Whole-Transcriptome Sequencing of Epidermal Mucus as a Novel Method for Oil Exposure Assessment in Juvenile Mahi-Mahi (Coryphaena hippurus)

Justin B. Greer,*† Nicolette E. Andrzejczyk,*‡ Edward M. Mager,* John D. Stieglitz,*† Daniel Benetti,*† Martin Grosell,** and Daniel Schlenk†§

1Department of Environmental Sciences, University of California, Riverside, California 92521, United States
2Department of Biological Sciences and Advanced Environmental Research Institute, University of North Texas, Denton, Texas 76201, United States
3Department of Marine Ecosystems and Society, University of Miami, Miami, Florida 33149, United States
4Department of Marine Biology and Ecology, University of Miami, Miami, Florida 33149, United States
5Institute of Environmental Health, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China

Supporting Information

ABSTRACT: Crude oil-derived polycyclic aromatic hydrocarbons (PAHs) are pervasive environmental pollutants with well-established detrimental effects on the health of marine organisms. Following large-scale oil spills in the marine environment, there is a critical need for noninvasive sampling methods to assess environmental exposure to PAHs without further perturbations to the population and for long-term monitoring following a spill. To test the efficacy of epidermal mucus mRNA as a source for noninvasive sampling, juvenile mahi-mahi (Coryphaena hippurus, ~28 days of age) were exposed to control seawater or two concentrations of high-energy water accommodated fractions (HEWAFs; 5% or 10%) of Deepwater Horizon surface oil for 48 h. Whole-transcriptome sequencing revealed differential expression of 501 transcripts in the low-HEWAF exposure (ΣPAH = 16.55 μg/L) and 196 transcripts in the high-HEWAF exposure (ΣPAH = 23.03 μg/L), suggesting differential regulation of mRNA in mucus following PAH exposure. In addition to differential expression of well-established biomarkers of PAH exposure such as cytochrome P450 enzymes, the mucosal transcriptome showed differential expression of transcripts involved in immune response, cardiotoxicity, and calcium homeostasis that parallel molecular responses in whole embryos. The consistency of the changes in expression in the epidermal mucus compared to that of tissues obtained from lethal sampling suggests that mucus is a promising source for noninvasive monitoring techniques.

INTRODUCTION

Large-scale oil spills such as the 2010 Deepwater Horizon (DWH) spill in the Gulf of Mexico can have significant ecological impacts on marine populations, for example, via exposure to polycyclic aromatic hydrocarbons (PAHs). The DWH oil spill resulted in extensive oiling of spawning regions for commercially important pelagic species in the Gulf of Mexico, including mahi-mahi (Coryphaena hippurus). Sublethal exposure to environmentally relevant concentrations of crude oil-derived PAHs has been shown to impair cardiac development, swimming performance, craniofacial development, and behavior, leading to reduced fitness and increased mortality risk.1–4 Therefore, there is a need for the development of monitoring tools that can estimate environmental exposure and monitor the health status of individuals following a spill in a noninvasive manner to minimize further population disruptions.

Fish epidermal mucus is a promising vector for the development of noninvasive sampling methods that can be repeatedly performed on the same individual without inducing physical harm associated with other methods such as biopsy or fin clipping.5 Epidermal mucus provides the first line of defense against pathogens and toxicants and is secreted by goblet cells, sacciform cells, and club cells located throughout the epidermis.6 Mucus contains a wide range of molecules such as proteins, lipids, carbohydrates, mRNA, and DNA, with the most well-characterized being antimicrobial and immune-related molecules.7,8 Furthermore, mucosal composition has been shown to be altered by a variety of stress conditions. For
example, oil exposure in dusky splitfin (*Goeda gracilis*) elicited antioxidant responses in the epidermal mucus that were greater than those found in the liver, brain, and muscle. Other studies focused on aquaculture applications have demonstrated that immune-related proteins such as lectins, heat shock proteins, and complement factors are highly abundant and dynamically regulated following bacterial infection, food deprivation, and overcrowding stress. Thus, changes in mucosal composition could be used to identify biomarkers of exposure for environmental contaminants or stressors.

