 179-188.

848 Montie, E.W., Fair, P.A., Bossart, G.D., Mitchum, G.B., Houde, M., Muir, D.C., Letcher, R.J.,
849 McFee, W.E., Starczak, V.R., Stegeman, J.J., Hahn, M.E., 2008. Cytochrome P4501A1
850 expression, polychlorinated biphenyls and hydroxylated metabolites, and adipocyte size
851 of bottlenose dolphins from the Southeast United States. Aquatic Toxicology 86, 397-
852 412.

853 Mulabagal, V., Yin, F., John, G.F., Hayworth, J.S., Clement, T.P., 2013. Chemical fingerprinting
854 of petroleum biomarkers in Deepwater Horizon oil spill samples collected from Alabama
855 shoreline. Marine Pollution Bulletin 70, 147-154.

856 Nebert, D.W., Roe, A.L., Dieter, M.Z., Solis, W.A., Yang, Y., Dalton, T.P., 2000. Role of the
857 aromatic hydrocarbon receptor and [Ah] gene battery in the oxidative stress response,
858 cell cycle control, and apoptosis. Biochemical Pharmacology 59, 65-85.

859 Nelson, D.R., Goldstone, J.V., Stegeman, J.J., 2013. The cytochrome P450 genesis locus: the
860 origin and evolution of animal cytochrome P450s. Philosophical Transactions of the
861 Royal Society of London. Series B, Biological Sciences 368, 20120474.

862 Park, H., Ahn, I.Y., Kim, H., Lee, J., Shin, S.C., 2009. Glutathione S-transferase as a biomarker
863 in the Antarctic bivalve *Laternula elliptica* after exposure to the polychlorinated biphenyl
864 mixture Aroclor 1254. Comp Biochem Physiol C Toxicol Pharmacol 150, 528-536.

865 Peters, L.D., Nasci, C.,Livingstone, D.R., 1998. Immunochemical investigations of cytochrome
866 P450 forms/epitopes (CYP1A, 2B, 2E, 3A and 4A) in digestive gland of *Mytilus ssp.*
867 Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and
868 Endocrinology 121, 361-369.

869 Piontkivska, H., Chung, J.S., Ivanina, A.V., Sokolov, E.P., Techa, S.,Sokolova, I.M., 2011.
870 Molecular characterization and mRNA expression of two key enzyme of hypoxia-sensing
871 pathways in eastern oysters *Crassostrea virginica* (Gmelin): hypoxia-inducible factor a
872 (HIF-a) HIF-prolyl hydroxylase (PHD). Comparative Biochemistry and Physiology Part D:
873 Genomics & Proteomics 6, 103-114.

874 Qui, J., Liu, Y., Yu, M., Pang, Z., Chen, W.,Xu, Z., 2013. Identification and functional
875 characterization of MRE-binding transcription factor (MTF) in *Crassostrea gigas* and its
876 conserved role in metal-induced response. Molecular Biology Reports 40, 3321-3331.

877 Reitzel, A.M., Passamaneck, Y.J., Karchner, S.I., Franks, D.G., Martindale, M.Q., Tarrant,
878 A.M.,Hahn, M.E., 2014. Aryl hydrocarbon receptor (AHR) in the cnidarian *Nematostella*
879 *vectensis*: comparative expression, protein interactions, and ligand binding.
880 Development Genes and Evolution 224, 13-24.

881 Schlezinger, J.J., Struntz, W.D., Goldstone, J.V.,Stegeman, J.J., 2006. Uncoupling of
882 cytochrome P450 1A and stimulation of reactive oxygen species production by co-planar
883 polychlorinated biphenyl congeners. Aquatic Toxicology 77, 422-432.

884 Schlezinger, J.J., White, R.D.,Stegeman, J.J., 1999. Oxidative inactivation of cytochrome P-450
885 1A (CYP1A) stimulated by 3,3',4,4'-tetrachlorobiphenyl: production of reactive oxygen by
886 vertebrate CYP1As. Molecular Pharmacology 56, 588-597.

