Transcriptomic responses to an Alaskan oil spill over time reveal a dynamic multisystem involvement in exposed mussels (*Mytilus trossulus*)

Lizabeth Bowen[1], William B. Driskell[2], Brenda Ballachey[3], James R. Payne[4], Shannon Waters[1], Eric Litman[5], Austin Love[6]

[1] *U.S. Geological Survey, Western Ecological Research Center, Davis, CA 95616,*

[2] *Consultant, Seattle, WA*

[3] *U.S. Geological Survey (Emeritus), Alaska Science Center, Anchorage, AK*

[4] *Payne Environmental Consultants, Encinitas, CA*

[5] *NewFields Environmental Forensics Practice LLC, Mansfield, MA*

[6] *Prince William Sound Regional Citizens' Advisory Council, Valdez, AK*

January 2023

PWSRCAC Contract 951.22.07

The opinions expressed in this commissioned report are not necessarily those of PWSRCAC.

The final report for this project is being prepared for publication in a peer-reviewed journal.

Corresponding author: L. Bowen, *lbowen@ucdavis.edu, 530-752-5365*

Transcriptomic responses to an Alaskan oil spill over time reveal a dynamic multisystem involvement in mussels (*Mytilus trossulus*)

Lizabeth Bowen[a*], William B. Driskell[b], Brenda Ballachey[c], James R. Payne[d], Shannon Waters[a], Eric Litman[e], Austin Love[f]

[a] *U.S. Geological Survey, Western Ecological Research Center, One Shields Avenue, Davis, CA 95616, USA; lbowen@usgs.gov; swaters@usgs.gov*
[b] *Consultant, 6536 20th Avenue NE, Seattle, WA 98115, USA; bdriskell@comcast.net*
[c] *U.S. Geological Survey (Emeritus), Alaska Science Center,4120 University Drive, Anchorage, AK 99508, USA; beballachey@gmail.com*
[d] *Payne Environmental Consultants., 1651 Linda Sue Lane, Encinitas, CA 92024, USA; jrpayne@sbcglobal.net*
[e] *NewFields Environmental Forensics Practice LLC, 300 Ledgewood Place, Suite 205, Mansfield, MA 02370, USA; elitman@newfields.com*
[f] *Prince William Sound Regional Citizens' Advisory Council, 130 S. Meals Suite 202, Valdez, AK 99686, USA; Austin.Love@pwrsrcac.org*
*Corresponding Author: Lizabeth Bowen lbowen@ucdavis.edu

Abstract

In response to a minor shoreline spill in Port Valdez, AK, a time series of mussels (*M. trossulus*) was collected and analyzed for oil burdens and transcriptome response. In general, transcription results show that following higher levels of tissue PAH burdens, significant physiological responses occurred. Whole tissues analyzed for the full suite of forensic oil hydrocarbons revealed weathering patterns and purging processes. The transcriptome of adductor muscle tissue showed multi-pathway effects and timing in gene activities related to the detoxification and recovery processes. In multiple pathways, gene activities did not return to reference-site levels, suggesting that recovery from hydrocarbons was not complete by the final sampling. This may have been due to residual sheening prolonging recovery. Genes that could potentially distinguish between ANS crude oil and harbor contaminants (pyrogenics and diesel) were identified with the goal of developing more robust monitoring tools.

Introduction

Historically, oil-spill monitoring in the marine environment has relied heavily on hydrocarbon identification and assessing tissue burdens as an indicator of a toxic insult, e.g., exceeding physiological threshold levels. Current assessment practices still emphasize polycyclic aromatic hydrocarbons (PAH) as the main concern in organismal toxicity (Marris et al., 2020). Yet, from a chemistry perspective, crude oil is a complex mixture of tens of thousands of different components (McKenna et al., 2013), and evaluating the toxicity of these hydrocarbons is an emerging science.

Working with just the established science, exposure to PAH can induce pathophysiological changes that may be subtle yet significant, and difficult to detect using classical diagnostic methods (Bodkin et al., 2014; Hylland, 2006; Incardona et al., 2014; Peterson et al., 2003). Because both immediate and long-term responses to oil are a function of gene expression, molecular techniques (e.g., transcriptomics) provide a set of tools for detecting exposure, understanding mechanisms of lethal and sublethal responses, and assessing population level effects (Portnoy et al. 2020). Altered gene expression may warn of broader ecosystem effects prior to visible manifestations in the organism or population (Farr and Dunn, 1999; McLoughlin et al., 2006; Poynton and Vulpe, 2009; Bowen et al., 2018).

When organisms are exposed to oil, molecular responses occur amongst multiple pathways to maintain homeostasis and facilitate excretion of toxins (Portnoy et al. 2020). But while organisms may initiate molecular responses focused on maintenance of homeostasis, these responses can be overburdened by high concentrations of oil, leading to the accumulation of bioactive stressors (Wills et al. 2009, Androutsopoulos et al. 2009). During the purging process, additional bioactive toxic compounds can also be produced that consequently may increase the exposure effects (Schlenk et al. 2008).

Much of our sublethal-oil-effects information has been derived from studies of vertebrates, with very little known about molecular and physiological responses of benthic invertebrates (Jenny et al. 2016). Filter-feeding mussels are excellent sentinels for the presence, bioavailability, exposure, and persistence of oil in the ecosystem (Lauenstein et al., 1993; Livingstone et al., 2000; Bolognesi and Cirillo, 2014; Short and Springman, 2016; Beyer et al. 2017; Bowen et al., 2018). They are integral to the nearshore food web as a significant food source for both marine and terrestrial predators, and where abundant, are major structural components of the intertidal community. The strength of evaluating coastal ecosystem health using

gene transcription in combination with biomarker assays has been demonstrated by studies in Alaska (Counihan et al. 2019; Bowen et al. 2020) as well as mussel biomonitoring studies in other coastal regions including the Mediterranean (Carella et al., 2018; Sforzini et al., 2018). In fact, alterations in gene expression (even in the presence of low body burdens of toxins) suggest physiological impairment predictive of population-level effects (Whitehead et al. 2012).

In some sense, accumulation and depuration of contaminants are mirror processes. Accumulation occurs through ingestion or absorption (e.g., finite oil droplets or dissolved components in siphon, mantle, or gill tissues) and inversely, depuration through gut elimination, diffusive desorption, or metabolism, e.g., conversion through the aryl-hydrocarbon-receptor (AHR) pathway (Zanette et al., 2013). Higher ambient oil concentrations and longer duration chronic exposures imply increased accumulation and longer depuration times. These factors, plus the variety of hydrocarbons each with individual water-versus-lipid solubilities, and even the prior health of the mussels, dictate that each spill event will be unique in degree of impacts and achieving recovery. We posit that examining transcription activities would be more insightful than the traditional approach of comparing chemistry body burdens against reported threshold levels of effects. Here, mussel transcription responses and body hydrocarbon burdens are documented following an acute oiling event and subsequent diminishing chronic exposure.

In April 2020, a minor oil spill occurred at Alyeska's Valdez Marine Terminal, Prince William Sound, Alaska, whereby an estimated 1,400-gallons (34 bbls) of Alaska North Slope (ANS) crude oil overflowed from a sump well, traversed undetected as a subsurface downslope plume (below ground and snow-cover), and subsequently emerged at the nearby shoreline, creating slicks and necessitating a full-scale marine response (Figure 1). Mussels were collected from the intertidal area at the spill site ("hotzone," HOT) as well as remote locations within the Port Valdez fjord at Jack Bay and Galena Bay (BAY1, negative controls), with the goals of determining oil concentrations for the impacted mussels and elucidating the dynamic transcriptomic processes triggered by the spill event. Mussels from the City of Valdez boat harbor (HAR) were sampled with the goal of providing a contrasting positive control to determine transcriptional differences between mussels exposed to spilled ANS crude oil and those exposed to a mixture of chronic hydrocarbons (e.g., diesel, lubricants, hydraulics, creosote residues, and combustion products) as commonly seen in a boat harbor. Related mussel samples were also collected from two Terminal sites, Saw Island and Jackson Point, as part of an annual Long Term Environmental Monitoring Program (LTEMP).

Materials and Methods

Sampling

The Valdez Marine Terminal (Prince William Sound, Alaska) is the terminus of the Trans-Alaska Pipeline bringing Alaska North Slope (ANS) crude oil to tankers in Port Valdez for shipment to west coast refineries. The spill comprised primarily ANS crude oil plus terminal runoff. As the spill reached the shoreline, the plume was containment boomed and partially removed, but sheening from lingering oil still leaking from the sump and the saturated intertidal sediments and riprap continued through the study. Sampling began 18 days after the initial spill event. Mussels for chemistry and transcriptome analyses were collected concurrently throughout the spill and perceived recovery period (Table 1).

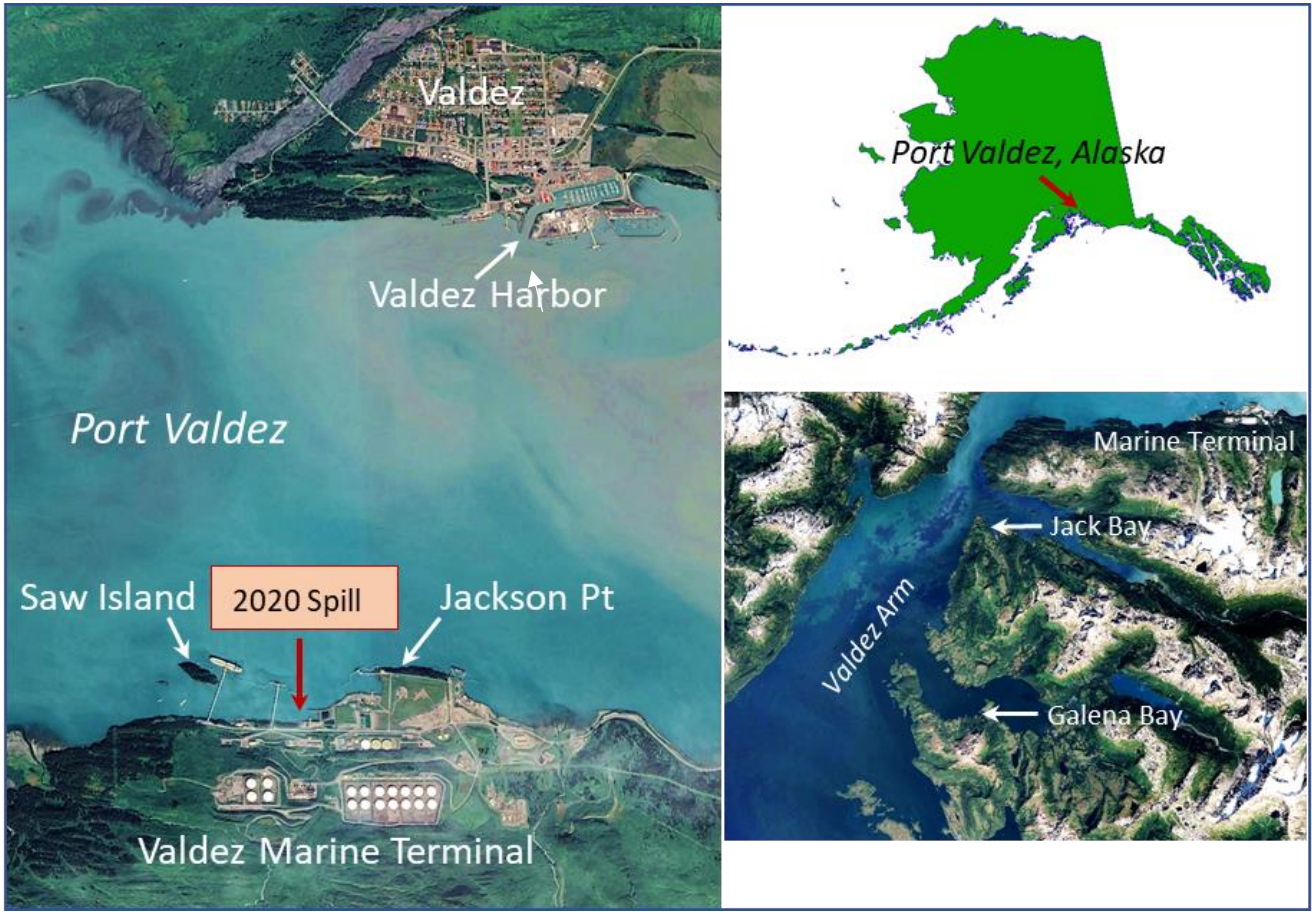


Fig. 1. Overview of Port Valdez showing the April 12, 2020 intertidal spill location at the Valdez Marine Terminal. Mussels were sampled at the Spill site ("Hot Zone") and LTEMP sites at Jackson Point and Saw Island. Regional background samples were collected at Jack Bay and Galena Bay (collectively, BAY1), and at the entrance to the Valdez boat harbor (HAR) in June 2020.

Table 1. *Mytilus trossulus* adductor muscle samples included in chemistry and transcription analyses. HARA is an average of samples collected from the left and right sides of the Valdez harbor mouth in 2020; BAY1A is an average of samples collected from Jack and Galena Bays in 2020.

Location	Sampling Date*	Elapsed days of study	Chemistry replicates**	Transcription replicates***
Spill site (HOTA)	4/30/2020	1	3	3
	5/13/2020	14	3	3
	5/20/2020	21	3	3
	5/27/2020	28	2	3
	6/3/2020	35	2	3
	6/11/2020	43	3	3
	6/18/2020	50	NA	3
	6/24/2020	56	NA	3
	7/6/2020	68	NA	3
	7/22/2020	82	2	3
	8/19/2020	111	NA	3
Valdez boat harbor (HARA)	6/8/2020	40	6	6
Negative control, Jack and Galena Bays (BAY1A)	6/9/2020	41	6	6

*Initial spill event (sheen detected) was 12 Apr 2020.
**Chemistry replicates were composites of up to 30 mussels.
***Transcriptome replicates consisted of pooled samples of 3-4 individuals per pool.

Chemistry

Collection and analytical methods are generally patterned after the NOAA Mussel Watch Program (Apeti et al., 2012; Payne et al., 2008; Payne et al., 2021). Briefly, three replicates of 30 mussels each were collected by hand at each site and immediately frozen for transport and archival before thawing, whole-body tissue compositing, homogenization, and extraction. Analyses were provided by Alpha Analytical Laboratory (Mansfield, MA) under the guidance of NewFields Environmental Forensics Practice LLC (Rockland, MA).

The PAH, alkylated PAH, and steranes/triterpanes (S/T) were analyzed as semi-volatile compounds using selected-ion-monitoring gas chromatography/mass spectrometry (SIM GC/MS) via a modified Environmental Protection Agency (EPA) Method 8270 aka 8270M (Stout and Wang, 2016). This analysis provides the concentration of approximately 80 PAH, alkylated PAH homologues, individual PAH isomers, and sulfur-containing aromatics, plus approximately 50 petroleum biomarkers (S/T) including tricyclic and pentacyclic triterpanes, regular and rearranged steranes, and triaromatic and monoaromatic steroids. For saturated hydrocarbons (SHC), a high-resolution gas chromatography/flame ionization detector (GC/FID) profile using modified EPA Method 8015B reports total extractable materials (TEM; C9-C44), n-alkanes (C9-C40) and selected (C15-C20) acyclic isoprenoids (e.g., pristane and phytane). The full list of analytes and plot abbreviations are presented in the appendix, SI-1.

Transcriptomics

Collection methods are generally patterned after standard sampling protocols (Bowen et al., 2018; Counihan et al., 2019). Briefly, mussels were collected by hand at each site and adductor muscles immediately excised and placed into RNAlater®. Samples were frozen and stored at -80°C until processing. In the lab, tissues were subsequently homogenized, and RNA was extracted according to standard protocols. All laboratory analyses follow MIQE guidelines (Bustin et al., 2009).

RNA quantification and qualification

RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library preparation for Transcriptome sequencing

Messenger RNA was purified from total RNA using poly-T oligo-attached magnetic beads. After fragmentation, the first strand cDNA was synthesized using random hexamer primers, followed by the second strand cDNA synthesis using either dUTP for directional library or dTTP for non-directional library. For the non-directional library, it was ready after end repair, A-tailing, adapter ligation, size selection, amplification, and purification. For the directional library, it was ready after end repair, A-tailing, adapter ligation, size selection, USER enzyme digestion, amplification, and purification. The library was checked with Qubit and real-time PCR for quantification and bioanalyzer for size distribution detection. Quantified libraries are pooled and sequenced on Illumina platforms, according to effective library concentration and data amount.

Clustering and sequencing

The clustering of the index-coded samples was performed according to Illumina's instructions. After cluster generation, the library preparations were sequenced on an Illumina platform and paired-end reads were generated.

Data Analysis

Quality control

Raw data (raw reads) of fastq format were first processed through in-house Perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low quality reads from raw data. At the same time, Q20, Q30 and GC content of the clean data were calculated. All the downstream analyses were based on the clean data with high quality.

Reads mapping to the reference genome

Reference genome and gene model annotation files were downloaded directly from the National Center for Biotechnology Information (NCBI) (Coordinators, 2016). Index of the reference genome was built using Hisat2 v2.0.5 and paired-end clean reads were aligned to the reference genome using Hisat2 v2.0.5. We selected Hisat2 as the mapping tool because it can generate a database of splice junctions based on the gene model annotation file and thus a better mapping result than other non-splice mapping tools.

Novel transcripts prediction

The mapped reads of each sample were assembled by StringTie (v1.3.3b) (Pertea et al. 2015) in a reference-based approach. StringTie uses a novel network flow algorithm as well as an optional de novo assembly step to assemble and quantitate full length transcripts representing multiple splice variants for each gene locus.

Quantification of gene expression level

The featureCounts v1.5.0-p3 software package was used to count the reads numbers mapped to each gene, and then FPKM of each gene was calculated based on the length of the gene and reads count mapped to the gene. FPKM, expected number of Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced, considers the effect of sequencing depth and gene length for the reads count at the same time, and is currently the most commonly used method for estimating gene expression levels.

Differential expression analysis

Differential expression analysis, i.e., determining whether the gene expression level differed between two or more groups/treatments/populations, was performed using the DESeq2 R package (1.20.0). DESeq2 provides statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting p-values were adjusted using the Benjamini- Hochberg procedure for controlling the false discovery rate. Genes with an adjusted p-value $\leq$ 0.05 found by DESeq2 were assigned as differentially expressed.

Prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by the edgeR program package through one scaling normalized factor. Differential expression analysis of two conditions was performed using the edgeR R package (3.22.5). The p-values were adjusted using the Benjamini-Hochberg method. Corrected p-value of 0.05 and absolute foldchange of 2 were set as the threshold for significantly differential expression.

GO and KEGG enrichment analysis

Gene Ontology (GO) enrichment analysis of differentially expressed genes was implemented by the clusterProfiler R package, in which gene length bias was

corrected. GO terms with adjusted p-value less than 0.05 were considered significantly enriched (either up or down regulated) by differentially expressed genes. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies. The clusterProfiler R package was used to test the statistical enrichment of differentially expressed genes in KEGG pathways. For clustering and plotting of GO functional groups, Revigo and Cytoscape programs were used.