There have been significantly fewer studies of the identification of RNA in fish epidermal mucus and its potential to be used as an ecotoxicological biomarker. RNA in the mucus may arise from dead cells on the epidermal surface or from active transport of RNA from the nucleus to the cell membrane by exosomes and microvesicles. However, the role of RNA in the mucus remains unknown. Mucosal mRNA in channel catfish (*Ictalurus punctatus*) presented with a bacterial challenge showed differential expression of immune-related transcripts using quantitative polymerase chain reaction (qPCR), suggesting transcriptional changes occur in mucus and can be used to assess the whole-animal stress response. Whole-transcriptome sequencing of mucus has yet to be performed and may provide insight into the potential use of mucus mRNA biomarkers. In this study, we examined the efficacy of whole-transcriptome profiling of epidermal mucus as a novel method for detection of oil exposure in mahi-mahi. Juvenile mahi-mahi were exposed to control seawater or high-energy water accommodated fraction (HEWAF) dilutions of DWH slick oil for 48 h, followed by mucus collection, RNA sequencing, differential expression analysis, and adverse biological pathway prediction.

## MATERIALS AND METHODS

### Oil Preparation and Exposures.

High-energy water accommodated fractions (HEWAFs) were prepared from crude oil obtained during surface skimming (OFS) following the DWH oil spill and transferred to the University of Miami under the chain of custody (sample ID OFS-20100719-Juniper-001 A00884). The HEWAF solutions were prepared according to established methods and were diluted to nominal concentrations of 5% or 10% using ultraviolet-sterilized seawater for testing (Supplemental Methods). Juvenile mahi-mahi were exposed to control seawater or high-energy water accommodated fraction (HEWAF) dilutions of DWH slick oil for 48 h, followed by mucus collection, RNA sequencing, differential expression analysis, and adverse biological pathway prediction.

### Water Quality and Water Chemistry Analysis.

Measured \( \Sigma \text{PAH} \) concentrations were calculated using a combination of gas chromatography/mass spectrometry-selective ion monitoring (GC/MS-SIM) and a previously described fluorescence method. The former was used to measure initial concentrations from the 10% HEWAF preparations at the start of exposure and following the 24 h water change. For this purpose, composite samples of approximately equal volumes were collected from each of the four replicate 10% HEWAF tanks in 250 mL amber glass bottles with no head space, immediately stored at 4 °C, and shipped overnight on ice to ALS Environmental (Kelso, WA). For fluorescence analysis, composite samples for both initial and final time points (initial only for controls) were collected from each treatment (including 10% HEWAF) for the first 24 h and following the water change. \( \Sigma \text{PAH} \) concentrations for all other treatments and time points were then calculated by using fluorescence values to scale samples relative to the measured concentrations for the 10% HEWAF GC/MS-SIM and fluorescence values. \( \Sigma \text{PAH} \) values presented represent geometric means ± the standard error of the mean (SEM) of initial and final \( \Sigma \text{PAH} \) concentrations from the 2 days of exposure (initials only for controls). Water quality parameters were measured throughout exposures (Table S1).

### Mucus Collection and Sequencing Preparation.

At the end of the 48 h exposure period, fish were collected in a small net and euthanized in a solution of tricaine methanesulfonate (MS-222). Fish were removed from solution and dabbed dry with a paper towel prior to the collection of mucus. Mucus was sampled by scraping a small weighing spatula lightly across the skin surface and then collected by scraping the mucus into a 2 mL collecting tube. Although fish were lightly scraped to avoid collecting epithelial cells, it is possible that some epithelial cells were present in each sample. All four fish per replicate tank were pooled into one sample, for a total of four samples per treatment group. Samples were flash-frozen in liquid nitrogen immediately upon collection and then stored at −80 °C until they were further processed.

Total RNA was extracted from mucus samples using the RNeasy Lipid Tissue Mini Kit (Qiagen, Germantown, MD) following the manufacturer’s protocol. Following extraction, total RNA concentrations were quantified using an Invitrogen Qubit 4 Flurometer (Invitrogen, Carlsbad, CA), and RNA integrity was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

To enrich mRNA collection, ribosomal RNA (rRNA) was removed from total RNA using the NEBNext mRNA Depletion Kit (New England Biolabs, Ipswich, MA) prior to library preparation. Sequencing libraries for each sample were then prepared with the NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs) using 500 ng of total RNA as input. Final libraries were sequenced as single-end 75 bp reads in one lane on an Illumina NextSeq 500 high-throughput sequencer at the Institute for Integrative Genome Biology at the University of California, Riverside. Raw sequencing reads were submitted to the NCBI SRA database (accession code PRJNA526742).