887 Sole, M.,Livingstone, D.R., 2005. Components of the cytochrome P450-dependent
888 monooxygenase system and 'NADPH-independent benzo[a]pyrene hydroxylase' activity
889 in a wide range of marine invertebrate species. Comparative Biochemistry and
890 Physiology Part C: Toxicology and Pharmacology 141.

891 Vidal-Linan, L., Bellas, J., Campillo, J.A.,Beiras, R., 2010. Integrated use of antioxidant
892 enzymes in mussels, *Mytilus galloprovincialis*, for monitoring pollution in highly
893 productive coastal areas of Galicia (NW Spain). Chemosphere 78, 265-272.

894 Vidal-Linan, L., Bellas, J., Etxebarria, N., Nieto, O.,Beiras, R., 2014. Glutathione S-transferase,
895 glutathione peroxidase and acetylcholinesterase activities in mussels transplanted to
896 harbour areas. Sci Total Environ 470-471, 107-116.

897 Vidal, M.L., Narbonne, J.F., 2000. Characterization of glutathione S-transferase activity in the
898 Asiatic clam *Corbicula fluminea*. Bulletin of Environmental Contamination & Toxicology
899 64, 455-462.

900 Volety, A.K., 2008. Effects of salinity, heavy metals and pesticides on health and physiology of
901 oysters in the Caloosahatchee Estuary, Florida. Ecotoxicology 17, 579-590.

902 Wassenberg, D.M., Giulio, R.T.D., 2004. Teratogenesis in *Fundulus heteroclitus* embryos
903 exposed to a creosote-contaminated sediment extract and CYP1A inhibitors. Marine
904 Environmental Research 58, 163-168.

905 Wassenberg, D.M., Nerlinger, A.L., Battle, L.P., Giulio, R.T.D., 2005. Effects of the polycyclic
906 aromatic hydrocarbon heterocycles, carbazole and dibenzothiophene, on *in vivo* and *in*
907 *vitro* CYP1A activity and polycyclic aromatic hydrocarbon-derived embryonic deformities.
908 Environmental Toxicology and Chemistry 24, 2526-2532.

909 Wilson, C., Scotto, L., Scarpa, J., Volety, A., Laramore, S., Haunert, D., 2005. Survey of water
910 quality, oyster reproduction and oyster health status in the St. Lucie Estuary. Journal of
911 Shellfish Research 24, 157-165.

912 Wu, X., Liu, W., Gou, X., 2013. Tissue-specific distribution and activity of glutathione S-
913 transferases as biomarkers in bivalve shellfishes. Water Environ Res 85, 404-410.

914 Zanette, J., Jenny, M.J., Goldstone, J.V., Woodin, B.R., Watka, L.A., Bainy, A.C., Stegeman,
915 J.J., 2009. New cytochrome P450 1B1, 1C2 and 1D1 genes in the killifish *Fundulus*
916 *heteroclitus*: basal expression and response of five killifish CYP1a to the AHR agonist
917 PCB126. Aquatic Toxicology 93, 234-243.

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

929 Zhang, L., Li, L., Zhu, Y., Zhang, G.,Guo, X., 2014. Transcriptome analysis reveals a rich gene
930 set related to innate immunity in the Eastern oyster (*Crassostrea virginica*). Marine
931 Biotechnology 16, 17-33.

932 Zhu, H., Itoh, K., Yamamoto, M., Zweier, J.L., Li, Y., 2005. Role of Nrf2 signaling in regulation of
933 antioxidants and phase 2 enzymes in cardiac fibroblasts: protection against reactive
934 oxygen and nitrogen species-induced cell injury. FEBS Letters 579, 3029-3036.