Results

Chemistry

Oil sourcing and weathering in the Alaskan environment have been well documented in the decades since the 1989 *Exxon Valdez* spill (Stout and Wang, 2016). For this study, we first focus on the background hydrocarbon concentrations and profiles from two uncontaminated (non-oiled) control sites and then a positive control, the Valdez boat harbor with anthropogenic hydrocarbon exposures and where transcriptomic induction was observed in a previous study (Bowen et al. unpublished). Finally, we focus on the time-series, component-specific signals observed in the mussels from the spill site (Hot Zone) during the depuration and recovery process. The project was also interested in depuration rates for these cold-clime mollusks. Those aspects of the project are covered in the appendices (SI2 and SI3).

Both chemistry and spill-response observations establish that this spill was not a single exposure event. The mussels initially were acutely oiled by the emerging land plume (reported April 12; source mitigated April 13). Subsequently, the subsurface plume continued to sheen in diminishing amounts through the spring and summer, and persisted after the project's last sampling in late July (light sheening confirmed in SCAT reports; pers comm, T. Larson, AKDEC). The containment booms stayed in place until October 2020.

As the plume abated, total PAH trends showed exponentially decreasing concentrations in the time series (Figure 2). Concentrations from Day 1 of the study were 1000-fold greater than those from the last sampling on Day 82, which had levels 50-100 times background. By 2021, both the background and spill site samples had equivalent TPAH.

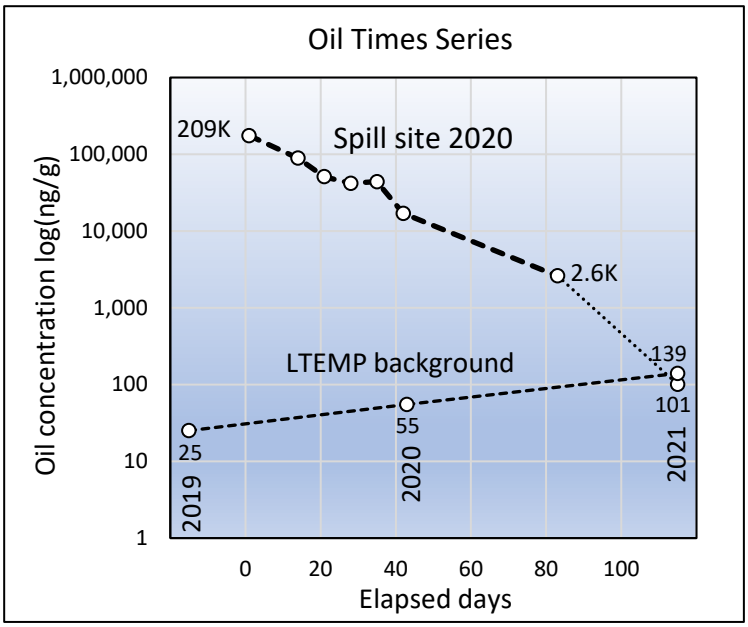


Figure 2. PAH trends at spill site and nearby LTEMP background samples, 2019-2021.

Chemistry differences in comparing the background control sites, the contaminated harbor, and the time-series profiles from the spill site (Hot Zone) are presented to examine the relationships between the PAH components and the transcription response. In the Supplemental Information (SI) section, Appendix 1, the standard forensic oil-hydrocarbon data are reported comprising the PAH, SHC, and S/T analytes along with their analyte abbreviations as used throughout this paper.

Background (negative control) Profiles

PAH profiles from the Jack Bay and Galena Bay sites (upper plots in Figure 3) show mostly below-method detection limit (MDL) naphthalenes, and combustion products at total PAH (TPAH) concentrations in the 25-35 ng/g dry weight (DW) range. Most of the individual analytes are less than 4 ng/g DW. Interestingly, the profiles obtained from those remote sites also closely resemble the signal obtained at a slightly higher TPAH level (55 ng/g DW mostly from phenanthrene, PHN) collected at the same time from the annual LTEMP monitoring station at Saw Island, which was not heavily impacted by the April 2020 sump overflow (or had returned to background levels and patterns at the time of sampling). All the PAH are at or just below the method

detection limit (MDL) and show the parent PAH greater than its alkylated homologues (e.g., PHN > PA1 > PA2) reflecting a combustion-product origin.

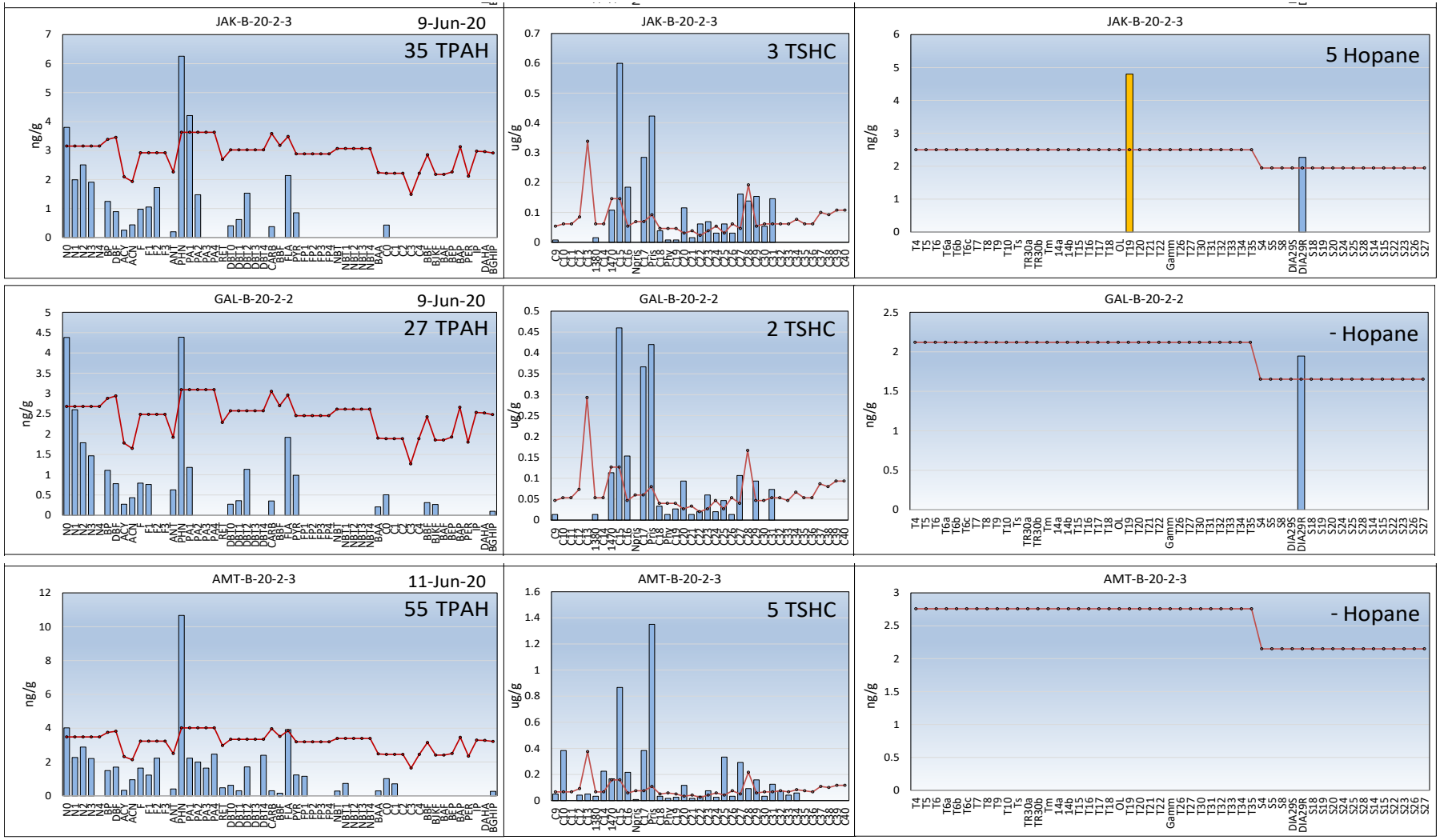


Figure 3. PAH, SHC, and S/T profiles from mussels sampled on 9 June 2020 at the remote sites, Jack Bay (JAK) and Galena Bay (GAL) and the unoiled Saw Island site sampled on 11 June 2020 approximately 200 m west of the Hot Zone (spill site). Totals are presented in the upper right corner of each graph. The profiles show only traces of above-MDL combustion products, and biogenic marine and terrestrial SHC. The red line is the sample-specific method detection limit.

Likewise, from the SHC and biomarker plots, it is apparent that neither Jack Bay nor Galena Bay were contaminated from the spill oil (there are no tell-tale biomarkers present). The SHC profiles show primarily marine biogenic components (n-C_{15}, n-C_{17}, pristane) and higher-molecular-weight, odd-carbon-number, terrestrial plant waxes (n-C_{23}, n-C_{25}, n-C_{27} and n-C_{29}). At these remote sites, there is no evidence of the evenly repeating series of n-C_{12} through n-C_{20} alkanes plus phytane associated with lighter distillates (Stout and Wang, 2016).

Valdez Boat Harbor (positive control) Profiles

The mussel samples from the Valdez Boat Harbor were collected from the intertidal zone beneath the red and green navigation lights on the breakwater riprap entrance to the harbor (Figure 1). Note that the TPAH levels in harbor mussels (Figure 4) were about 30 times higher than in mussels from unoiled control sites (Figure 3). The PAH profiles, from June (Figure 4), showed elevated TPAH concentrations (977 and 916 ng/g, respectively) with combustion products dominating the higher-molecular-weight components (parent phenanthrene, fluoranthene, pyrene, and chrysene with trailing alkylated homologues). The SHC show primarily biogenic n-alkanes and isoprenoids (n-C_{15}, n-C_{17}, pristane), and higher-molecular-weight, odd-carbon-number, terrestrial plant waxes (n-C_{23}, n-C_{25}, n-C_{27} and n-C_{29}). The more complex, underlying pattern of odd and even-carbon numbered n-alkanes in the n-C_{12} to n-C_{20} range plus phytane suggest traces of lighter distillate products (e.g., IFO 180, diesel fuel oil #4; Stout and Wang, 2016). However, the PAH profiles lack the expected patterns for distillates, which either offers little support for the conjecture or they are overwhelmed by the dominant combustion products. In the biomarker plots, the descending T4-T6 terpanes hint at diesel but relative to T4-T6, the other biomarker levels exceed any expectation for diesel and are instead a close match to the crude oil (ANS) from the VMT spill. Thus, it is possible that traces of the spilled oil may have reached the entrance to the Valdez Boat Harbor, but the signal is confounded by harbor pyrogenics and diesel contaminants.

Hot Zone (spill site) Profiles

Weathering of the spill site mussel PAH profiles is seen by the time-series loss of the more volatile and water-soluble parent and lower-alkylated PAH within each homologous group (Figure 5). The red line superimposed over the PAH and S/T profiles represents the pattern expected from the source oil collected on study day 1 (18 days after the initial spill) normalized to C2-chrysene measured in each sample. The SHC are normalized against the measured n-C_{32}. The S/T overlay is the expected S/T profile from the study day 1 sample normalized against the measured hopane.

Significant quantities of the spill oil remain in the spill site samples through study day 82 (July 21, 100 days from the initial spill).

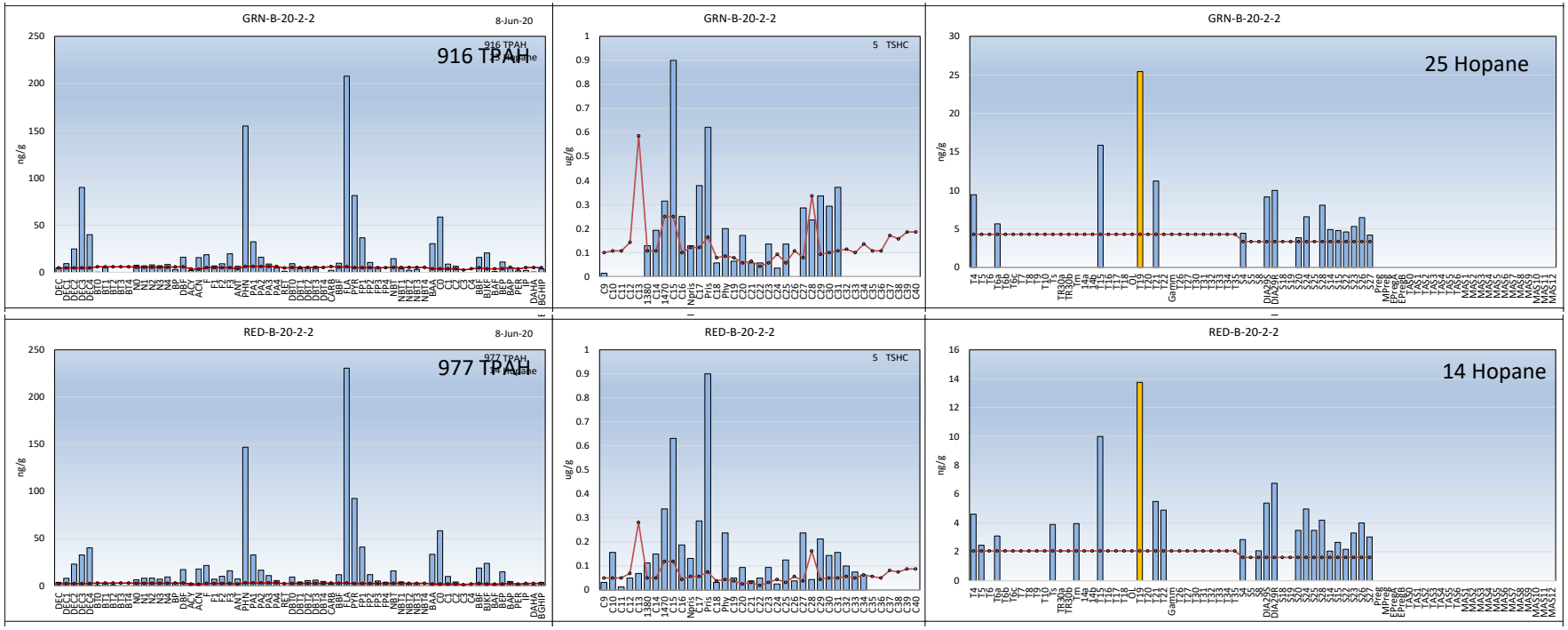


Figure 4. PAH, SHC, and S/T profiles from mussels at the entrance to the Valdez Boat Harbor. The PAH profiles show water-washed naphthalenes of petrogenic original and dominant parent-PAH reflecting pyrogenic (combustion) products, i.e., mixed hydrocarbon contaminants. The SHC profiles reflect traces of diesel oil, plus biogenic marine and terrestrial hydrocarbons. The S/T patterns suggest a trace of diesel and are also consistent with weathered ANS oil (possibly from the Terminal spill). The red line is the sample-specific method detection limit.

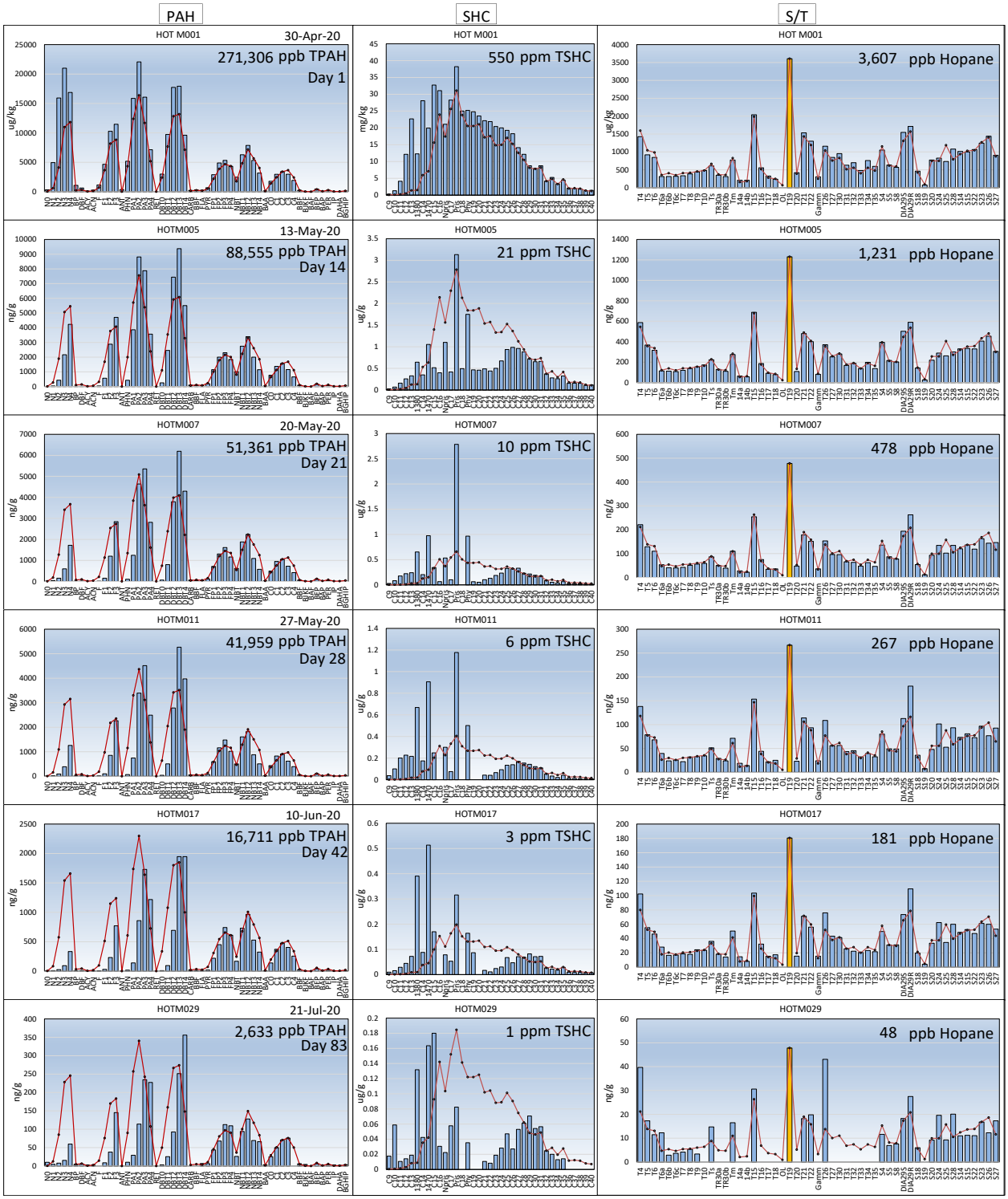


Figure 5. PAH, SHC, and S/T profiles of spill-site (Hot Zone) mussels (ppb) showing the weathering trends over time. Red lines denote the Day 1 oil profile with its PAH scaled to each sample's C-2 chrysene, SHC to n-C_{32}, and S/T to hopane.

The time-series plots of PAH, SHC, and S/T in mussels from the spill site (Figure 5) show monotonic declines in lower-molecular-weight PAH components and elimination of the fence-post pattern of odd- and even-carbon number, petroleum-derived n-alkanes from n-C9 through n-C40 in the fresher spilled oil (top profile, Figure 5), while biogenic isoprenoids (1380, 1470, and pristane) become more dominant with time (NRC, 1985).

Transcriptomics

Four approaches were used in examining the transcription data: Venn diagrams, KEGG pathway functionality, Gene Ontology functions, and selected gene trends.

