REPORT: PORT VALDEZ MUSSEL TRANSCRIPTOMICS

PWSRCAC Contract number: 951.20.06

November 20, 2019

Lizabeth Bowen[1], Austin Love[2], Shannon Waters[1], Katrina Counihan[3], Brenda Ballachey[4], and Heather Coletti[5]

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

lbowen@ucdavis.edu. 530-752-5365

[2]Prince William Sound Regional Citizens' Advisory Council, Valdez, AK, USA.

[3]Alaska SeaLife Center, Seward, AK 99664 USA

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

[5]National Park Service, Inventory & Monitoring Program, Southwest Alaska Network, Anchorage, AK, USA

TABLE OF CONTENTS

ACRONYMS

AMT – Saw Island sample site
B[a]P – benzo[a]pyrene
CaM – Calmodulin gene
Casp8 – Caspase 8 gene
CCOIV – Cytochrome C Oxidase IV gene
cDNA – complementary deoxyribonucleic acid
CHI – Chitinase gene
CNN – Calponin gene
Cyp3 – Cytochrome P450 family 3 gene
C_T – cycle threshold crossing values
GOC – Gold Creek sample site
Harbor – Whittier Harbor, Seward Harbor and Cordova Harbor sample sites
HIFa – hypoxia-inducible factor alpha gene
HSP70 – heat shock protein 70 gene
HSP90 – heat shock protein 90 gene
JPT – Jackson Point sample site
KATM – Katmai
LACL – Lake Clark
mRNA – messenger ribonucleic acid
MIF – macrophage migration inhibitory factor gene
MT20 – Metallothionein 20 gene
Myt – Mytilin gene
MytB – Myticin B gene
NMDS – non-parametric multidimensional scaling
PCR – polymerase chain reaction
PWS – Prince William Sound
PWS Field – Herring Bay, Hogan Bay, Iktua Bay, Johnson Bay, and Whale Bay sample sites
PWSRCAC – Prince William Sound Regional Citizens' Advisory Council
p53 – tumor protein 53 gene
SAC – Scientific Advisory Committee
USGS – United States Geological Survey
VH – Valdez Small Boat Harbor sample site
VMT – Valdez Marine Terminal
18S – ribosomal reference gene

INTRODUCTION

This project was inspired by a Prince William Sound Regional Citizens' Advisory Council (PWSRCAC) 2018 Science Night presentation by Dr. Katrina Counihan from the Alaska Sea Life Center. Dr. Counihan conducted blue mussel and razor clam transcriptomics experiments in Lake Clark National Park. Additionally, the Council was introduced to another researcher, Dr. Lizabeth Bowen, an ecologist with the United States Geological Survey (USGS), who was the lead author of a 2018 research paper titled *Gene transcription patterns in response to low level petroleum contaminants in Mytilus trossulus [blue mussels] from field sites and harbors in southcentral Alaska*, in Deep-Sea Research Part II.

Dr. Bowen's 2018 paper looked at whether or not blue mussel gene transcription is affected by oil pollution and other environmental factors. Dr. Bowen's research included five locations in Prince William Sound (PWS), Alaska that were oiled during the Exxon Valdez Oil Spill and three harbors in the Exxon Valdez Oil Spill Region: Seward, Whittier, and Cordova.

This project was designed to mirror previous mussel transcriptomics research conducted by Dr. Bowen, with the exception of sampling locations. Additionally, this project is meant to be a feasibility study for evaluation of whether this type of environmental monitoring could potentially become a part of the Long Term Environmental Monitoring Program (LTEMP).

Exposure to contaminants, including polycyclic aromatic hydrocarbons (PAHs), can lead to pathophysiological changes that may be subtle but significant, and difficult to detect using classical diagnostic methods. The earliest observable indications of physiological impairment can be altered levels of gene transcripts, evident prior to clinical signs (Farr and Dunn, 1999; McLoughlin et al., 2006; Poynton and Vulpe, 2009). Gene transcription is the process by which information from the DNA template of a particular gene is transcribed into messenger ribonucleic acid (mRNA) and eventually translated into a functional protein. Quantitative

analysis of mRNA therefore is used as a measure of gene transcription (Heid et al., 1996). The amount of a particular gene that is transcribed is physiologically dictated by a number of intrinsic and extrinsic factors, including stimuli such as infectious agents, toxin exposure, trauma, or neoplasia. Thus, gene transcription assays measure the physiological response of an organism to xenobiotics. In addition, the deleterious effects of toxic exposure may persist beyond metabolism and excretion of the toxin. The advantage of using gene expression assays is the ability to measure the chronic physiologic responses of an individual to the metabolic stimuli, independent of the continued presence of the original toxin or its metabolites.