935

936

Figure 1

Percent (%) of Total Gene Ontology Annotation

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

Digestive Gland

Gill

Table 1. Hatchery-reared Oyster Deployment Dates

	Deployment Site		
Deployment #	Grand Bay	Fort Morgan	Orange Beach
GPS Coordinates	30° 22.993'N 88° 18.767'W	30° 14.394'N 87° 52.189'W	30° 18.535'N 87° 30.888'W
1	No Deployment	May 4, 2010 to June 30, 2010	May 5, 2010 to June 30, 2010
2	Aug 3, 2010 to Oct 6, 2010	Aug 3, 2010 to Oct 7, 2010	Aug 3, 2010 to Oct 7, 2010
3	Nov 2, 2010 to Jan 7, 2011	Nov 1, 2010 to Jan 6, 2010	Nov 1, 2010 to Jan 6, 2010
4	Feb 4, 2011 to Mar 31, 2011	Feb 14, 2011 to Apr 6, 2011	Feb 4, 2011 to Apr 6, 2011
5	May 4, 2011 to July 6, 2011	May 4, 2011 to July 7, 2011	May 5, 2011 to July 6, 2011

Table 2. Temperature and Salinity of Local Coastal Seawater at Time of Post-Deployment Oyster Collection.

	Grand Bay		Fort Morgan		Orange Beach	
Month of Initial Deployment	Temperature	Salinity	Temperature	Salinity	Temperature	Salinity
May 2010	27.8 °C	17.9 ‰	28.4 °C	14.5 ‰	29.5 °C	12.7 ‰
Aug 2010	18.5 °C	25.6 ‰	25.0 °C	17.0 ‰	21.0 °C	17.0 ‰
Nov 2010	10.5 °C	28.5 ‰	14.5 °C	25.5 ‰	13.0 °C	24.0 ‰
Feb 2011	25.5 °C	14.5 ‰	18.0 °C	7.5 ‰	20.5 °C	17.0 ‰
May 2011	31.0 °C	20.0 ‰	30.0 °C	17.5 ‰	31.5 °C	20.5 ‰

Table 3. Summary of Roche 454 and Illumina Next-Generation Sequencing Assembly

	Digestive Gland 454	Gill 454	454 Combined	Illumina	Illumina & 454
Total # of Reads	783,963	678,822	66,725 ^a	95,871,895	130,845 ^b
Singleton #	n/a	n/a	29,756	64,120 ^c	35,010 ^d
Contig #	31,772	34,953	14,743	39,182 ^c	27,277 ^d
Contig (N50 ^e)	1,400	1,376	1,577	1,766	2,000
Contig (mean bp)	848	947	1,329	1,429	1,674
Singleton (mean bp)	n/a	n/a	566	n/a	863

^aThe sum of the assembled contigs from Digestive Gland and Gill 454.

^bThe sum of the assembled contigs from Digestive Gland 454, Gill 454 and Illumina Singletons.

^cThe Illumina "Singleton #" includes potential splice variants and "Contig #" represents the reduced number after related spliced variants are grouped into one contig.

^dOf the 130,845 sequences compared in the this assembly, 95,835 condensed into the 27,277 contigs leaving 35,101 unique "singletons" from the three original assemblies.

^eThe N50 statistic is commonly used in computational biology and is the measure of the average length of a set of sequences.

Table 4. Summary of Pertinent Phase I, Phase II, Phase III Genes Involved in Biotransformation and Excretion of Xenobiotics

Gene Family	Identified from Contigs	Identified from Singletons
Cytochrome P450, Family 1	3	9
Cytochrome P450, Family 2	18	25
Cytochrome P450, Family 3	13	16
Cytochrome P450, Family 4	15	15
Flavin-containing monooxygenase	5	2
Glutathione S-transferases	22	16
Multidrug Resistance Transporters	64	69

Table 5. Summary of Major Stress Response Genes Identified by Next-Gen Sequencing