Spill site, control, and harbor

The Venn diagram (Figure 6) shows the overlaps of how many genes had levels of expression in common between site groups (left) and sampling dates at spill site (right). Between site groups, approximately 1,526 genes were identified that were differentially expressed among mussels from the three different sites, the spill site (HOTA), negative control (Bay1A), and Valdez harbor (HARA). Transcription of 4,453 genes (central portion in Figure 6 left) was similar among the three locations, while transcription of 360 genes (HOTA), 586 genes (HARA), and 857 genes (BAY1A) were unique to the individual sites (Figure 6 left). Genes that were differentially expressed between HOTA and Bay1A, as well as HOTA and HARA were identified as potential genes for a transcript panel focused on discerning mussels exposed to ANS crude oil as opposed to diesel and petrogenic chemicals found primarily in the harbor areas.

Looking at spill-site genes for individual samples (Figure 6 right), the dynamic nature of gene expression across the time series becomes evident. Each sampling had 1,200-1,700 genes uniquely expressed on their respective dates as various pathway components are turned on or off in response to physiological status and environmental conditions. From just the 30 April and 19 Aug samples, there are 918 genes that turned off/on or expressed at significantly different levels between the series start and finish.

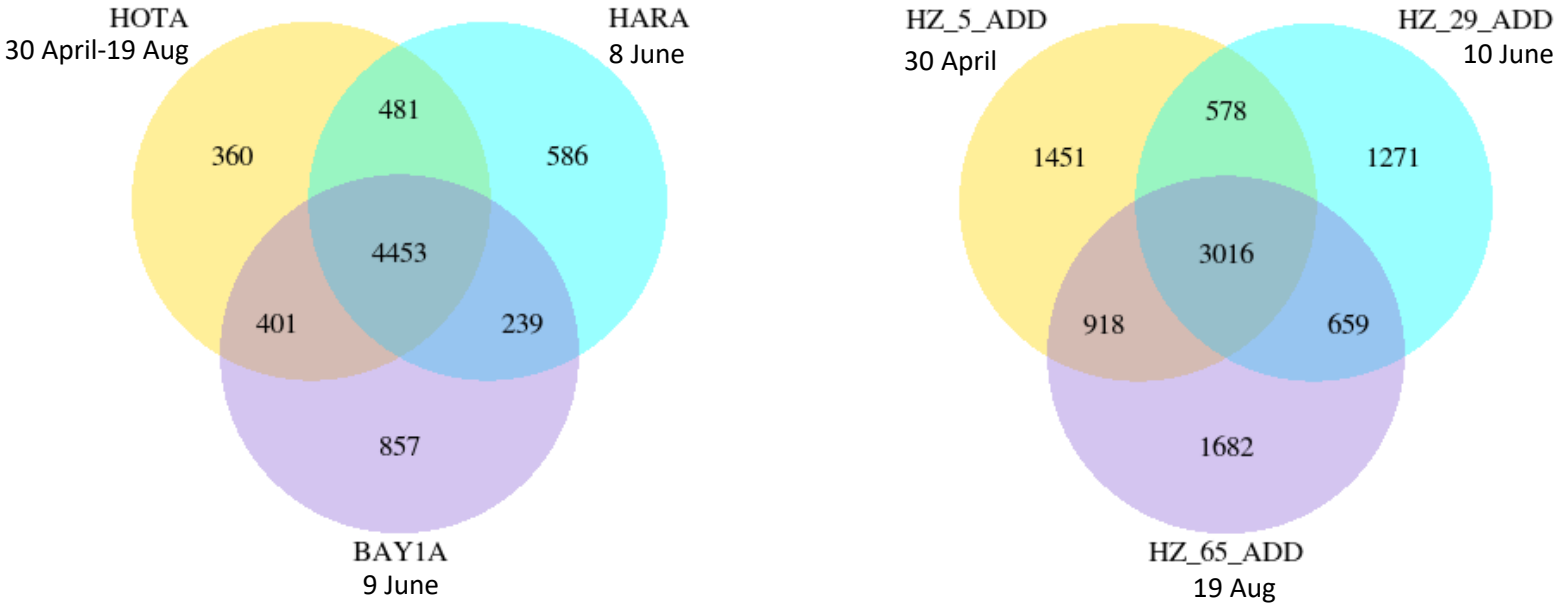


Figure 6. Venn diagrams of gene activity overlap between sampling groups (left) showing the number of differentially expressed genes in mussel adductor tissue that were unique or shared between samples collected at the spill site (HOTA, all dates), the Valdez small boat harbor (HARA, 8 June), and the control sites Jack and Galena Bay (BAY1A, 9 June). Right diagram compares beginning, mid and end-point dates from individual spill-site (hot zone, HZ) samples (i.e., samples collected on 30 April, 10 June, and 19 Aug).

From a more complex perspective, the field of bioinformatics has progressed to such an extent that many genes can be identified and assigned a putative function. Two separate approaches, KEGG and Gene Ontology (GO), were used to examine the active pathways and functional groupings for oiling effects. These analyses compared spill-site samples (i.e., from each of 11 sampling dates) to the unoiled control sample (BAY1A; Jack Bay and Galena Bay combined, one sampling date only) and plotted statistical differences or pathway linkages.

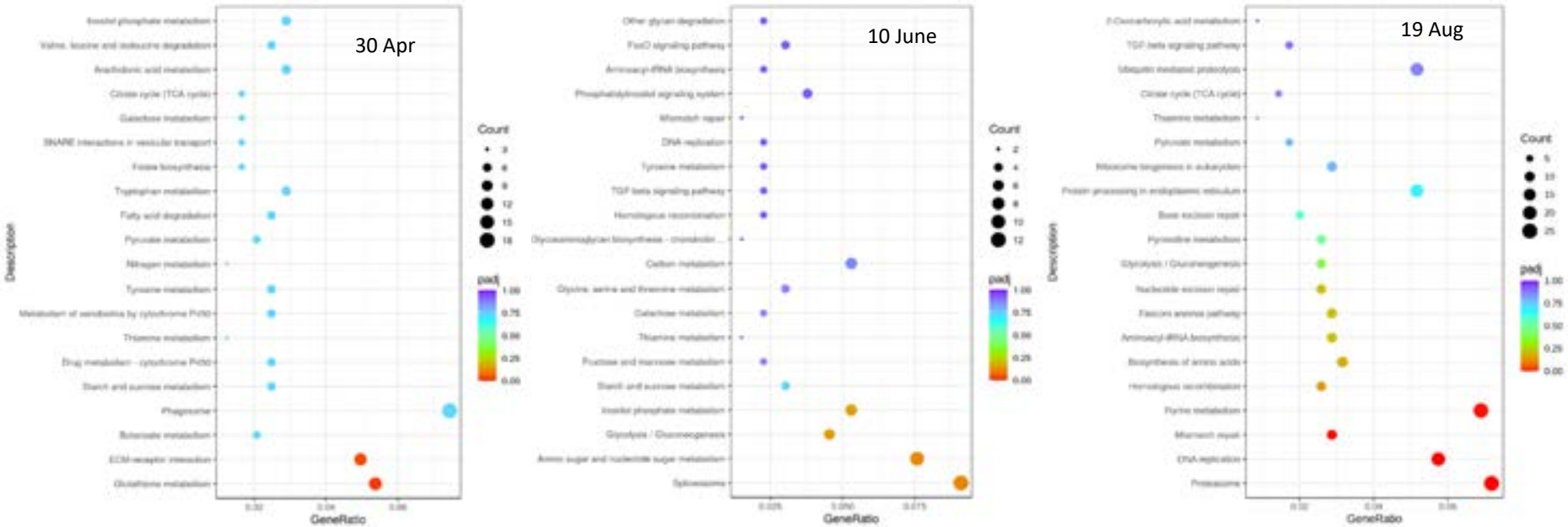
KEGG pathways

A multitude of KEGG pathways were enriched (either up or down regulated) in spill-site mussels in comparison with control mussels (Figure 7a). Significantly enriched pathways changed over time but included extracellular matrix (ECM)-receptor interaction, glutathione metabolism, purine metabolism, mismatch repair, DNA replication, and proteasome. When compared with harbor mussels (Figure 7b), significantly enriched pathways included ECM-receptor interaction, aminoacyl-rRNA biosynthesis, and proteasome.

Volcano plots

While Venn diagrams show a very generic view of the response of all differentially expressed genes (Figure 6), the dot plots show just the top twenty with maximum KEGG differences (Figure 7). In contrast, volcano plots (Figure 8) are multiparameter summaries. showing the counts of up- and down-regulated genes (i.e., increased vs. decreased transcription, respectively) and the significance and magnitude of their differential expression with their comparative site samples.

A. Spill site vs. Bay controls, single sampling dates



B. Spill sites vs. Harbor, single sampling dates

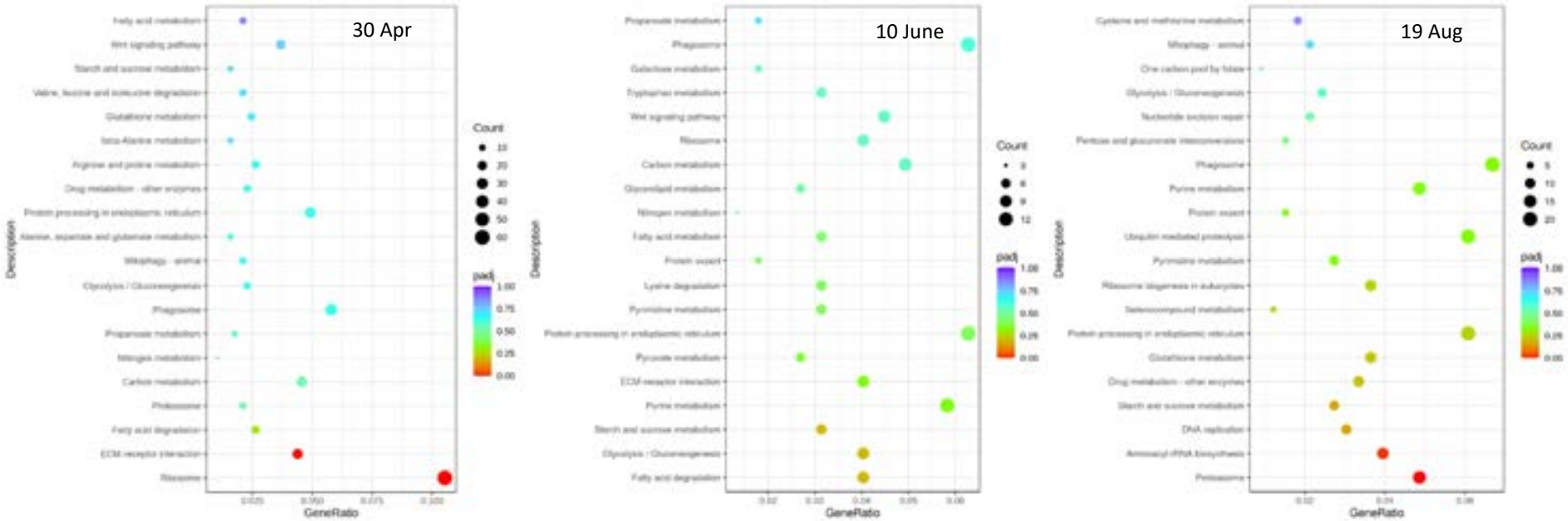
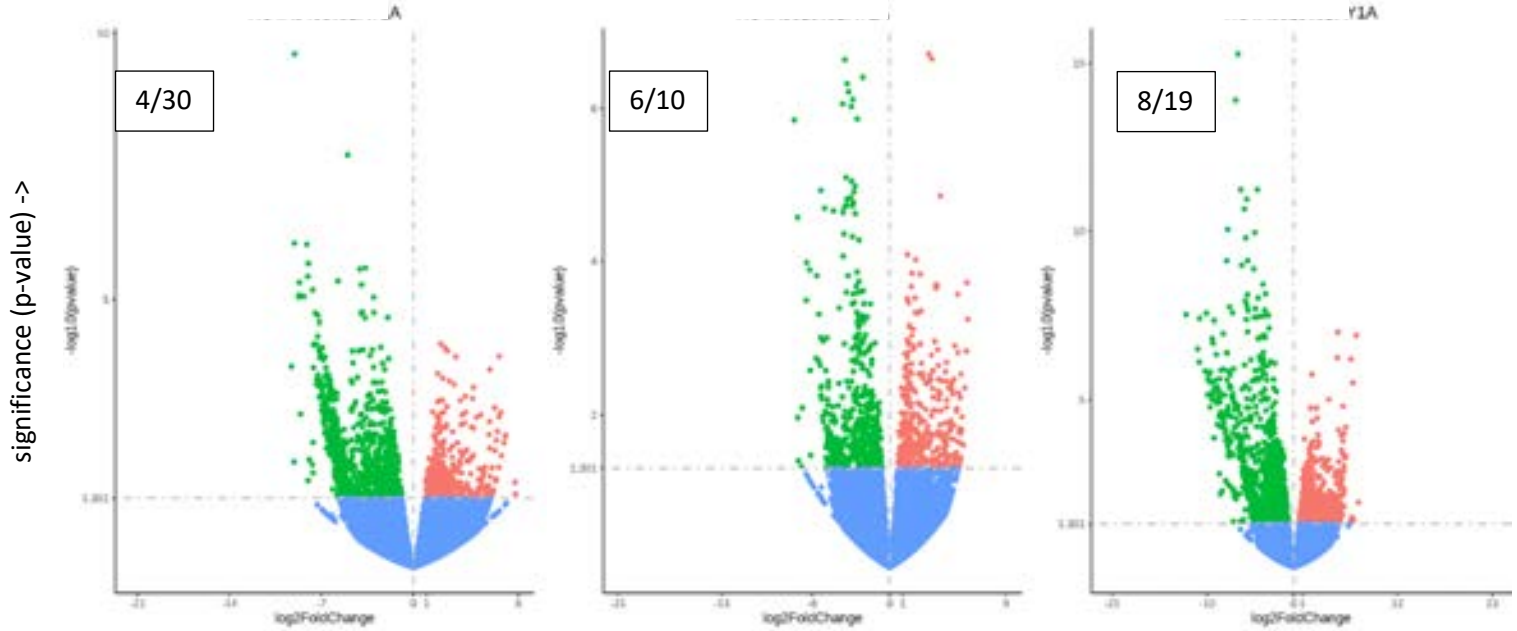


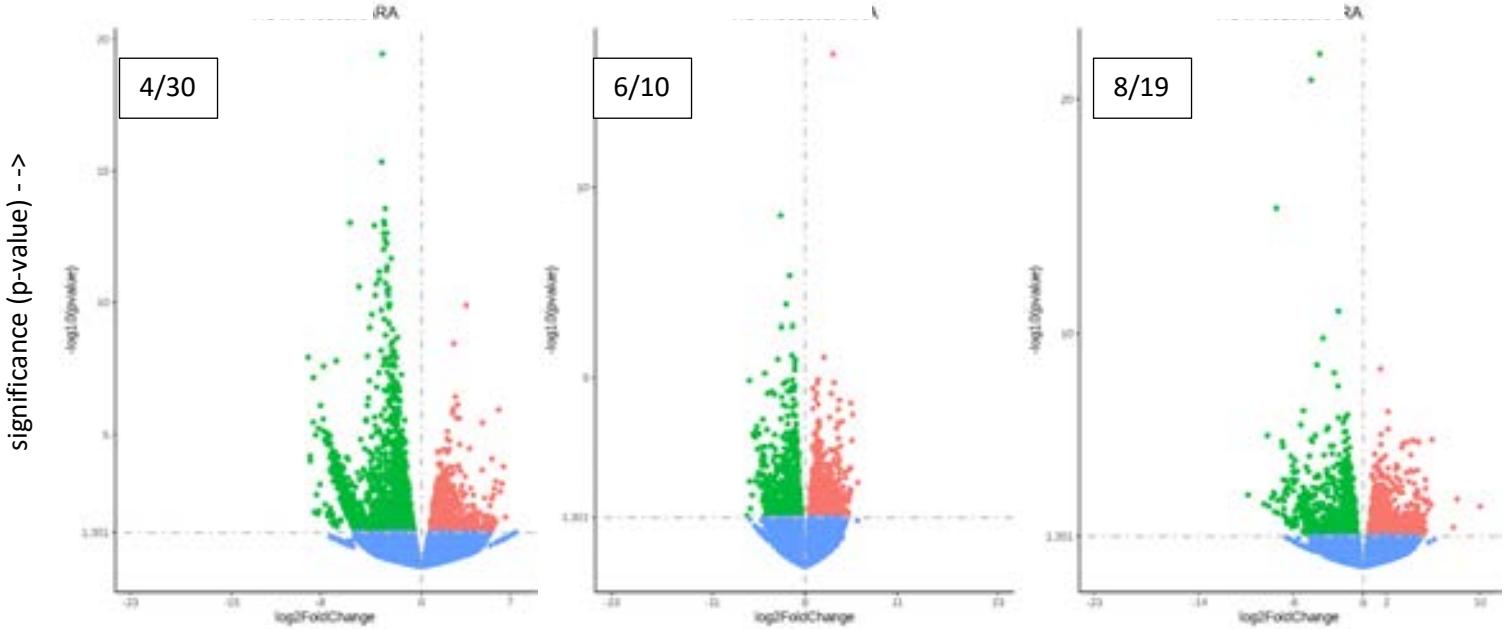
Figure 7. Dot plots of enriched KEGG pathways in mussels (*M. trossulus*) (adductor tissue) exposed to ANS crude oil compared with mussels from (A) unoiled control sites within the Valdez Fjord and (B) the Valdez boat harbor within the Valdez Fjord, showing beginning, mid, and ending time points. The size of the dot is based on gene count enriched in the pathway, and the color of the dot shows the pathway enrichment significance. KEGG pathway descriptions are listed on the vertical axis. The horizontal axis represents gene ratio (k/n) where k is the number of genes participating in the current KEGG pathway, and n is the number of genes annotated as participants of any KEGG pathway.