### Identification of Differentially Expressed Genes and Pathways.

Processing of raw sequencing reads, *de novo* transcriptome assembly, and differential expression analysis were performed as previously described, with slight modifications. Briefly, rRNA sequences were first removed from each sample using sortmerna version 2.1 (\( e \) value of <0.001). Poor quality reads and adapters were then removed using trimmomatic version 0.36 and read quality was assessed with FastQC version 0.11.7. The remaining reads were used to build a *de novo* transcriptome with Trinity version 2.8.4, followed by transcript quantification with RSEM version 1.3.1 and differential expression analysis with DESeq2 version 1.23.6 at the \( p \leq 0.05 \) level after Benjamini–Hochberg false discovery rate (FDR) correction.
Functional annotation of differentially expressed genes was performed using Trinotate version 3.1.1. Differentially expressed (FDR-adjusted $p \leq 0.05$) and annotated transcripts were input into Ingenuity Pathway Analysis (IPA), along with their log-fold change values, to identify altered molecular and biological pathways (FDR-adjusted $p \leq 0.05$). A full list of bioinformatics parameters can be found in Supplemental Methods.

qPCR Validation of Differential Expression. qPCR was used to verify the expression status of select genes differentially expressed in the RNASeq analysis. cDNA was generated using the Promega Reverse Transcription System (Promega, Madison, WI) with 500 ng of total RNA as input. The 20 μL final volume was diluted to 100 μL using nuclease-free H2O to provide the final cDNA concentration. qPCR mixtures consisted of 10 μL of SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, Hercules, CA), 0.2 μM forward and reverse primers, and 2 μL of cDNA in a final volume of 20 μL. Each sample was performed in triplicate on a CFX Connect Real-Time PCR Detection System with the following thermal cycler conditions: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s, 58 °C for 30 s, and 72 °C for 30 s. Relative expression was determined by the $2^{-\Delta\Delta Ct}$ method with 18S rRNA as the normalizing gene. Primer sequences and amplicon lengths are listed in Table S2.

**RESULTS AND DISCUSSION**

Geometric mean $\sum$PAH concentrations in the low- and high-HEWAF preparations were 16.55 ± 7.69 and 23.03 ± 1.23 μg/L, respectively (Table S3). Due to the extensive aeration required to meet the high metabolic demands of juvenile mahi-mahi, significant depletion of soluble PAHs occurred over each 24-h period and may have contributed to the relatively similar PAH concentrations in the low and high HEWAF. The PAH composition determined by GC/MS-SIM showed that three-ring PAHs such as phenanthrenes/anthracenes comprised ~70% of the total PAHs (Table S4). The observed PAH composition was similar to the results of other studies showing that weathering of DWH oil at the surface resulted in a loss of low-molecular weight PAHs and a relative enrichment of three- and four-ring PAHs compared to nonweathered source oil.18 Following rRNA removal from the raw reads, 179 million reads were used for de novo transcriptome assembly (Table S5). Trinity assembled 194,282 genes with an N50 contig length of 1121 bp and an average contig length of 708 bp based on the longest isoform per gene (Table S6), showing that an abundance of RNA molecules are present in mahi-mahi mucus. A Euclidian distance heat map was then generated to assess if HEWAF exposure induces overall transcriptomic differences in the mucosal RNA. Clearly defined clades were

---

**Figure 1.** Results from RNA sequencing of mahi-mahi epidermal mucus, including (A) a heat map of Euclidean distances between samples in the control, low-HEWAF ($\sum$PAH = 16.55 μg/L), and high-HEWAF ($\sum$PAH = 23.03 μg/L) exposure groups calculated from DEseq2 variance stabilizing transformation of the RSEM count data, (B) a volcano plot of differentially expressed transcripts within the low- and high-HEWAF exposure groups, and (C) the top canonical pathways, physiological system development and function (D & F) pathways, and toxicity functions predicted to be altered in the low-oil exposure group by Ingenuity Pathway Analysis. Full lists of all pathways for both low- and high-HEWAF exposures can be found in Tables S8–S12.
depicted for each treatment, indicating transcriptome-wide expression changes in the mucus of HEWAF-exposed fish compared to non-exposed controls (Figure 1A). Furthermore, clear separation of low- and high-HEWAF treatments shows that increasing PAH concentrations further alters the mucosal transcriptome.