Our objective was to explore the utility of gene transcript analysis for discerning altered physiological responses in Port Valdez mussels in comparison with mussels from PWS, Katmai, Lake Clark, and southcentral Alaska harbors selected as positive controls.

SCOPE OF WORK

1. **Collect mussel samples in Port Valdez** – This work was done by PWSRCAC staff, Austin Love. Love collected blue mussels from four locations in Port Valdez: Saw Island, Jackson Point, Gold Creek, and the Valdez Small Boat Harbor. Saw Island, Jackson Point, and Gold Creek are "traditional" LTEMP sites and mussels were collected from rocks in the intertidal zone at these locations. Saw Island and Jackson Point are both adjacent to the Valdez Marine Terminal (VMT) and are used for monitoring environmental impacts attributable to the operation of the VMT. Gold Creek is across Port Valdez from the VMT and is relatively far from known sources of hydrocarbon pollution; therefore, it is considered the clean control site to compare to the three other sampling locations. The Valdez Small Boat Harbor is not an LTEMP sample site; however, it supports recreational and commercial boating activity. It is suspected to be contaminated by hydrocarbons (e.g., diesel fuel, gasoline, motor oil, hydraulic oil);

therefore, it serves as a positive (likely contaminated) reference site to compare to the three other locations. Mussels from the Valdez Small Boat Harbor were collected from preserved (creosote) wood pilings. The Scientific Advisory Committee (SAC) reviewed and approved of the four locations sampled by this project.

Love collected ten mussels per site, selected randomly from each sampling location. After the mussels were collected their gill and adductor muscles were extracted and preserved. This means each mussel became two samples, one gill and one adductor muscle tissue sample. Therefore, each site yielded 20 samples, and including all four sampling locations, a total of 80 samples were collected. All mussels were sampled, dissected, and preserved on June 15, 2019. USGS provided the sample preservation fluid and vials. Those 80 mussel samples were shipped by Love to Dr. Lizabeth Bowen.

2. **Analyze mussel gene transcription** – This work was managed by Dr. Lizabeth Bowen and her team at the USGS Davis Field Station, at the University of California, in Davis, California. The transcription of the 15 genes listed in Table 1 was analyzed in all 80 samples. With the exception of one additional gene, tumor protein 53, this table was copied from Dr. Bowen's 2018 Deep-Sea Research Part II paper. Members of SAC agreed that this was an appropriate list of genes to analyze.
3. **Report results of mussel gene transcription analysis** – Dr. Lizabeth Bowen was the lead author of this report with input from PWSRCAC staff, SAC members, and LTEMP contractor William (Bill) B. Driskell.

METHODS

Mussel collection

Ten blue mussels were collected in June 2019 from each of four locations in Port Valdez: Saw Island (AMT), Jackson Point (JPT), Gold Creek (GOC), and the Valdez Small Boat Harbor (VH) (Figure 1). Saw Island, Jackson Point, and Gold Creek are "traditional" LTEMP sites. Saw Island and Jackson Point are both adjacent to the VMT; therefore, sampling there is meant to monitor any environmental impacts attributable to the operation of the VMT. Gold Creek is across Port Valdez from the VMT and is relatively far from known sources of hydrocarbon pollution; therefore, it was chosen as the clean control site to compare to the three other locations.