A) Spill site vs. Reference



	30-Apr	10-Jun	19-Aug
Up (red)	382	378	690
Down (green)	1,263	528	1,391
Non-significant (blue)	31,691	34,044	33,544

B) Spill site vs Harbor

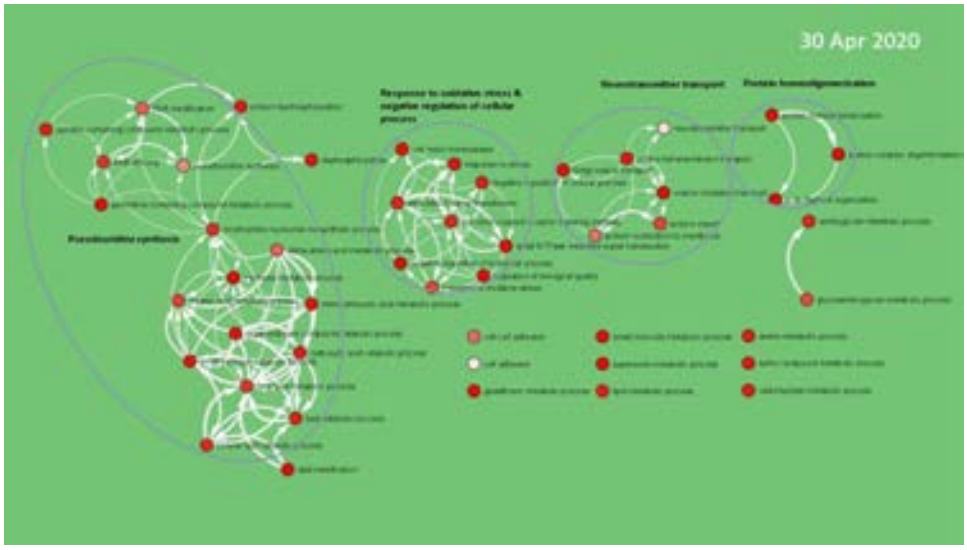


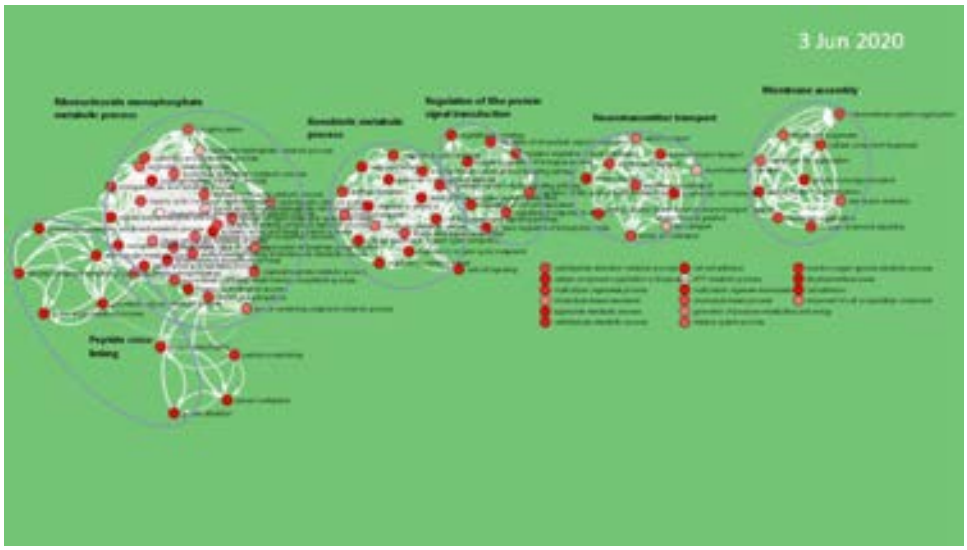
	30-Apr	10-Jun	19-Aug
Up (red)	829	665	791
Down (green)	2,354	797	977
Non-significant (blue)	30,989	34,260	34,644

Figure 8. Volcano plots of (A) spill site vs. unoiled reference samples and (B) spill site vs. harbor samples, showing up-regulated (green dots), down-regulated (red dots), and non-significant (blue dots) counts of differentially expressed genes from beginning, mid and end-point sampling dates. Differences in gene expression below p-values cutoff (0.05) were not significantly different from their respective comparison sites. Horizontal axes show magnitude of difference between comparisons (as log2 fold change), and the vertical axes show increasing significance (decreasing p-value, -log(p)). Gene numbers for each plot are presented in the tables below the plots.

Gene Ontology

Using a separate mega-database of gene functions (GO rather than KEGG), the GO assignments were assessed for significant differences and then plotted as functional cluster groups (Figure 9) with linkages indicated between gene functions. The groups are then somewhat subjectively titled with the highest common-linking clustering gene's label to identify cluster group functionality. Of note in this series is an increase over time in numbers and complexity of functional groupings (Figure 9).


30 Apr 2020


3 Jun 2020

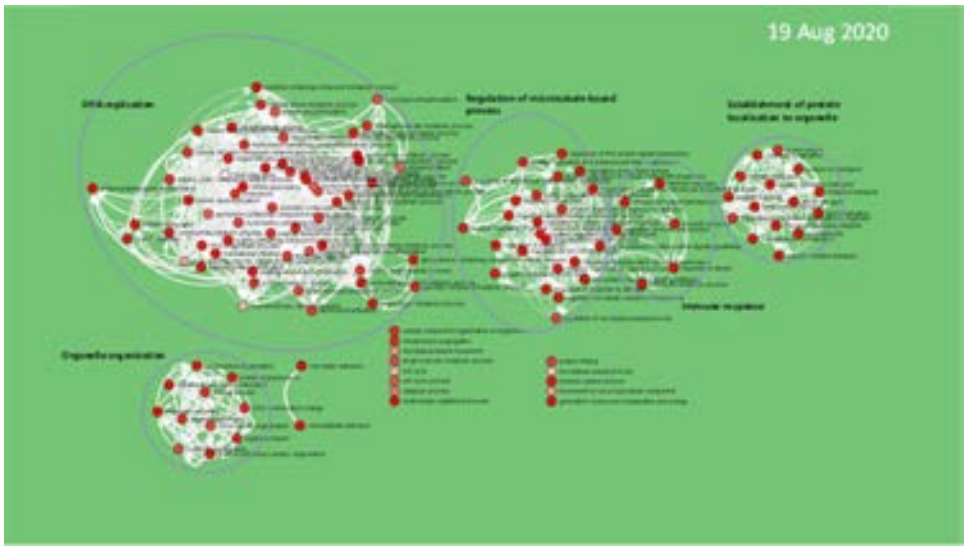

19 Aug 2020

Figure 9. Time series of GO pathway linkages. At the beginning, middle and end of the time series, genes that are expressed differently between mussels sampled at the spill site (HOTA) and mussels sampled at the unoiled control site (BAY1A) are clustered into groupings by their similar metabolic functions. White arrows link gene functions in their respective pathways. Listed single genes have no linkage. Over time, an increasing complexity of linkages is noted.

Spill site vs control

Genes representing many gene families were significantly differentially expressed between mussels from the spill site (HOTA) and mussels from the control site (Bay1A). Gene families particularly relevant to detoxification of hydrocarbons include cytochrome P450, aryl hydrocarbon receptor, and glutathione. Systems affected by initiation of detoxification include, but are not limited to, cytokines, neurotransmitters, and heat shock proteins. At most timepoints, there was a mixture of genes up- and down-regulated in no discernible pattern. From this group, we selected those genes which were significantly differentially expressed (Table 2) and that formed discernible trends of interest.

Table 2. Genes identified by differential expression analysis (DEG) of selected gene ontology functions for potential use in new transcript panel and formed discernible trends of interest. Listed genes were chosen (A) to distinguish between spill-site responses and unoiled reference sites (Jack and Galena Bays) and (B) to potentially differentiate between spill site and harbor (ANS crude oil versus pyrogenic and diesel components at harbor).

(A) Spill site vs Reference sites

Gene or gene family	Number of genes
ABC (ATP-binding cassette)	8
Glutathiones	18
Heat shock proteins	9

Helicase	26
Immune related	21
Kinesin	11
Meiosis	7
Neurotransmitter	9
Oxidative stress response	1
Cytochrome P450	7
RNA recognition motif	24
General stress	3
Superoxide dismutase	2
Tumor necrosis family	16
Ubiquitin	39

(B) Spill site vs Harbor

Gene or gene family	Number of genes
ABC (ATP-binding cassette)	13
Glutathiones	16
Heat shock, HSP, chaperone	21
Helicase	1
Immune related	35
Neurotransmitter	11
Cytochrome P450	7
RNA recognition motif	4
General stress	9

Superoxide dismutase	3
Tumor necrosis factor	11
Ubiquitin	49

Discussion

Overview

This project provides a limited snapshot of impacted Alaskan mussels depurating and recovering from a minor, single source, spring-to-summer oil spill with diminishing chronic exposures. While these data are accurate, they represent just one study; the degree that the observational results can be generalized to another event requires additional validation. But the data do show general patterns of oil weathering, depuration and gene transcription that provide multiple lines of evidence for oil exposure, uptake, metabolism, elimination, and impacts in a widely used biomonitoring organism. Further, the study supports a growing area of research that uses gene transcription as a sensitive biomonitoring endpoint for contaminant exposure and physiological responses.

The results support and expand on previously identified molecular responses to oil (Whitehead et al. 2012, Connon et al. 2019, Portnoy et al. 2020), and characterize the physiological response to crude oil from the initial exposure through detoxification and initial recovery, a period of approximately five months in this case. In general, exposure to ANS crude oil resulted in significant physiological responses in mussels, as indicated by the enriched biological processes and pathway activations throughout the study. These physiological responses persisted even as exposure diminished to relatively lower levels and tissue burdens of chemicals returned towards baseline.

Chemistry

From the spill timeline, the monotonically decreasing TPAH values and the profiles' weathering-consistent trends (Figure 5 and Figure SI2-1), it is reasonable to assume that oil concentrations in the mussels were higher in the 18-day interval between the initial spill and this study's inception. Thus, this project's data only address the diminishing exposure series but not the true initial exposure. Spill-responder accounts (SCAT reports) and our data both

suggest that by the end of sampling (Day 82), sheening exposure had not completely stopped nor had the spill-site mussels reached full recovery. The spill-site mussels were still above the concentrations seen at the adjacent LTEMP monitoring sites sampled either mid-study (Day 41) or from before or after the spill (in 2019 and 2021; Figure 2). Transcription values also show incomplete recovery (discussed below).

PAH and SHC weathering patterns were, as expected, based mostly on log K_{OW} properties of the individual PAH and preferential microbial consumption of SHC. From previous LTEMP collections, the expected local background/recovery point appears as a sparse suite of mixed dissolved and pyrogenic components, each in single digit, ppb concentrations detected at or below MDLs, and without petroleum biomarkers (Figure 3). Background data from the remote Jack Bay and Galena Bay sites also reflected the trace-level pyrogenic and marine/terrestrial profiles anticipated in these pristine Alaskan intertidal environments (Figure 3).

Transcription

The Venn diagram (Figure 6) identified genes that were differentially expressed among mussels from the three different sites, the spill zone (HOTA), negative control (Bay1A), and harbor (HARA). Transcription levels of 360 genes were unique to the spill zone, implying that some subset could be sufficient to differentiate between physiological responses induced by ANS crude oil vs hydrocarbon mixtures from a small boat harbor. However, in addition to the hydrocarbon differences between these habitats, other environmental differences potentially exist as well (e.g., nutrient sources, water chemistry, niche microhabitats), necessitating post-development validation of a new transcript panel.

From volcano plots (Figure 8), the comparison of spill site to unoiled reference samples (Table 3) shows a high count of down-regulated genes in the initial sample (1,263) and a similar upturn in the final sampling (1,391). Comparing spill site samples to harbor samples, an even larger count of down-regulated genes occurred in the initial sampling (2,354) while the final sampling plateaued, with no changes in up-regulated gene count. It is speculation to attribute this behavior to initial homeostatic disruption or to a shift to recovery phase in the last sampling. This form of summary data is insufficient for making conjectures, but definitely points to large and dynamic pathway shifts.

Table 3. Gene counts by up/down expression and significance in site comparisons (see Figure 8).

Spill vs unoiled reference	30-Apr	10-Jun	19-Aug
Up (red)	382	378	690
Down (green)	1,263	528	1,391
Non-signif (blue)	31,691	34,044	33,544
Spill vs harbor			
Up (red)	829	665	791
Down (green)	2,354	797	977
Non-signif (blue)	30,989	34,260	34,644

The molecular response to oil exposure obviously involves a dynamic host of genes and functional processes which are revealed in the time series shown in Figures 7 and 9. The Venn diagrams (Figure 6) and volcano plots (Figure 8), however, only provide information on counts of differentially expressed genes but not on biological processes. Due to the different chemical compositions of the spill site and harbor environments, we anticipated differences in respective enriched biological processes. While the spill-site/controls comparison primarily elicited differences in "molecular" and "cellular maintenance and function" processes, the spill-site/harbor comparison also identified a relatively strong enrichment in the "response to organic cyclic compound" process. Since both ANS crude oil and harbor contaminants contain a plethora of mostly unmeasured organic compounds, this difference is not explained by the available chemistry data.

Comparison of spill site and controls revealed, not surprisingly, gene responses to ANS crude oil in cellular systems associated with detoxification of polycyclic aromatic hydrocarbons. These systems include well-known targets of the ligand-activated AHR signaling pathway (cytochrome P450 and glutathione-S-transferase). The aryl hydrocarbon receptor (AHR) pathway is the major

regulatory pathway known in vertebrates as responsible for the molecular response to hydrocarbon exposure (Jenny et al. 2016). The primary AHR target genes, involved in the biotransformation of hydrocarbons into a water-soluble, more excretable product, include cytochrome P450 1A (a Phase 1 enzyme), and glutathione S-transferases (Phase II conjugating enzymes) (Nebert et al. 2000, Boutros et al. 2004). In this two-phase intracellular process, PAH bind with the AHR, which is then chaperoned by heatshock proteins as the AHR complex is translocated into the nucleus of the cell and eventually binds with the promoter region of cytochrome P450 genes (Zhou et al., 2010; Murray et al., 2014). The PAH are then chemically transformed, which facilitates excretion. Although not necessarily expected, the presence of both up and down-regulated genes at single timepoints reflects the complex and interconnected nature of molecular responses in general, and/or reflects later intermittent exposure episodes to sheening.

As well, genes involved in PAH detoxification in particular have been shown to be differentially regulated based on tissue type and contamination level (Chatel et al 2014). The minimal perceived involvement of the AHR itself (Table 2) is perplexing but may be an artifact of the limited genomic information for the study species (*Mytilus trossulus*). Alternatively, since expression of many genes is tissue specific, AHR may not be sufficiently expressed to be detected in adductor muscle tissue. Chatel et al. (2014) found AHR expression was 10 X lower in mussels from contaminated sites than from reference mussels after one month of exposure. Furthermore, although the activation of the AHR-CYP pathway helps to remove toxicants, some intermediate metabolites could be even more toxic than in their original form (Sun et al. 2020). If this was occurring at high contaminant loads, it may be mitigated by a negative feedback regulation via the suppression of AHR by an AHR repressor or the export or degradation of AHR (Sun et al. 2020).

Inflammatory signaling is becoming increasingly recognized as an important mechanism mediating the toxic effects of AHR agonists (Dubansky et al. 2013). In fact, the immune system of bivalves might be one of the main targets for PAHs (Wootton et al., 2003; von Moos et al., 2012; Canesi et al., 2015; Tang et al., 2020). Indirectly, PAHs may exert immunotoxic impacts by affecting physiological processes such as metabolism, energy supply, and neural-immune regulation (Sun et al. 2020). Bivalves such as mussels have well-developed nervous systems (Brenneis and Richter, 2010) and are able to regulate their immune response and maintain homeostasis in a manner similar to that of vertebrates through the coordination of neural and immune systems

via neurotransmitters and cytokines (Tang et al. 2020). The effects of PAHs on neurotransmitters in bivalves has been well-documented (Balbi et al. 2021, Dellali et al. 2021). PAHs can also induce intracellular reactive oxygen species (ROS) which disrupts immune-related molecular signaling pathways including the cytokine network, facilitating an inflammatory response (Guan et al., 2019; Zha et al., 2019; Tang et al., 2020; Dubansky et al. 2013). Cytokines were identified as significantly differentially expressed in seven of the 11 spill site samples in our study (Table 2). Altered immune function capabilities may increase health risks for organisms exposed to significant pathogen challenges, perhaps including some pathogens occurring with increasing frequency due to climate change. Immune function is critical to the health of marine bivalves; any disruption in immunity may pose a serious threat to individuals, populations, and ecosystems (Matozzo et al., 2012; Liu et al., 2016; Guan et al., 2018; Zha et al., 2019).

While we have focused here on the affected systems most directly related to detoxification of and recovery from oil exposure, we identified many other affected systems of interest. Other genes and pathways of interest include: dynenin, kinesin, RNA Recognition Motif, Helicase, Oxidative stress, Ubiquitin, TNFalpha, and ABC, to name a few.

Ultimately, energy availability in bivalves can be significantly reduced by exposure to pollutants (Shi et al. 2020). Continued gene transcription to mitigate stressors, e.g., detoxification, can thus be physiologically costly (Graham et al., 2010). Long-term chronic or high-concentration acute oil exposures may overwhelm metabolic pathways and/or simultaneously cause shifts in resources away from normal cell functions (Portnoy et al., 2020). Stress mitigation imposes extra demands, above those normally required to maintain homeostasis, which may reduce fitness. Fitness loss is usually evidenced by decreased growth and reproductive capability, increased susceptibility to disease, or disadvantageous behavioral changes (Graham et al., 2010; Martin et al., 2010). The individual, and hence the population, suffers.

While gene transcription provides relevant information on physiological status in exposed organisms, synergistic effects among PAH, other potential contaminants, and varying environmental stressors will complicate our interpretations of these physiological responses. For example, PAH exposure initiates gene transcription related to xenobiotic metabolism, but the resulting metabolic products may be toxic (Portnoy et al., 2020). Likewise, temperature, dissolved oxygen, salinity and tidal emersion all vary in the nearshore marine

environment and can influence physiological responses of organisms to chemical exposures (Whitehead, 2013). While the acute effects of oil toxicity are well known (Whitehead, 2013), sublethal effects that are difficult to identify may be critically important for predicting long-term, population-level impacts (Whitehead et al., 2012). Additionally, there are suggestions that chronically oiled populations will adapt (alter their epigenetic regulation) in response to continued stressors, a topic not evaluated in this project but relevant to differentiating gene responses of an acutely or chronically distressed population versus a control group.

Implications of enriched biological systems shifting over time

Gene ontology (GO) enrichment analysis of sets of transcriptional responses over time can offer insight into the molecular mechanisms that underpin biological responses. For example, when exposed to a stressor an organism will likely have a complex and non-linear transcriptional response. But at which point in the response curve have we sampled and how has this influenced our interpretations?

In general, transcription results show that following higher levels of tissue PAH burdens, significant physiological responses occurred and progressed as PAH burdens declined. However, at least at the molecular level, initial responses appeared to be dampened or lagged. The numbers of enriched biological systems change over time, increasing from our initial timepoint to our ending timepoint. The most probable interpretation was that the mussels had limited energy budgets and, as other physiological processes needed the same resources to respond to oil, the mussels were unable to immediately increase transcription of the detoxification genes evaluated in this study. This pattern of progression concurs with other studies that have detected similar delayed relationships between chemical exposures and transcription, and suggests overwhelmed or inhibited mitigating responses in organisms when initially exposed to high oil concentrations (Poynton and Vulpe, 2009; Pilcher et al., 2014). Pilcher et al. (2014) examined Gulf killifish (*Fundulus grandis*) exposed to water-accommodated fractions from weathered south Louisiana crude oil, and found that exposure to lower oil concentrations translated into a greater number of activated pathways, compared to exposure to higher concentrations.

Conclusions

This observational study has been a limited snapshot of impacted mussels recovering from a minor spring-season oil spill in Alaska while enduring diminishing chronic exposures. The chemistry and transcriptomic data showed general patterns of oil weathering, depuration and gene transcription that provide multiple lines of evidence for oil exposure, uptake, and gene pathway responses for intertidal mussels, a widely used biomonitoring organism. Furthermore, the study supports a growing research area using gene transcription for a sensitive biomonitoring perspective on contaminant exposures and their physiological responses as a proxy for impacts and recovery.

In this time series study, gene activities related to the detoxification and recovery processes were identified in multiple pathways. The gene activities and tissue PAH burdens did not return to reference-site levels, suggesting that recovery from hydrocarbons was not complete by final sampling. This may have been due to the presumed residual sheening.

Transcriptome differences among the three sites (spill site, harbor and unoiled controls) indicated that gene responses unique to oil exposure were identified in the spill site versus the unoiled site samples. Also, genes that could potentially distinguish between ANS crude oil and harbor contaminants (pyrogenics and diesel) were identified. These gene suites have yet to be validated under controlled conditions.

Acknowledgements

Funding for this research was provided by Prince William Sound Regional Citizens' Advisory Council (PWSRCAC) under Contract No. PWSRCAC Contract 951.22.07. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the PWSRCAC or the US Government. Mussel collections were under the Aquatic Resources Permit of Alaska Department of Fish and Game (CF-20-066).