Analysis of differentially expressed transcripts showed that the mucosal transcriptome was significantly altered following PAH exposure. There were 501 differentially expressed transcripts in the low-HEWAF exposure (227 upregulated, 274 downregulated) and 196 differentially expressed transcripts in the high-HEWAF exposure (121 downregulated, 75 upregulated) (Figure 1B). The low and high treatments shared 136 of the same differentially expressed transcripts (Figure S1), suggesting common transcriptional responses occur within the range of PAH concentrations tested. Cytochrome P450 enzymes (cyp1a1 and cyp1b1) are well-established biomarkers of PAH exposure and were among the most highly upregulated transcripts in both the low- and high-HEWAF exposures (Table S7). The efficacy of cyp1a1 induction as a biomarker for PAH exposure has also been examined with other nonlethal sampling methods, with greater cyp1a1 upregulation observed in the caudal fin than in the liver of marine diesel-exposed juvenile coho salmon (Oncorhynchus kisutch).28 Thus, upregulation of cyp1a1 appears to show a robust molecular response using nonlethal sampling methods and may be a useful biomarker candidate for PAH exposure.

To further confirm the RNASeq results, the expression of several differentially expressed transcripts was validated with qPCR. Genes that are important for xenobiotic metabolism (cyp1b1), Ca²⁺ homeostasis (rr2), and immune function (socs3, il18, and il17f) were selected and confirmed the expression patterns observed in RNASeq (Figure S2).

Predictions of biological pathways, functions, and diseases most significantly affected by gene expression changes were identified using IPA. The top-ranked canonical pathways (ranked by p value) in the low treatment included EIF2 signaling, the Th17 activation pathway, the 2-ketoglutarate dehydrogenase complex, nicotine degradation III, the osteoarthritis pathway, the STAT3 pathway, and aryl hydrocarbon receptor signaling (Figure 1C and Table S8). Many of these top canonical pathways, such as EIF2 signaling, the Th17 activation, and STAT3, are involved in cancer pathways [number 1 ranked disease and disorder (Table S9)], immune-related responses, and general stress responses.29–31 In the high treatment, the top-ranked canonical pathways were Wnt/β-catenin signaling, nicotine degradation III, nicotine degradation I, melatonin degradation I, the superpathway of melatonin degradation, the osteoarthritis pathway (Table S8).

Next, the top physiological system and development functions predicted to be altered were examined. Immune cell trafficking was the top-ranked function, followed by skeletal and muscular system development and function, and tissue development (Figure 1C and Table S10). Other immune-related functions were also highly ranked in this category, including perturbations in lymphoid tissue structure and development (rank 6) and cell-mediated immune response (rank 12). Furthermore, five of the top 21 diseases and disorders were related to immune function (Table S9). In contrast to the full data set, which contained a nearly equal number of up- and downregulated genes in the low-HEWAF exposure, downregulation of immune genes was observed and...
subsequently predicted suppression of immune function (Figure 2A and Table S11). PAH-induced suppression of immune-related functions, gene transcription, and lymphocyte function has been well-documented in several fish species and may be related to AhR activation.\textsuperscript{32}−\textsuperscript{35} Transcriptional repression of immune-related genes also occurred in the epidermal mucus of channel catfish following bacterial exposure, which could have downstream effects on the susceptibility to pathogen infection.\textsuperscript{15,32,36} In our study, downregulation of immune-related genes was predicted to inhibit the inflammatory response, leukocyte migration, activation of leukocytes, activation of phagocytes, T cell development, and antibacterial responses (Figure 2A).\textsuperscript{33} Moreover, predicted alterations in immune function are consistent with proteomic studies of mucus in which the abundance and compositional profile of immune-related proteins are altered in response to biotic and abiotic stressors.\textsuperscript{8,14} Together, our data suggest that mucosal mRNA abundance may be indicative of whole-animal changes in immune-related function in PAH-exposed fish.