Sequence ID	Transcript Description	Species Match	Accession # Match	E-Value
Transcription Factors – Master Regulators of the Stress Response				
S-16101	Aryl Hydrocarbon Receptor (AHR)	<i>Crassostrea gigas</i>	EKC23432.1	0.0
C-6628	AHR nuclear translocator-like protein (ARNT)	<i>C. gigas</i>	EKC32806.1	0.0
C-1673	Hypoxia-inducible factor 1 α (HIF-1 α)	<i>C. gigas</i>	AED87588.1	0.0
C-388	Nuclear factor erythroid 2-related factor 2 (NRF2)	<i>C. gigas</i>	EKC19603.1	0.0
C-5703	Heat shock transcription factor 1 (HSF-1)	<i>C. gigas</i>	BAK61501.1	0.0
C-26099	Metal Response Transcription Factor 2 (MTF-2)	<i>C. gigas</i>	AGB56249.1	0.0
Cytochrome P450 Gene Family Members				
C-14418	Cytochrome P450 1A1	<i>C. gigas</i>	EKC38032.1	0.0
C-18472	Cytochrome P450 1A2	<i>C. gigas</i>	EKC40102.1	0.0
C-20744	Cytochrome P450 2B4	<i>C. gigas</i>	EKC35518.1	1e-49
C-6475	Cytochrome P450 2B5	<i>C. gigas</i>	EKC34198.1	1e-180
C-6432	Cytochrome P450 2C8	<i>C. gigas</i>	EKC28816.1	0.0
C-6228	Cytochrome P450 2C14	<i>C. gigas</i>	EKC26766.1	0.0
C-8604	Cytochrome P450 2D14	<i>C. gigas</i>	EKC18231.1	0.0
C-4711	Cytochrome P450 2F2	<i>C. gigas</i>	EKC23711.1	0.0
C-10486	Cytochrome P450 2G1	<i>C. gigas</i>	EKC24419.1	0.0
C-3890	Cytochrome P450 3A9	<i>C. gigas</i>	EKC23465.1	0.0
C-17777	Cytochrome P450 3A24	<i>C. gigas</i>	EKC24008.1	4e-165
C-3942	Cytochrome P450 4A25	<i>C. gigas</i>	EKC39293.1	0.0
C-5026	Cytochrome P450 4D2	<i>C. gigas</i>	EKC18482.1	3e-157
C-20266	Cytochrome P450 4F8	<i>C. gigas</i>	EKC34231.1	2e-82
Oxidative Stress Response Genes				
C-2345	Catalase	<i>C. hongkongensis</i>	ADZ93495.1	0.0
C-2052	Superoxide dismutase [Cu-Zn]	<i>C. gigas</i>	EKC41617.1	0.0
C-14853	Glutaredoxin-3	<i>C. gigas</i>	EKC37761.1	1e-170
C-12729	Glutathione reductase, mitochondrial	<i>C. gigas</i>	EKC35400.1	0.0
C-12976	Glutathione peroxidase 3	<i>C. gigas</i>	EKC20036.1	4e-93
Glutathione S-transferase Gene Family				
C-16939	Glutathione S-transferase 1	<i>C. gigas</i>	EKC23981.1	2e-159
C-13785	Glutathione S-transferase <i>omega</i> 1	<i>C. gigas</i>	EKC37935.1	1e-161
C-12750	Glutathione S-transferase <i>kappa</i> 1	<i>C. gigas</i>	EKC18414.1	2e-123
C-8207	Glutathione S-transferase <i>mu</i> class protein	<i>C. ariakensis</i>	ACJ06748.1	1e-138
C-14019	Glutathione S-transferase <i>pi</i> class protein	<i>C. gigas</i>	EKC32272.1	3e-122
C-13653	Glutathione S-transferase <i>theta</i> class protein	<i>C. gigas</i>	EKC30389.1	1e-155
ATP-Binding Cassette Transporter Gene Family				
C-3348	Multidrug resistance protein 1	<i>C. gigas</i>	EKC35579.1	0.0
C-17449	Multidrug resistance protein 3	<i>C. gigas</i>	EKC22340.1	0.0
C-8441	Multidrug resistance protein 4	<i>C. gigas</i>	EKC31450.1	0.0