References

Androutsopoulos VP, Tsatsakis AM, Spandidos DA. Cytochrome P450 CYP1A1: wider roles in cancer progression and prevention. BMC cancer. 2009 Dec;9(1):1-7.

Apeti DA, Johnson WE, Kimbrough KL, Lauenstein GG. National Status and Trends Mussel Watch Program: Sampling Methods 2012 Update.

Balbi T, Auguste M, Ciacci C, Canesi L. Immunological responses of marine bivalves to contaminant exposure: Contribution of the -omics approach. Frontiers in Immunology. 2021 Feb 18;12:618726.

Beyer J, Green NW, Brooks S, Allan IJ, Ruus A, Gomes T, Bråte IL, Schøyen M. Blue mussels (Mytilus edulis spp.) as sentinel organisms in coastal pollution monitoring: a review. Marine environmental research. 2017 Sep 1;130:338-65.

Bodkin JL, Esler D, Rice SD, Matkin CO, Ballachey BE. The effects of spilled oil on coastal ecosystems: lessons from the Exxon Valdez spill. Coastal conservation. 2014 Mar 27;19:311.

Bolognesi C, Cirillo S. Genotoxicity biomarkers in aquatic bioindicators. Current Zoology. 2014 Apr 1;60(2):273-84.

Boutros PC, Moffat ID, Franc MA, Tijet N, Tuomisto J, Pohjanvirta R, Okey AB. Dioxin-responsive AHRE-II gene battery: identification by phylogenetic footprinting. Biochemical and biophysical research communications. 2004 Aug 27;321(3):707-15.

Bowen L, Miles AK, Ballachey B, Waters S, Bodkin J, Lindeberg M, Esler D. Gene transcription patterns in response to low level petroleum contaminants in *Mytilus trossulus* from field sites and harbors in southcentral Alaska. Deep Sea Research Part II: Topical Studies in Oceanography. 2018 Jan 1;147:27-35.

Bowen L, Counihan KL, Ballachey B, Coletti H, Hollmen T, Pister B, Wilson TL. Monitoring nearshore ecosystem health using Pacific razor clams (*Siliqua patula*) as an indicator species. PeerJ. 2020 Mar 5;8:e8761.

Brenneis G, Richter S. Architecture of the nervous system in *Mystacocarida* (Arthropoda, Crustacea)—an immunohistochemical study and 3D reconstruction. Journal of morphology. 2010 Feb;271(2):169-89.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J. The MIQE Guidelines: M inimum I nformation for Publication of Quantitative Real-Time PCR Experiments.

Canesi L, Ciacci C, Bergami E, Monopoli MP, Dawson KA, Papa S, Canonico B, Corsi I. Evidence for immunomodulation and apoptotic processes induced by cationic polystyrene nanoparticles in the hemocytes of the marine bivalve Mytilus. Marine environmental research. 2015 Oct 1;111:34-40.

Carella F, Aceto S, Mangoni O, Mollica MP, Cavaliere G, Trinchese G, Aniello F, De Vico G. Assessment of the health status of mussels *Mytilus galloprovincialis* along the Campania coastal areas: a multidisciplinary approach. Frontiers in Physiology. 2018 Jun 12;9:683.

Châtel A, Faucet-Marquis V, Pfohl-Leszkowicz A, Gourlay-Francé C, Vincent-Hubert F. DNA adduct formation and induction of detoxification mechanisms in Dreissena polymorpha exposed to nitro-PAHs. Mutagenesis. 2014 Nov 1;29(6):457-65.

Connon RE, Hasenbein S, Brander SM, Poynton HC, Holland EB, Schlenk D, Orlando JL, Hladik ML, Collier TK, Scholz NL, Incardona JP. Review of and recommendations for monitoring contaminants and their effects in the San Francisco Bay− Delta. San Francisco Estuary and Watershed Science. 2019;17(4).

Counihan KL, Bowen L, Ballachey B, Coletti H, Hollmen T, Pister B, Wilson TL. Physiological and gene transcription assays to assess responses of mussels to environmental changes. PeerJ. 2019 Oct 4;7:e7800.

Dellali M, Hedfi A, Ali MB, Noureldeen A, Darwish H, Beyrem H, Gyedu-Ababio T, Dervishi A, Karachle PK, Boufahja F. Multi-biomarker approach in Mytilus galloprovincialis and Ruditapes decussatus as a predictor of pelago-benthic responses after exposure to Benzo [a] Pyrene. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology. 2021 Nov 1;249:109141.

Dubansky B, Whitehead A, Miller JT, Rice CD, Galvez F. Multitissue molecular, genomic, and developmental effects of the Deepwater Horizon oil spill on resident Gulf killifish (*Fundulus grandis*). Environmental Science & Technology. 2013 May 21;47(10):5074-82.

Farr S, Dunn 2nd RT. Concise review: gene expression applied to toxicology. Toxicological sciences: an official journal of the Society of Toxicology. 1999 Jul 1;50(1):1-9.

Graham AL, Shuker DM, Pollitt LC, Auld SK, Wilson AJ, Little TJ. Fitness consequences of immune responses: strengthening the empirical framework for ecoimmunology. Functional ecology. 2011 Feb;25(1):5-17.

Guan X, Tang Y, Zha S, Han Y, Shi W, Ren P, Yan M, Pan Q, Hu Y, Fang J, Zhang J. Exogenous Ca2+ mitigates the toxic effects of TiO2 nanoparticles on phagocytosis, cell viability, and apoptosis in haemocytes of a marine bivalve mollusk, *Tegillarca granosa*. Environmental Pollution. 2019 Sep 1;252:1764-71.

Hylland K. Polycyclic aromatic hydrocarbon (PAH) ecotoxicology in marine ecosystems. Journal of Toxicology and Environmental Health, Part A. 2006 Jan 1;69(1-2):109-23.

Incardona JP, Gardner LD, Linbo TL, Brown TL, Esbaugh AJ, Mager EM, Stieglitz JD, French BL, Labenia JS, Laetz CA, Tagal M. Deepwater Horizon crude oil impacts the developing hearts of large predatory pelagic fish. Proceedings of the National Academy of Sciences. 2014 Apr 15;111(15):E1510-8.

Jenny MJ, Walton WC, Payton SL, Powers JM, Findlay RH, O'Shields B, Diggins K, Pinkerton M, Porter D, Crane DM, Tapley J. Transcriptomic evaluation of the American oyster, *Crassostrea virginica*, deployed during the Deepwater Horizon oil spill: evidence of an active hydrocarbon response pathway. Marine environmental research. 2016 Sep 1;120:166-81.

Lauenstein, G.G., 1993. *Sampling and Analytical Methods of the National Status and Trends Program National Benthic Surveillance and Mussel Watch Projects, 1984-1992: Comprehensive descriptions of trace organic analytical methods* (Vol. 71). US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Office of Ocean Resources Conservation and Assessment, Coastal Monitoring and Bioeffects Assessment Division.

Liu S, Shi W, Guo C, Zhao X, Han Y, Peng C, Chai X, Liu G. Ocean acidification weakens the immune response of blood clam through hampering the NF-kappa β and toll-like receptor pathways. Fish & shellfish immunology. 2016 Jul 1;54:322-7.

Livingstone DR, Chipman JK, Lowe DM, Minier C, Pipe RK. Development of biomarkers to detect the effects of organic pollution on aquatic invertebrates: recent molecular, genotoxic, cellular and immunological studies on the

common mussel (*Mytilus edulis* L.) and other mytilids. International Journal of Environment and Pollution. 2000 Jan 1;13(1-6):56-91.

Marris CR, Kompella SN, Miller MR, Incardona JP, Brette F, Hancox JC, Sørhus E, Shiels HA. Polyaromatic hydrocarbons in pollution: a heart-breaking matter. The Journal of physiology. 2020 Jan;598(2):227-47.

Martin LB, Hopkins WA, Mydlarz LD, Rohr JR. The effects of anthropogenic global changes on immune functions and disease resistance. Annals of the New York Academy of Sciences. 2010 May;1195(1):129-48.

Matozzo V, Chinellato A, Munari M, Finos L, Bressan M, Marin MG. First evidence of immunomodulation in bivalves under seawater acidification and increased temperature. PloS one. 2012 Mar 27;7(3):e33820.

McLoughlin K, Turteltaub K, Bankaitis-Davis D, Gerren R, Siconolfi L, Storm K, Cheronis J, Trollinger D, Macejak D, Tryon V, Bevilacqua M. Limited dynamic range of immune response gene expression observed in healthy blood donors using RT-PCR. Molecular Medicine. 2006 Jul;12(7):185-95.

Murray IA, Patterson AD, Perdew GH. Aryl hydrocarbon receptor ligands in cancer: friend and foe. Nature Reviews Cancer. 2014 Dec;14(12):801-14.

Nebert DW, Roe AL, Dieter MZ, Solis WA, Yang YI, Dalton TP. Role of the aromatic hydrocarbon receptor and [Ah] gene battery in the oxidative stress response, cell cycle control, and apoptosis. Biochemical pharmacology. 2000 Jan 1;59(1):65-85.

National Research Council. Oil in the sea: inputs, fates, and effects. National Academies Press; 1985.

Payne JR, Driskell WB, Short JW, Larsen ML. Long term monitoring for oil in the Exxon Valdez spill region. Marine Pollution Bulletin. 2008 Dec 1;56(12):2067-81.

Payne JR, Driskell WB, Janka D, Ka'aihue L, Banta J, Love A, Litman E. EVOS and the Prince William Sound Long Term Environmental Monitoring Program. In International Oil Spill Conference 2021 May (Vol. 2021, No. 1, p. 688991).

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature biotechnology. 2015 Mar;33(3):290-5.

Peterson CH, Rice SD, Short JW, Esler D, Bodkin JL, Ballachey BE, Irons DB. Long-term ecosystem response to the Exxon Valdez oil spill. Science. 2003 Dec 19;302(5653):2082-6.

Pilcher W, Miles S, Tang S, Mayer G, Whitehead A. Genomic and genotoxic responses to controlled weathered-oil exposures confirm and extend field studies on impacts of the Deepwater Horizon oil spill on native killifish. PLoS One. 2014 Sep 10;9(9):e106351.

Portnoy DS, Fields AT, Greer JB, Schlenk D. Genetics and oil: Transcriptomics, epigenetics, and population genomics as tools to understand animal responses to exposure across different time scales. Deep Oil Spills. 2020:515-32.

Poynton HC, Vulpe CD. Ecotoxicogenomics: emerging technologies for emerging contaminants 1. JAWRA Journal of the American Water Resources Association. 2009 Feb;45(1):83-96.

Schlenk D, Handy R, Steinert S, Depledge MH, Benson W (2008) Biotransformation in fishes. In: Di Giulio R, Hinton D (eds) The toxicology of fishes, vol 1. CRC Press, Boca Raton, FL, pp 153–234

Sforzini S, Oliveri C, Orrù A, Chessa G, Pacchioni B, Millino C, Jha AN, Viarengo A, Banni M. Application of a new targeted low density microarray and conventional biomarkers to evaluate the health status of marine mussels: A field study in Sardinian coast, Italy. Science of the Total Environment. 2018 Jul 1;628:319-28.

Shi W, Han Y, Sun S, Tang Y, Zhou W, Du X, Liu G. Immunotoxicities of microplastics and sertraline, alone and in combination, to a bivalve species: size-dependent interaction and potential toxication mechanism. Journal of hazardous materials. 2020 Sep 5;396:122603.

Short JW, Springman KR. Identification of hydrocarbons in biological samples for source determination. In: Standard Handbook Oil Spill Environmental Forensics 2016 Jan 1 (pp. 1039-1069). Academic Press.

Stout SA, Wang Z. Chemical fingerprinting methods and factors affecting petroleum fingerprints in the environment. In: Standard handbook oil spill environmental forensics 2016 Jan 1 (pp. 61-129). Academic Press.

Sun M, Ma N, He T, Johnston LJ, Ma X. Tryptophan (Trp) modulates gut homeostasis via aryl hydrocarbon receptor (AhR). Critical reviews in food science and nutrition. 2020 May 30;60(10):1760-8.

Tang Y, Zhou W, Sun S, Du X, Han Y, Shi W, Liu G. Immunotoxicity and neurotoxicity of bisphenol A and microplastics alone or in combination to a bivalve species, *Tegillarca granosa*. Environmental Pollution. 2020 Oct 1;265:115115.

Von Moos N, Burkhardt-Holm P, Köhler A. Uptake and effects of microplastics on cells and tissue of the blue mussel *Mytilus edulis* L. after an experimental exposure. Environmental science & technology. 2012 Oct 16;46(20):11327-35.

Whitehead A, Dubansky B, Bodinier C, Garcia TI, Miles S, Pilley C, Raghunathan V, Roach JL, Walker N, Walter RB, Rice CD. Genomic and physiological footprint of the Deepwater Horizon oil spill on resident marsh fishes. Proceedings of the National Academy of Sciences. 2012 Dec 11;109(50):20298-302.

Whitehead A. Interactions between oil-spill pollutants and natural stressors can compound ecotoxicological effects. Integrative and comparative biology. 2013 Oct 1;53(4):635-47.

Wills LP, Zhu S, Willett KL, Di Giulio RT. Effect of CYP1A inhibition on the biotransformation of benzo [a] pyrene in two populations of *Fundulus heteroclitus* with different exposure histories. Aquatic toxicology. 2009 May 5;92(3):195-201.

Wootton EC, Dyrynda EA, Pipe RK, Ratcliffe NA. Comparisons of PAH-induced immunomodulation in three bivalve molluscs. Aquatic toxicology. 2003 Oct 8;65(1):13-25.

Zanette J, Jenny MJ, Goldstone JV, Parente T, Woodin BR, Bainy AC, Stegeman JJ. Identification and expression of multiple CYP1-like and CYP3-like genes in the bivalve mollusk *Mytilus edulis*. Aquatic toxicology. 2013 Mar 15;128:101-12.

Zha S, Shi W, Su W, Guan X, Liu G. Exposure to 2, 3, 7, 8-tetrachlorodibenzo-paradioxin (TCDD) hampers the host defense capability of a bivalve species, *Tegillarca granosa*. Fish & Shellfish Immunology. 2019 Mar 1;86:368-73.

Zhou H, Wu H, Liao C, Diao X, Zhen J, Chen L, Xue Q. Toxicology mechanism of the persistent organic pollutants (POPs) in fish through AhR pathway. Toxicology mechanisms and methods. 2010 Jul 1;20(6):279-86.

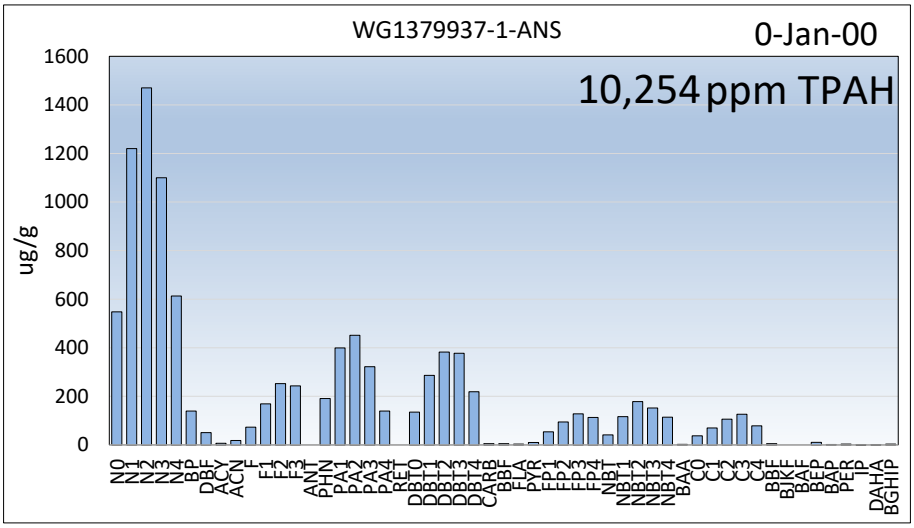
Supplemental Information

SI-1 Appendix 1 Hydrocarbon Analytes

Polyaromatic Hydrocarbons (PAH)

Analytes	Abbreviation
Naphthalene	N
C1-Naphthalene	N1
C2-Naphthalene	N2
C3-Naphthalene	N3
C4-Naphthalene	N4
Biphenyl	BI
Acenaphthylene	ACY
Acenaphthene	ACN
Fluorene	F
C1-Fluorene	F1
C2-Fluorene	F2
C3-Fluorene	F3
C4-Fluorene	F4
Anthracene	A
Phenanthrene	Ph
C1-	PA1
C2-	PA2
C3-	PA3
C4-	PA4
Retene	RET
Dibenzothiophene	DBT
C1-Dibenzothiophene	DBT1
C2-Dibenzothiophene	DBT2
C3-Dibenzothiophene	DBT3
C4-Dibenzothiophene	DBT4
Benzo(b)fluorene	BF

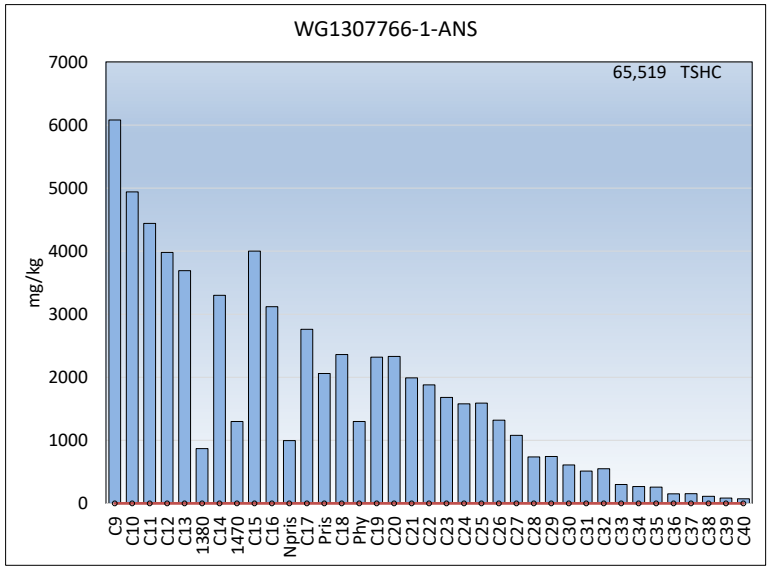
Analytes	Abbreviation
Fluoranthene	FL
Pyrene	PY
C1-	FP1
C2-	FP2
C3-	FP3
C4-	FP4
Napthobenzothiophene	NBT
C1-	NBT1
C2-	NBT2
C3-	NBT3
C4-	NBT4
Benzo(a)Anthracene	BAA
Chrysene	C
C1-Chrysene	C1
C2-Chrysene	C2
C3-Chrysene	C3
C4-Chrysene	C4
Benzo(b)fluoranthene	BBF
Benzo(k)fluoranthene	BKF
Benzo(e)pyrene	BEP
Benzo(a)pyrene	BAP
Perylene	PER
Indeno(1,2,3-cd)pyrene	IND
Dibenzo(a,h)anthracene	DAHA
Benzo(g,h,i)perylene	BGH
Total PAH	TPAH



ANS Crude oil example

Saturated hydrocarbons (SHC or n-alkanes)