IPA also predicted alterations in toxicity functions induced by oil exposure. The top-ranked toxicity functions in the low-HEWAF exposure were liver fibrosis, liver hyperplasia, congenital heart anomaly, and cardiac enlargement (Figure 1C and Table S12). Similarly, liver enlargement, pulmonary hypertension, liver hyperbilirubinemia, and liver hyperplasia were the top-ranked toxicity functions in the high-HWAF exposure. Of all toxicity functions, 17 cardiac functions were predicted to be altered in the low-HEWAF exposure and 11 in the high-HEWAF exposure, with many overlapping categories such as cardiac enlargement, cardiac arrhythmia, cardiac fibrosis, and cardiac necrosis/cell death (Table S12). In addition, there was a predicted inhibition of cardiac muscle function, cardiac muscle contractility, and abnormality of the heart ventricle (Figure 2B), which are cardiac phenotypes known to be altered by crude oil-derived PAH exposure in fish.\textsuperscript{37,38} Alterations in Ca\textsuperscript{2+} homeostasis were also predicted from the mucus transcriptional profile, such as decreased mobilization of Ca\textsuperscript{2+} and an increased quantity of Ca\textsuperscript{2+} (Figure 2B). Transcripts of ryanodine receptor 2 (\textit{ryr2}), the primary mediator of calcium-induced Ca\textsuperscript{2+} release required for cardiomyocyte contraction, were upregulated in the mucus and have also been shown to be dysregulated in oil-exposed mahi-mahi and Atlantic haddock (\textit{Melanogrammus aeglefinus}) embryos (Table S11).\textsuperscript{39,40} The predicted changes in both Ca\textsuperscript{2+} homeostasis and cardiac functions are likely linked, supported by previous studies showing that inhibition of cardiomyocyte Ca\textsuperscript{2+} fluxes is a key component of functional defects in the developing heart following embryonic PAH exposure.\textsuperscript{38,41} Aside from its role in cardiac function, the modulation of intracellular Ca\textsuperscript{2+} levels may also be linked to PAH-induced immunosuppression and, therefore, may play a role in our predicted suppression of immune-related functions.\textsuperscript{42} Overall, the differential expression of cardiomyocyte Ca\textsuperscript{2+} cycling and homeostasis genes as well as the associated dysregulation of cardiovascular pathways in the mucus of juveniles is consistent with altered expression patterns in whole-embryo homogenates of mahi-mahi and other marine fish.\textsuperscript{39,40,43}

To the best of our knowledge, this is the first study to sequence a fish mucosal transcriptome and characterize transcriptional dysregulation in the mucus following pollutant exposure. HEWAF exposure in juvenile mahi-mahi induced transcriptional changes in cytochrome P450 enzymes as well as immune and cardiac transcripts in the mucus that resemble whole-animal changes in transcription following PAH exposure. The consistency of our findings with other studies suggests that transcript quantification in the epidermal mucus is an effective method for the assessment of oil exposure. Moreover, the proposed mRNA-based analysis of mucus could be expanded for assessment of other environmental contaminants.

**ASSOCIATED CONTENT**

## Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acselett.9b00479.

Venn diagram of overlapping differentially expressed genes (Figure S1), a comparison of qPCR and RNAseq (Figure S2), water quality (Table S1), qPCR primer sequences (Table S2), PAH measurements by fluorescence and GC/MS-SIM (Tables S3 and S4), read depth and assembly statistics (Tables S5 and S6), differentially expressed genes (Table S7), IPA-predicted pathway enrichment, diseases, physiological development (Tables S8−S10), immune, cardiac, and stress genes differentially expressed (Table S11), toxicity function enrichment (Table S12), and Supplemental Methods (PDF)

**AUTHOR INFORMATION**

### Corresponding Authors

*Address: 2460A Geology, Department of Environmental Sciences, University of California, Riverside, CA 92521. Telephone: 1-951-827-2018. Fax: 1-951-827-3993. E-mail: jgreer@ucr.edu.*

*E-mail: nandr009@ucr.edu.*

### ORCID

Justin B. Greer: 0000-0001-6660-9976

### Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

This research was made possible by a grant from The Gulf of Mexico Research Initiative [Grant SA-1520, Relationship of Effects of Cardiac Outcomes in Fish for Validation of Ecological Risk (RECOVER)]. Data are publicly available through the Gulf of Mexico Research Initiative Information & Data Cooperative (GR11DC) at [https://data.gulfresearchinitiative.org](https://data.gulfresearchinitiative.org) (doi: 10.7266/n7-m7d5-nk45). M.G. is a Maytag Professor of Ichthyology.

### REFERENCES