Analyte	Abbrev
Nonane (C9)	C9
Decane (C10)	C10
Undecane (C11)	C11
Dodecane (C12)	C12
Tridecane (C13)	C13
2,6,10 Trimethyldodecane (1380)	1380
Tetradecane (C14)	C14
2,6,10-Trimethyltridecane (1470)	1470
Pentadecane (C15)	C15
Hexadecane (C16)	C16
Norpristane (1650)	Pristane
Heptadecane (C17)	C17
Pristane	Phytane
Octadecane (C18)	C18
Phytane	Phy
Nonadecane (C19)	C19
Eicosane (C20)	C20
Heneicosane (C21)	C21
Docosane (C22)	C22
Tricosane (C23)	C23
Tetracosane (C24)	C24
Pentacosane (C25)	C25
Hexacosane (C26)	C26
Heptacosane (C27)	C27
Octacosane (C28)	C28



Nonacosane (C29)	C29
Triacontane (C30)	C30
Hentriacontane (C31)	C31
Dotriacontane (C32)	C32
Tritriacontane (C33)	C33
Tetratriacontane (C34)	C34
Pentatriacontane (C35)	C35
Hexatriacontane (C36)	C36
Heptatriacontane (C37)	C37
Octatriacontane (C38)	C38
Nonatriacontane (C39)	C39
Tetracontane (C40)	C40
Total SHC	TSHC

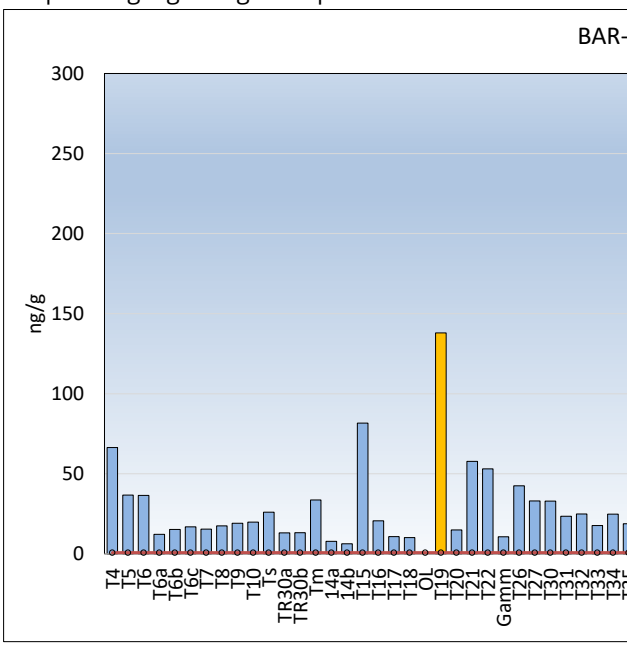
Petroleum Biomarkers (S/T)

Class	Biomarker	Abbrev
Terpanes	C23 Tricyclic Terpane (T4)	T4
	C24 Tricyclic Terpane (T5)	T5
	C25 Tricyclic Terpane (T6)	T6
	C24 Tetracyclic Terpane	T6a
	C26 Tricyclic Terpane-22S	T6b
	C26 Tricyclic Terpane-22R	T6c
	C28 Tricyclic Terpane-22S	T7
	C28 Tricyclic Terpane-22R	T8
	C29 Tricyclic Terpane-22S	T9
	C29 Tricyclic Terpane-22R	T10
	18a-22,29,30-	Ts
	C30 Tricyclic Terpane-22S	C30Ts
	C30 Tricyclic Terpane-22R	C30Tr
Hopanes	17a(H)-22,29,30-	Tm
	17a/b,21b/a 28,30-	14a
	17a(H),21b(H)-25-	14b
	30-Norhopane (T15)	T15
	18a(H)-30-Norneohopane-	T16
	17a(H)-Diahopane (X)	X
	30-Normoretane (T17)	T17
	18a(H)&18b(H)-Oleananes	T18
	Hopane (T19)	T19*
	Moretane (T20)	T20
	30-Homohopane-22S (T21)	T21
	30-Homohopane-22R	T22
	Gammacerane/C32-	T22a
	30,31-Bishomohopane-	T26
	30,31-Bishomohopane-	T27
	30,31-Trishomohopane-	T30
	30,31-Trishomohopane-	T31
	Tetrakishomohopane-22S	T32
	Tetrakishomohopane-22R	T33
	Pentakishomohopane-22S	T34
	Pentakishomohopane-22R (T35)	T35
Steranes	13b(H),17a(H)-20S-Diacholestane (S4)	S4
	13b(H),17a(H)-20R-Diacholestane (S5)	S5

Class	Biomarker	Abbrev
	13b,17a-20S-Methyldiacholestane (S8)	S8
	17a(H)20SC27/C29dia	DIA29S
	17a(H)20rc27/C29dia	DIA29R
	Unknown Sterane (S18)	S18
	13a,17b-20S-Ethyldiacholestane (S19)	S19
	14a,17a-20S-Methylcholestane (S20)	S20
	14a,17a-20R-Methylcholestane (S24)	S24
	14a(H),17a(H)-20S-Ethylcholestane (S25)	S25
	14a(H),17a(H)-20R-Ethylcholestane (S28)	S28
	14b(H),17b(H)-20R-Cholestane (S14)	S14
	14b(H),17b(H)-20S-Cholestane (S15)	S15
	14b,17b-20R-Methylcholestane (S22)	S22
	14b,17b-20S-Methylcholestane (S23)	S23
	14b(H),17b(H)-20R-Ethylcholestane (S26)	S26
	14b(H),17b(H)-20S-Ethylcholestane (S27)	S27
	C20 Pregnane	Preg
	C21 20-Methylpregnane	MPreg
	C22 20-Ethylpregnane (a)	EPregA
	C22 20-Ethylpregnane (b)	EPregB
Triaromatic Steroids	C26,20S TAS	TAS0
	C26,20R+C27,20S TAS	TAS1
	C28,20S TAS	TAS2
	C27,20R TAS	TAS3
	C28,20R TAS	TAS4
	C29,20S TAS	TAS5
	C29,20R TAS	TAS6

Mono-aromatic Steroids	5b(H)-C27 (20S) MAS+	MAS1
	5b(H)-C27 (20R) MAS+	MAS2
Class	**Biomarker**	**Abbrev**
	5a(H)-C27 (20S) MAS	MAS3
	5b(H)-C28 (20S) MAS+	MAS4
	5a(H)-C27 (20R) MAS	MAS5
	5a(H)-C28 (20S) MAS	MAS6
	5b(H)-C28 (20R) MAS+	MAS7
	5b(H)-C29 (20S) MAS+	MAS8
	5a(H)-C29 (20S) MAS	MAS9
	5a(H)-C28 (20R) MAS	MAS10
	5b(H)-C29 (20R) MAS+	MAS11
	5a(H)-C29 (20R) MAS	MAS12

*hopane highlighted gold in plots



ANS spill oil from 2017

SI-2 Appendix 2 Depuration Kinetics

Depuration

The biomarker hopane, 17β(H),21β(H)-hopane (abbreviated as T19 in plots), is considered to be non-water soluble and mostly non-biodegradable and thus is among the best conserved of the petroleum components. Its non-biodegradable nature and the declining logarithmic shape of its depuration curve (Figure SI2-1, left panel) implies that simple elimination (gut-purging) dominates versus a mix of elimination, metabolic processing and/or weathering. For the latter processes, the depuration curve would be expected to flatten or increase (bioaccumulate). A similar assessment was made for the 43-analyte PAH sum (TPAH43) (Figure SI2-1, right panel). Note that the TPAH43 index combines a wide range of solubilities with octanol/water partitioning coefficients ranging from 3-7 (log K_{OW}), which implies that each analyte is endowed with unique diffusive-desorption and gut-elimination rates. The resulting TPAH43 depuration curve represents a composite of those effects. Both hopane and TPAH43 depletions closely fit logarithmic-regression models (R^2=0.96 for each). From these models, the half-life for hopane was calculated as 8.4 days while TPAH43 was 9.2 days, remarkably close values considering that certain members in the TPAH suite can bioaccumulate.

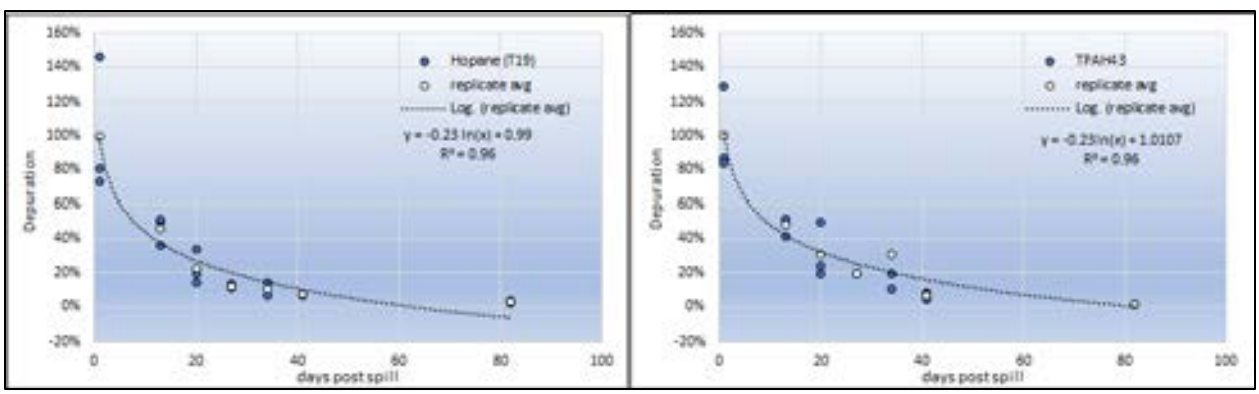


Figure . SI-2-1. Depletion plots of hopane and PAH composite, TPAH43, in whole mussel tissues from the spill site. Both show simple depletion without bioaccumulation. Half-lives are 8.4 and 9.2 days, respectively. Solid dots are single replicates; white dots are averages.

A few PAH examples of forensic interest show their varied rates of depletion (Figure SI2-2). In this study, most of the individual PAH present as various rates of simple depletion, by either log or exponential decay, and without accumulation, i.e., no initially increasing trends. However, there were exceptions. Higher-molecular-weight and higher-log- K_{OW} PAH showed either accumulation or a more complex pattern (Figure SI2-3). As discussed below and in SI Appendix 5, there appears to be a transition from simple depletion mode into accumulation or mixed response modes for compounds with log K_{OW} > ~ 6.3.

In contrast, petroleum biomarkers plotted relative to hopane showed mostly flat lines with little depletion (Figure SI2-3). This suggests that similar to hopane (Figure SI2-2), the petroleum biomarkers in mussels undergo a simple depletion process (gut purging) without bioaccumulation. Depletion curves for all PAH are presented in the SI Appendix 3.

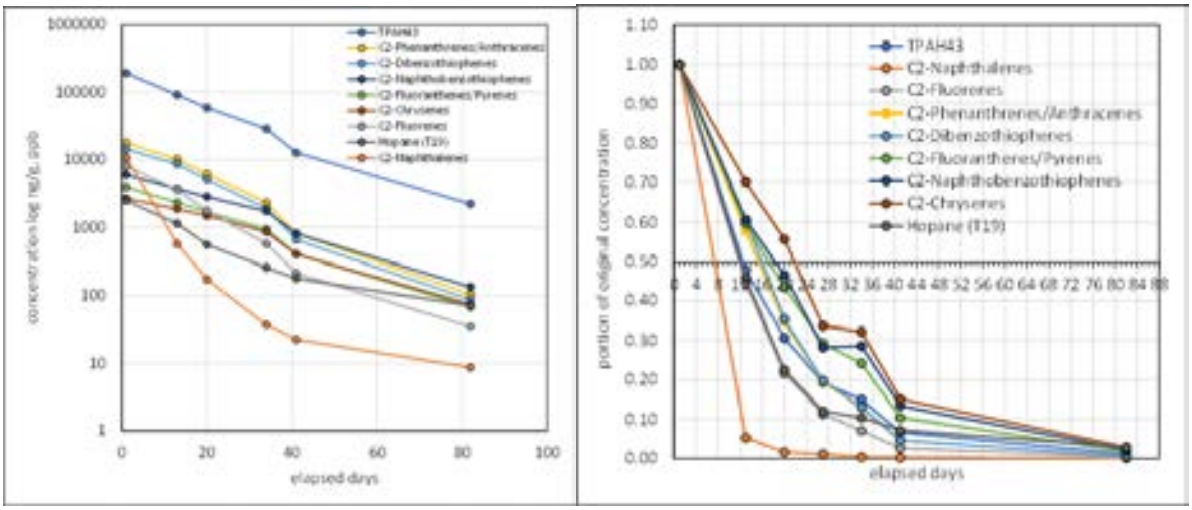


Figure SI2-2. Time series of measured and proportional concentrations (left & right plots respectively) for select PAH of forensic interest in spill site mussels. All show log or exponential depletion. Proportional plot (right) estimates half-life (days) at 0.5 axis crossing.

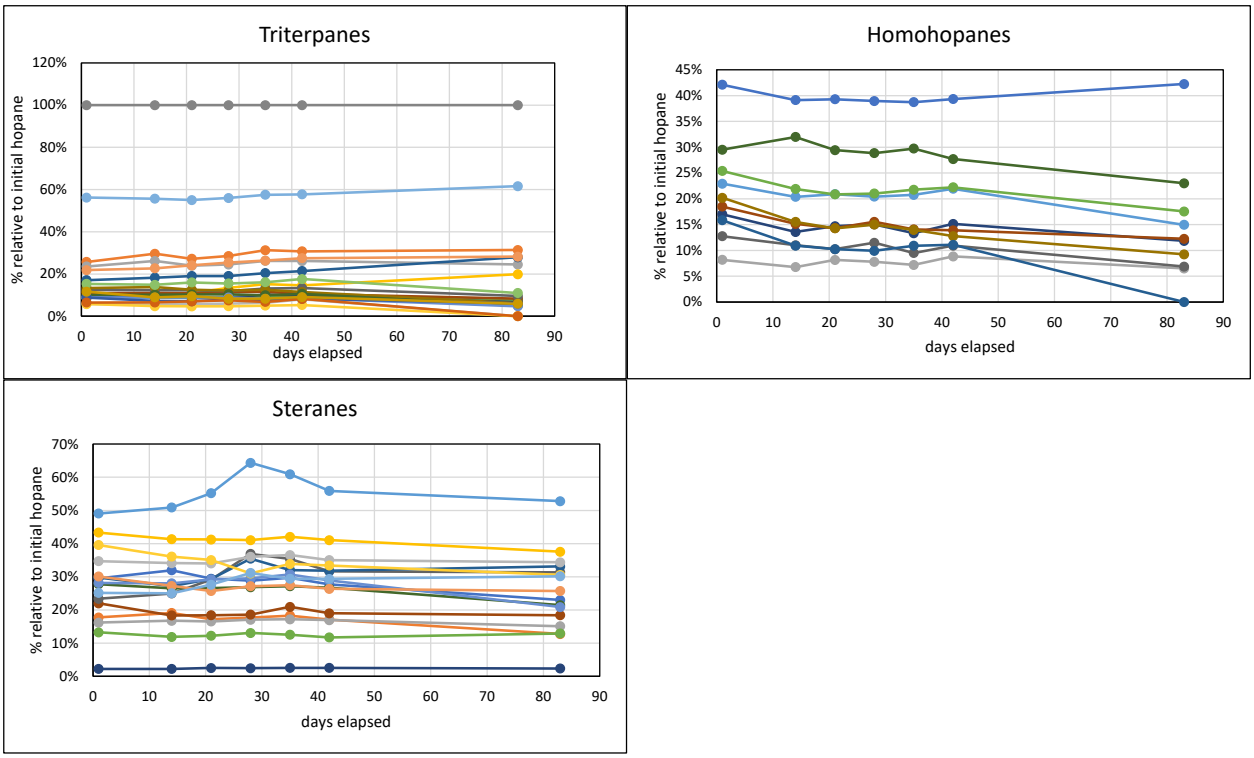


Figure SI2-3. Time series of petroleum biomarkers relative to hopane, showing essentially flat or slightly decreasing concentrations.

Depuration Rates

As mentioned above, from depuration calculations, the half-life for the hopane biomarker in mussel tissue was 8.4 days (Figure SI2-1). This petroleum biomarker is non-biodegradable, essentially not water-soluble (est. log K_{OW} 10-15), and displayed a simple log-decay depletion curve. These characteristics imply it is ingested and eliminated without bioaccumulating, behaving as a simple gut-content tracer (discussed further in SI3). However, the PAH components showed more complexity. PAH that are more lipophilic (higher molecular weights and higher log K_{OW}) are both more easily transported/diffused into body tissues from the gut tract and tend to bioaccumulate. In contrast, the more water-soluble PAH (lower log K_{OW}) tend to absorb as dissolved compounds via the siphon, mantle, and gills and quickly deplete. Thus, the PAH have different depletion curves whereby they appear to either simply deplete, accumulate in the body tissues, or show some complex mixed-response anomaly (Figure SI3-1).

Seen in PAH depletion plots grouped by the homologous analyte families (parent and alkylated homologs, e.g., fluorene, C1-fluorene, C2-fluorene, and C3-fluorene [F, F1, F2, F3]), and normalized relative to the most recalcitrant PAH, C2-chrysene (C2), the depletion rate slows. They are more persistent with degree of alkylation and increase in log K_{OW} partitioning values (Figure SI3-1, 4th panel on left). Then a transition occurs. In this study, PAH-depletion curves switch to an accumulation modality that initially appears with the mid-weight, three-ringed PAH, phenanthrene. Here, C4-phenanthrene (PA4) homologue concentration increases in proportion to the sample's C2 over time (Figure SI3-1, 5th panel on left). Seemingly, PAH with log K_{OW} > ~6.3 tended to show bioaccumulation in these mussels (discussed in SI; Table SI3-1). In contrast, the four-ringed chrysenes are so recalcitrant, it was difficult to characterize the nearly flat depletion curves in the 83-day time series.

There is other disparate behavior in the depletion curves where depletion trends are unexpectedly slower or even flat given the component's log K_{OW}, e.g., C3-fluorene, and all five- and six-ringed, primarily pyrogenic, PAH (Figure SI3-1). The latter are mostly absent in the source oil, so their appearance plus the depletion anomalies suggest that the mussels are chronically re-exposed throughout this study. SCAT reports and communications with agency responders corroborate that light sheening was still present on Day 83.

SI-3 Appendix 3. Tissue Depuration

As mentioned in SI-2, viewed by their homologous groupings (Figure SI3-1), depuration rates slowed with increasing alkylation (parent, C1-, C2-, C3-, C4-homologs) and increasing molecular weight (left column then right column). Accumulation patterns first appear in C4-phenanthrene. Decalins, a non-aromatic cyclic compound (not a PAH), are accumulating from an unknown secondary source (top left panel).

Viewed by log K_{OW} values (Table SI3-1), the patterns just described appear related to log K_{OW} values. Sorted by homologous families, note the log K_{OW} values increase with alkylation (left columns). Sorted by log K_{OW} (right columns), the spark-line plots appear as unit-less, sequential (not time-scaled) trend lines. Note that accumulating (increasing) trends do not appear before log K_{OW} ~6.3. Uniquely anomalous are several PAH, like acenaphthene, fluoranthene, C3-fluoranthene/pyrene and benzo[b]fluoranthene with a non-trending, concentration bump midway through the time series. The source of the patterns is uncertain. Also, several of the trend lines have an uptick to end the time series. These are considered lab artifacts (flagged as estimated values) whereby detection was confirmed but below calibration limits. Also, several analytes flat-line as non-detects. Yet behind this noisy data, log K_{OW} -related depuration trends are apparent.

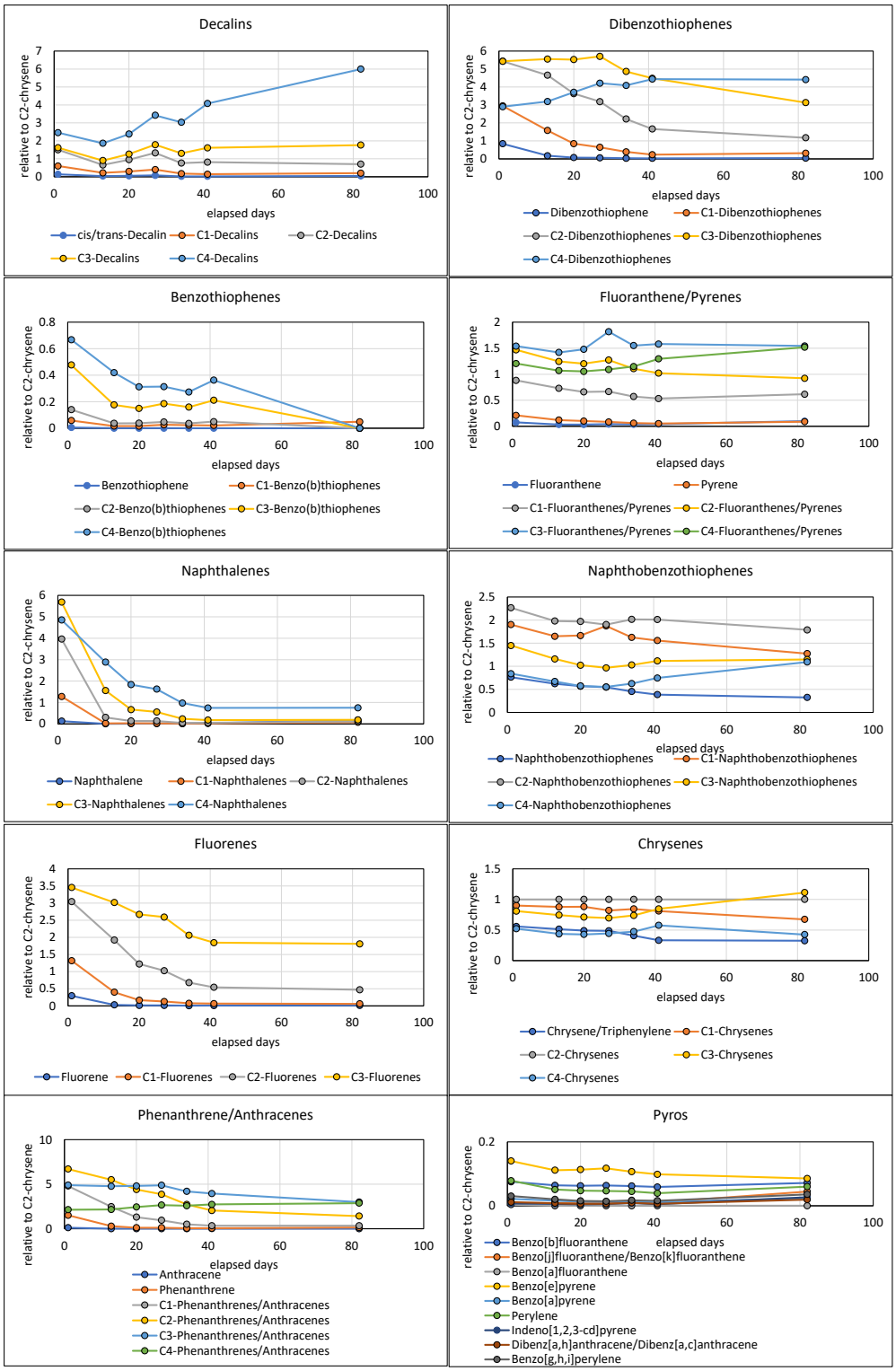


Figure SI3-1. Time series of proportional concentrations of PAH families (relative to C2-chrysene) in spill site (Hot Zone) mussels. Most show simple log decay depletion curves but others show accumulation (e.g., C4-phenanthrene/anthracene) or a mixed response (e.g., C1-, C2- and C3-naphthobenzothiophene).

Table SI3-1. PAH analytes and depuration trends sorted by homologous families (in GC elution order) (left columns) and by solubility, log K_{OW} (right columns). Lipophilicity, log KLIPW, is calculated from log K_{OW}. Depuration curves based on C2-chrysene normalized values, depict a modality of simple depletion vs accumulation. Here, the trend lines are unit-less and just sequential, not time-scaled. PAH ordered by log K_{OW} (right columns) show a general transition from simple depletion to accumulation or mixed response when log K_{OW} value > 6.3 (grey highlighted K_{OW} values). Singular upticks at the end of trend lines are trace-level lab artifacts.

PAH by homologus families	log Kow	log Klipw	Depuration
Benzo(b)thiophene	3.12	3.12	
C1-Benzo(b)thiopenes	3.65	3.72	
C2-Benzo(b)thiophenes	4.17	4.31	
C3-Benzo(b)thiopenes	4.66	4.87	
C4-Benzo(b)thiopenes	5.18	5.46	
Naphthalene	3.37	3.40	
C1-Naphthalenes	3.87	3.97	
C2-Naphthalenes	4.38	4.55	
C3-Naphthalenes	5.00	5.26	
C4-Naphthalenes	5.30	5.60	
Acenaphthene	3.92	4.03	
Acenaphthylene	4.00	4.12	
Fluorene	4.18	4.32	
C1-Fluorenes	4.97	5.22	
C2-Fluorenes	5.20	5.48	
C3-Fluorenes	5.70	6.05	
Anthracene	4.54	4.73	
Phenanthrene	4.57	4.77	
C1-Phenanthrenes/anthracenes	5.12	5.39	
C2-Phenanthrenes/anthracenes	5.25	5.54	
C3-Phenanthrenes/anthracenes	5.92	6.30	
C4-Phenanthrenes/anthracenes	6.32	6.76	
Dibenzothiophene	4.34	4.51	
C1-Dibenzothiophenes	4.86	5.10	
C2-Dibenzothiophenes	5.33	5.63	
C3-Dibenzothiophenes	5.81	6.18	
C4-Dibenzothiophenes	6.34	6.78	
Benzo(b)fluorene	5.75	6.11	
Fluoranthene	5.22	5.51	
Pyrene	5.18	5.46	
C1-Fluoranthenes/pyrenes	5.29	5.59	
C2-Fluoranthenes/pyrenes	5.56	5.89	
C3-Fluoranthenes/pyrenes	6.28	6.71	
C4-Fluoranthenes/pyrenes	6.69	7.18	
Naphthobenzothiophene	5.60	5.94	
C1-Naphthobenzothiophenes	6.38	6.82	
C2-Naphthobenzothiophenes	6.87	7.38	
C3-Naphthobenzothiophenes	7.37	7.95	
C4-Naphthobenzothiophenes	7.79	8.43	
Benzo(a)anthracene	5.91	6.29	
Chrysene + Triphenylene	5.55	5.88	
Chrysene	5.86	6.23	
C1-Chrysenes	6.14	6.55	
C2-Chrysenes	6.59	7.06	
C3-Chrysenes	6.97	7.50	
C4-Chrysenes	7.42	8.01	
Benzo(b)fluoranthene	5.80	6.17	
Benzo(j+k)fluoranthene	6.20	6.62	
Benzo(a)fluoranthene	6.54	7.01	
Benzo(e)pyrene	6.44	6.89	
Benzo(a)pyrene	6.35	6.79	
Perylene	6.25	6.68	
Indeno(1,2,3-c,d)pyrene	7.53	8.13	
Dibenz(a,h+a,c)anthracene	6.75	7.25	
Benzo(g,h,i)perylene	6.50	6.96	

PAH by log Kow	log Kow	log Klipw	Depuration
Benzo(b)thiophene	3.12	3.12	
Naphthalene	3.37	3.40	
C1-Benzo(b)thiophenes	3.65	3.72	
C1-Naphthalenes	3.87	3.97	
Acenaphthene	3.92	4.03	
Acenaphthylene	4.00	4.12	
C2-Benzo(b)thiophenes	4.17	4.31	
Fluorene	4.18	4.32	
Dibenzothiophene	4.34	4.51	
C2-Naphthalenes	4.38	4.55	
Anthracene	4.54	4.73	
Phenanthrene	4.57	4.77	
C3-Benzo(b)thiophenes	4.66	4.87	
C1-Dibenzothiophenes	4.86	5.10	
C1-Fluorenes	4.97	5.22	
C3-Naphthalenes	5.00	5.26	
C1-Phenanthrenes/anthracenes	5.12	5.39	
C4-Benzo(b)thiophenes	5.18	5.46	
Pyrene	5.18	5.46	
C2-Fluorenes	5.20	5.48	
Fluoranthene	5.22	5.51	
C2-Phenanthrenes/anthracenes	5.25	5.54	
C1-Fluoranthenes/pyrenes	5.29	5.59	
C4-Naphthalenes	5.30	5.60	
C2-Dibenzothiophenes	5.33	5.63	
Chrysene + Triphenylene	5.55	5.88	
C2-Fluoranthenes/pyrenes	5.56	5.89	
Naphthobenzothiophene	5.60	5.94	
C3-Fluorenes	5.70	6.05	
Benzo(b)fluorene	5.75	6.11	
Benzo(b)fluoranthene	5.80	6.17	
C3-Dibenzothiophenes	5.81	6.18	
Chrysene	5.86	6.23	
Benzo(a)anthracene	5.91	6.29	
C3-Phenanthrenes/anthracenes	5.92	6.30	
C1-Chrysenes	6.14	6.55	
Benzo(j+k)fluoranthene	6.20	6.62	
Perylene	6.25	6.68	
C3-Fluoranthenes/pyrenes	6.28	6.71	
C4-Phenanthrenes/anthracenes	6.32	6.76	
C4-Dibenzothiophenes	6.34	6.78	
Benzo(a)pyrene	6.35	6.79	
C1-Naphthobenzothiophenes	6.38	6.82	
Benzo(e)pyrene	6.44	6.89	
Benzo(g,h,i)perylene	6.50	6.96	
Benzo(a)fluoranthene	6.54	7.01	
C2-Chrysenes	6.59	7.06	
C4-Fluoranthenes/pyrenes	6.69	7.18	
Dibenz(a,h+a,c)anthracene	6.75	7.25	
C2-Naphthobenzothiophenes	6.87	7.38	
C3-Chrysenes	6.97	7.50	
C3-Naphthobenzothiophenes	7.37	7.95	
C4-Chrysenes	7.42	8.01	
Indeno(1,2,3-c,d)pyrene	7.53	8.13	
C4-Naphthobenzothiophenes	7.79	8.43	

These log K_{OW} transition patterns may suggest why some PAH would simply deplete while others accumulate. Early studies proposed a "multiple compartment model" whereby some of the accumulated hydrocarbons are released rapidly before a much slower release of those remaining (Stegeman and Teal, 1973; Farrington et al., 1982; Mason, 1988). They report that chronically polluted bivalves lose their

burden more slowly because the hydrocarbons have been accumulated into "stable compartments" that are not available for rapid depuration. Mason proposed that after 46 days, the depuration curve of 0.15 ppm oil exposure could be broken into two exponential rate losses: "fast" during the first 28 days and "slower" thereafter. Two depuration rates support a concept of gut clearance of an accumulated substance followed by slower tissue clearance, the latter either diffusively or being metabolized via the mussel's AHR detox system.

Lipophilicity (lipid affinity) is a molecule's key property in transport processes, including intestinal absorption, membrane permeability, protein binding, and distribution to different tissues and organs. Endo et al. (2011) address this affinity mechanism in tissues by modeling differences in accumulation properties of neutral organic molecules (including PAH) between different types of lipids, i.e., accumulation into storage lipids versus membrane lipids. They find that a molecule's lipid affinity, expressed as log KLIPW (liposome/water partitioning coefficient, the inverse to water solubility) reasonably correlates to log K_{OW} (octanol/water partitioning coefficient) but suggests that poly-parameter linear free energy relationships (PP-LLFERs) are more concise. Note that accumulation may also result from PAH interactions with non-lipids such as proteins, however, lipid storage models match well with organism results (Endo et al., 2011).

These lipophilicity concepts work well for the traditional suite of PAH but not so much for polar or extreme hydrophobic compounds, including the entire suite of petroleum (S/T) biomarkers. For example, the recalcitrant forensic tracer, 17β(H),21β(H)-hopane, is mostly insoluble in water with high log K_{OW} values (estimated to be in the 10-15 range depending on the modeling program vs 3-7 for traditional PAH). From its depletion plot in mussel tissue (Figure 4) hopane shows a simple logarithmic depletion trend and demonstrates no sign of accumulation despite its log K_{OW} implying it to be highly lipophilic. The same is true for all the reported petroleum biomarkers; triterpenes, hopanes and steranes show only simple depletion trends (n=13, 20, 17 analytes respectively) (Figure SI3-2) with only a couple of exceptions noted in the results section.

A mechanistic/free energy lipid-storage paradigm (Endo et al., 2011) suggests a reason for their lipophobic behavior. Hopanes, as a group, are non-aromatic C30 pentacyclic triterpanes (i.e., composed of six isoprene subunits) consisting of four six-

membered rings and one five-membered ring (Figure SI5-3). Commonly containing 27-35 carbon atoms, they derive from precursors in bacterial bi-lipid membranes.

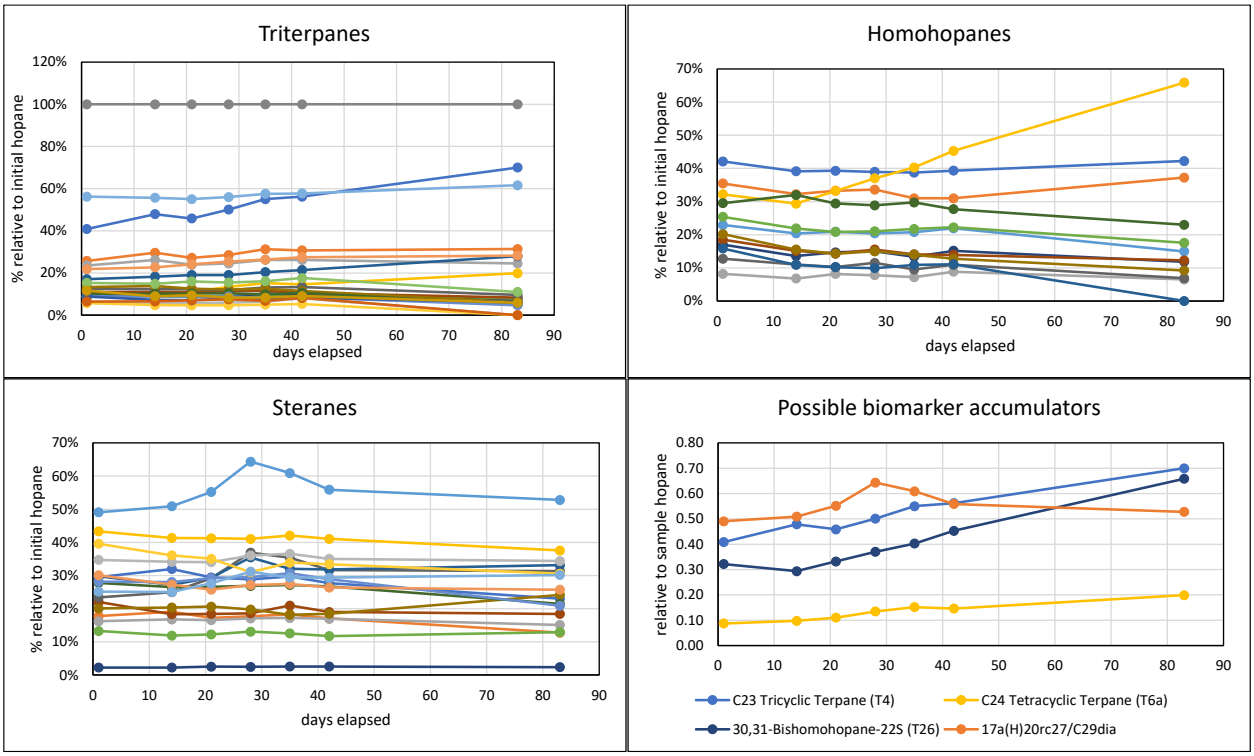


Figure SI3-2. Biomarker depletion relative to hopane depicts mostly simple depletion (mildly decreasing or flat trends) rather than accumulation. The few accumulating examples (lower right panel – duplicated from the three other plots) are exceptions from trace-level lab artifacts.

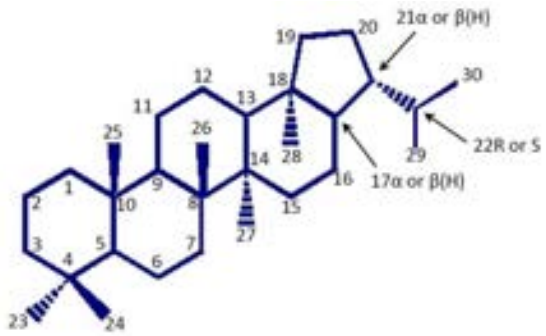


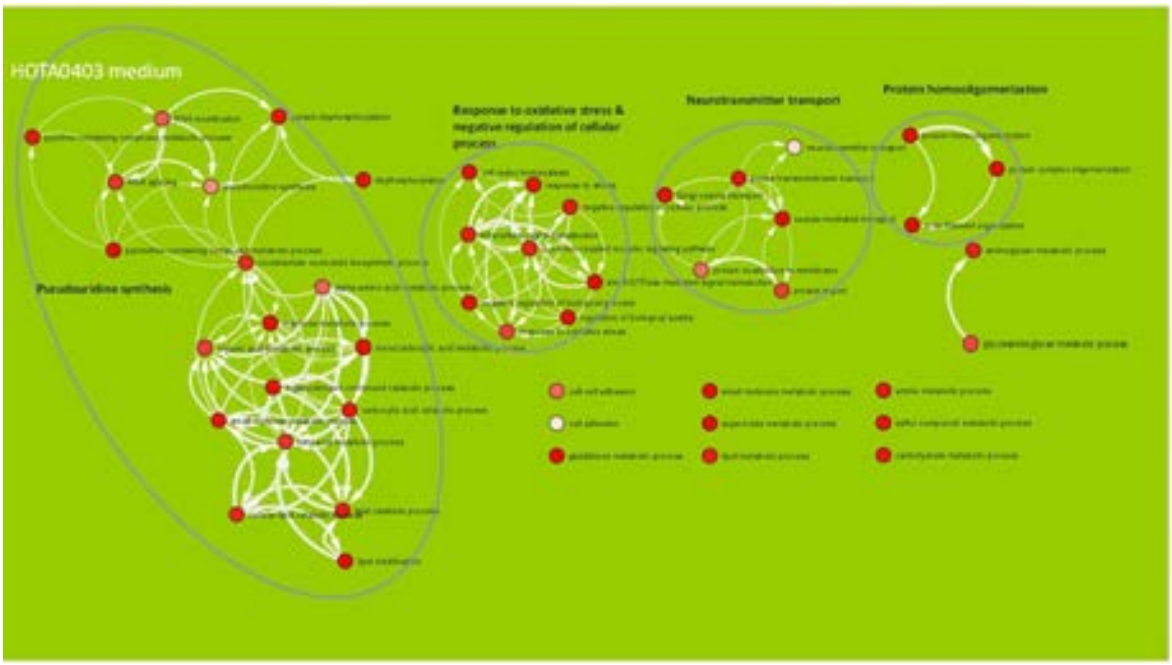
Figure SI3-3. Carbon model of 17α(H),21β(H)-hopane. The node of each solid line represents a carbon atom (here numbered, 1-30). Solid and dashed triangles indicate upper or downward projection of carbons. The αβ, ββ, or βα stereoisomers describe alternative hydrogen projections.

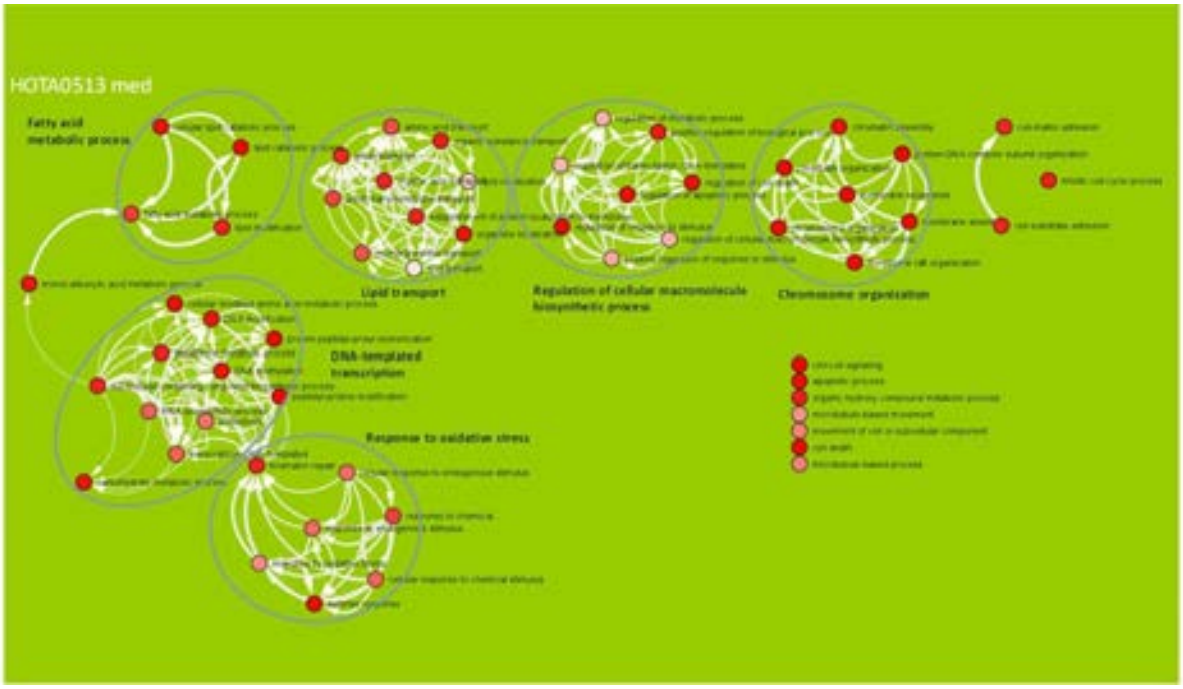
But even if derived from lipid membranes, hopanes seem unlikely to

reincorporate or for there to be sufficient "free energy" to slip through lipid membranes. Hopanes comprise three stereoisomeric series: 17α(H),21β(H)-, 17β(H), 21β(H)-, and 17β(H),21α(H)-hopanes, with the α and β notations indicating whether the hydrogen atoms are below or above the plane of the page, respectively (Figure SI3-3). Hopanes with the αβ configuration are characteristic of petroleum because of their greater thermodynamic stability relative to the other configurations (ββ, βα) after eons of microbial, heat, and pressure maturation in the oil formation. The major precursors for the hopanes in living organisms have "biological" or ββ stereochemistry, which is almost flat but not a planar aromatic molecule like the 4 and 5-ring PAH. Those saturated six-member hopane rings can exist in either a "boat or chair" configuration as controlled by the stereochemistry and the position of alkyl and hydrogen substituents. Hopanes are also amphipathic (e.g., having both hydrophilic and lipophilic structural components), which along with the flatter ββ configuration, appears to be necessary for insertion into lipid membranes. In oil formations, because the ββ stereochemical arrangement is thermodynamically less stable, diagenesis and catagenesis of the precursors result in transformation of ββ precursors to αβ hopane, the petroleum biomarker. Per free-energy models (Endo et al., 2011), we speculate that the αβ hopanes with their protruding methyl groups and hydrogens (Figure SI3-3) are unlikely to slip into alignment with the ββ forms in order to enter living membranes nor are configured to be efficiently processed by metabolic enzymes into storage lipids. And perhaps, this accounts for their anomalous lipophobic behavior in mussels (Figure SI3-3).

If the molecular configuration produces both hydrophobic and lipophobic behavior in tissues, then despite its high theoretical KLIPW, the hopane depletion curve makes sense. It suggests biomarkers are not bioavailable to the mussel or its gut microbes. Regarding hopane's short half-life in mussels, recent work by Staniszewska et al. (2017) corroborates that less lipophilic hydrocarbons (bisphenol A, BPA in their study), even the portion collected by lipids in the digestive tract, can theoretically be eliminated faster. Similarly, their seabird guano study (Staniszewska et al., 2014) reported smaller BPA accumulation in tissues, with a simultaneous increase in elimination compared to the removal of higher K_{OW} alkylphenols. Work by Farrington et al. (2020) supports this concept, demonstrating longer retention as PAH increase in alkylation and molecular weight in spill-contaminated mussels.

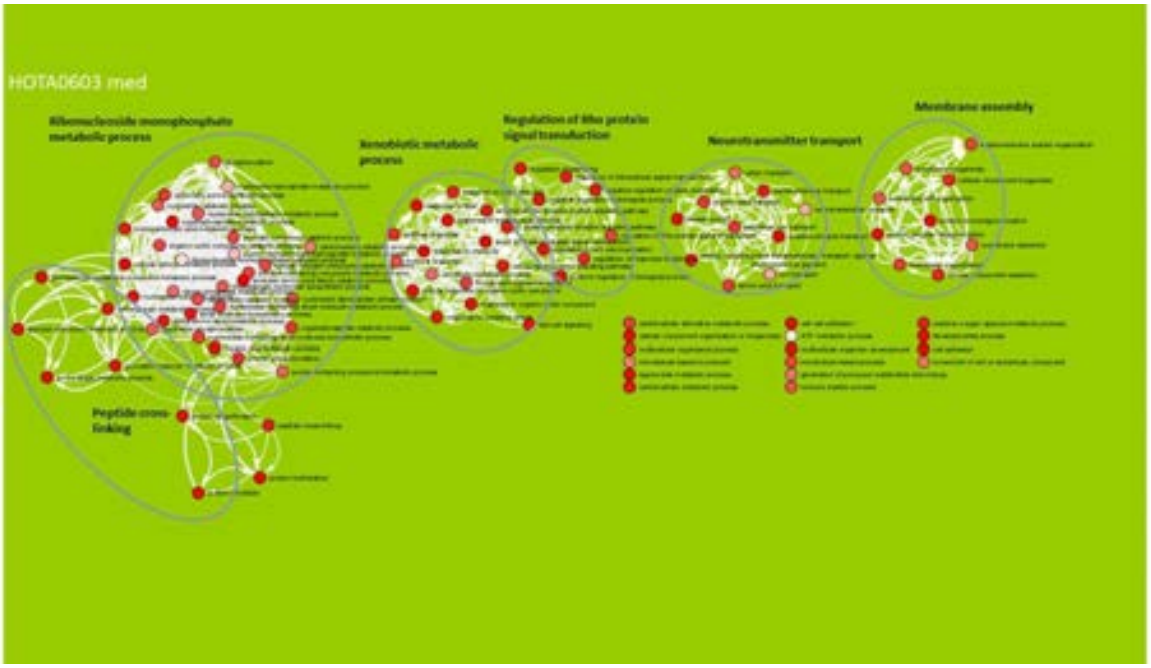
SI-4 Appendix 4 - Revigo Cluster plots of Gene Ontology Functions





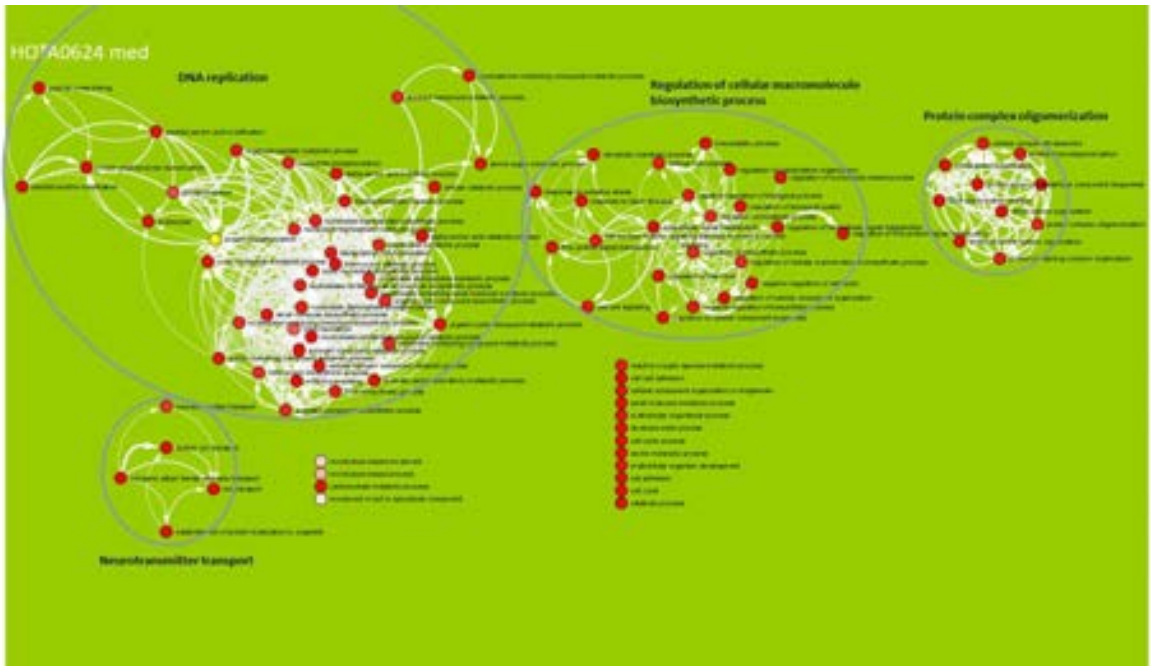
HOTA0520 med
Glycerophospholipid metabolic process
Neurotransmitter transport
Immune response
Cell redox homeostasis
Protein ubiquitination
Amino sugar metabolic process

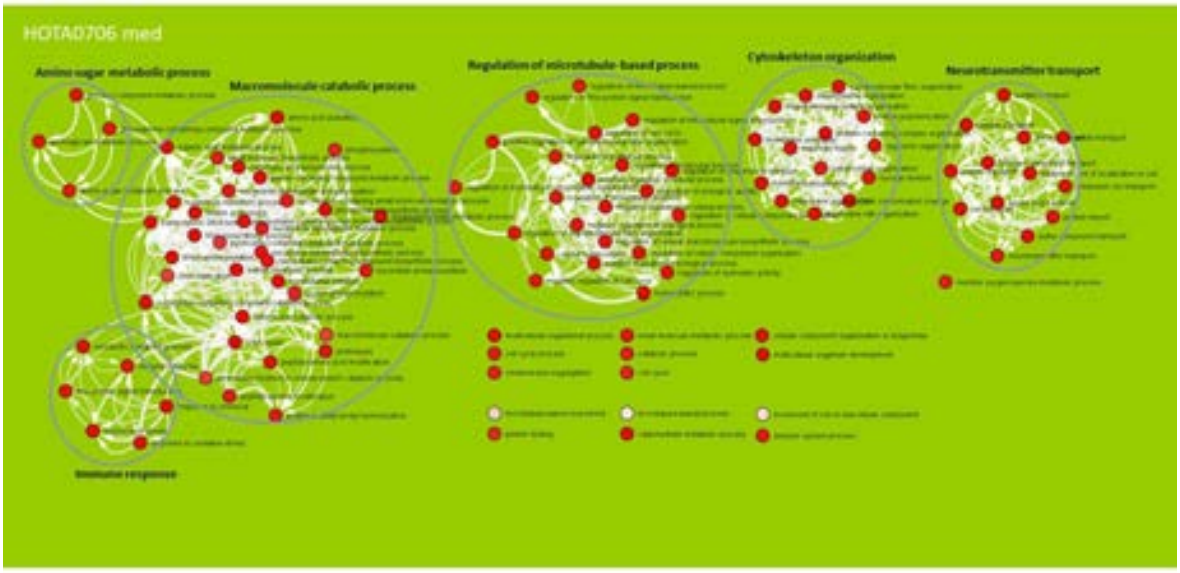
HOTA0527 med
Ribonucleoside monophosphate metabolic process
Nucleotide metabolic process
Neurotransmitter transport
Protein complex oligomerization
Regulation of biological quality
Amino sugar metabolic process

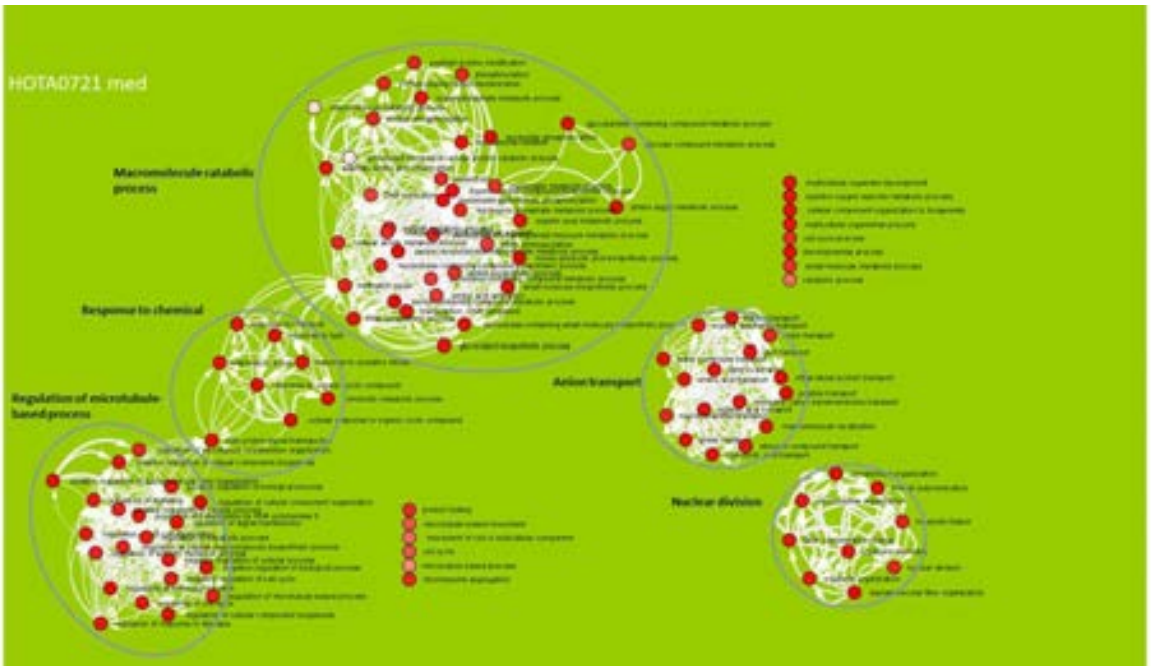

HOTA0603 med
Ribonucleoside monophosphate metabolic process
Xenobiotic metabolic process
Regulation of Rho protein signal transduction
Neurotransmitter transport
Membrane assembly
Peptide cross-linking

HOTA0610 med
Negative regulation of cellular process
Xenobiotic metabolic process
Amino sugar metabolic process
Metal ion transport
Alpha-amino acid catabolic process

HOTA0618 med
Amino sugar metabolic process
Cellular component assembly
Neurotransmitter transport
Protein phosphorylation & organic cyclic compound catabolic process
Xenobiotic metabolic process
Regulation of cytoskeleton organization


HOTA0624 med
DNA replication
Regulation of cellular macromolecule biosynthetic process
Protein complex oligomerization
Neurotransmitter transport


HOTA0706 med
Amino sugar metabolic process
Macromolecule catabolic process
Regulation of microtubule-based process
Cytoskeleton organization
Neurotransmitter transport
Immune response


HOTA0721 med
Macromolecule catabolic process
Response to chemical
Regulation of microtubule-based process
Anion transport
Nuclear division

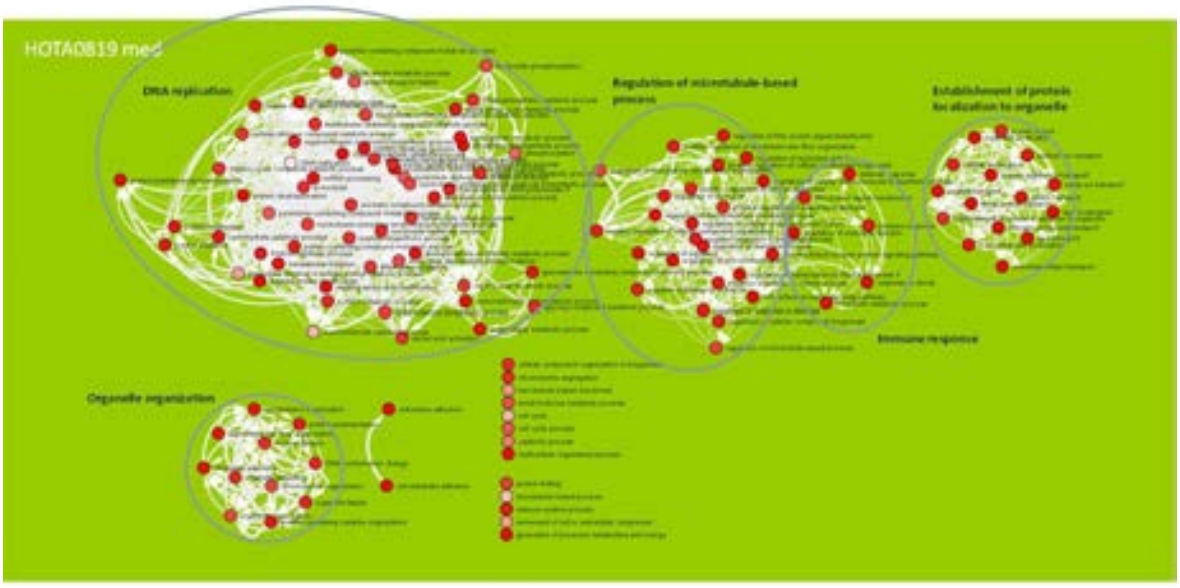


Figure S4-1. Time series of GO pathway linkages. Groupings of cellular functions from the time-series of mussel adductor tissues sampled over time (11 timepoints, starting 30 April and ending on 19 August 2020). Groupings depict gene expression differences between samples from the spill site (HOTA) and samples from the unoiled control site (BAY1A). White arrows link gene functions in their respective pathways. Listed single genes have no linkage. Over time, an increasing complexity of linkages is noted.