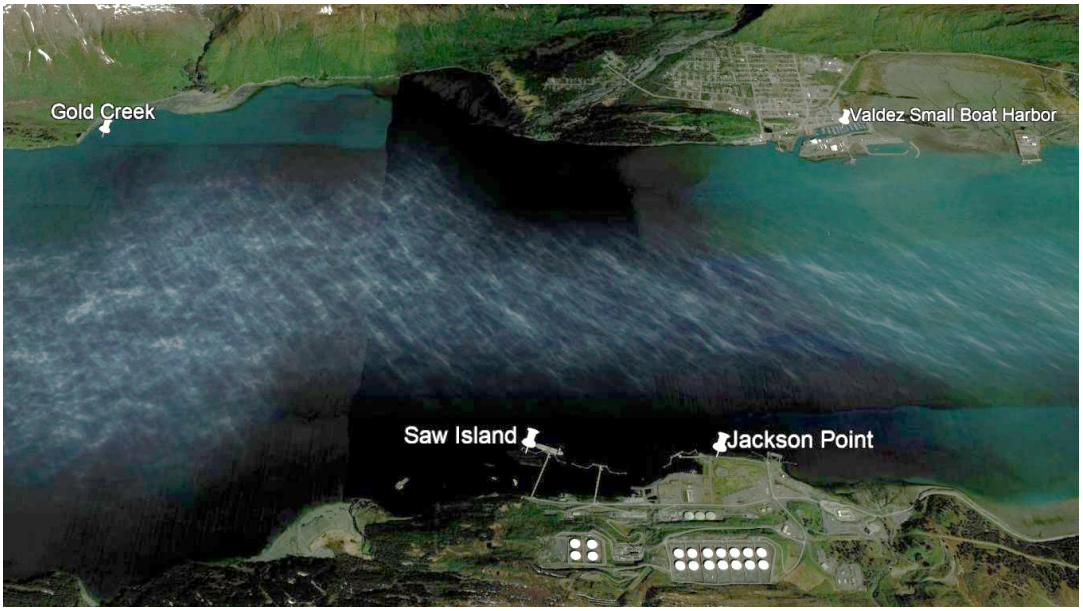


Figure 1. Map of Port Valdez showing the four mussel transcriptomics sampling locations (Source: Google Earth).

After the mussels were collected their gill and adductor muscles were extracted, preserved in RNAlater, and frozen until further analysis.

RNA extraction

Total RNA was extracted from pulverized adductor muscle and gill tissue using the RNeasy Lipid Tissue Mini Kit (Qiagen, Valencia, CA). To remove contaminating genomic DNA, the spin columns were treated with 10 U µl−1 of RNase-free DNase I (DNase, Amersham Pharmacia Biotech Inc.; www.apbiotech.com) at 20 °C for 15 min followed by extraction of total RNA and stored at −80 °C.

cDNA synthesis

A standard complementary deoxyribonucleic acid (cDNA) synthesis was performed on 2 µg of RNA template from each tissue. Reaction conditions included 4 units reverse transcriptase (Omniscript, Qiagen, Valencia, CA), 1 µM random hexamers, 0.5 mM each dNTP, and 10 units RNase inhibitor, in RT buffer (Qiagen, Valencia, CA). Reactions were incubated for 60 min at 37 °C, followed by an enzyme inactivation step of 5 min at 93 °C, and then stored at –30 °C until further analysis. Real-time polymerase chain reaction (PCR) reactions for the individual, mussel-specific housekeeping gene (18S) and genes of interest were run in separate wells (Table 1). Briefly, 1 µl of cDNA was added to a mix containing 12.5 µl of Applied Biosystems Fast SYBR Green® Master Mix [5 mM Mg2+] (Qiagen, Valencia, CA), 0.5 µl each of forward and reverse sequence specific primers (Invitrogen, Carlsbad, CA), and 10.5 µl of RNase-free water; total reaction mixture was 25 µl. The reaction mixture cDNA samples for each gene of interest and 18S were loaded into Fast 96 well plates in duplicate and sealed with optical sealing tape (Applied Biosystems, Foster City, CA). Reaction mixtures that contained water but no cDNA were used as negative controls. Amplifications were conducted on a Step-One Plus Real-time Thermal Cycler (Applied Biosystems, Foster City, CA). Reaction conditions were as follows: 50 °C for 2 min, 95 °C for 15 min, 40 cycles of 94 °C for 30 s, 60 °C for 30 s, 72 °C for 31 s, an

extended elongation phase at 72 °C for 10 min. Reaction specificity was monitored by melting curve analysis using a final data acquisition phase of 60 cycles of 65 °C for 30 s and verified by direct sequencing of randomly selected amplicons. Cycle threshold crossing values (C_T) for the genes of interest were normalized to the 18S housekeeping gene. The C_T value of a reaction is defined as the cycle number when the fluorescence of a PCR product can be detected above the background signal and is associated with the amount of PCR product in the reaction. Thus, the lower the C_T value, the more PCR product that is present.

Table 1. Genes selected for the transcription panel, the primary biological processes they are associated with, and what types of environmental interactions are known to affect their transcription.

Gene	Biological Process	Environmental Interaction
Calmodulin (CaM)	Metabolism, shell formation	Ocean acidification Temperature Dissolved oxygen
Caspase 8 (Casp8)	Apoptosis, necrosis, inflammation	Pathogens Contaminants
Macrophage migration inhibitory factor (MIF)	Innate immunity	Pathogens
Calponin (CNN)	Hypoxia	Ocean acidification Dissolved oxygen
Chitinase (CHI)	Metabolism, hypoxia	Ocean acidification Dissolved oxygen
Cytochrome C Oxidase IV (CCOIV)	Hypoxia	Dissolved oxygen
Heat shock protein 70 (HSP70)	Thermal stress	Temperature Pathogens Contaminants
Heat shock protein 90 (HSP90)	Thermal stress	Temperature Pathogens Contaminants
Hypoxia-inducible factor alpha (HIFa)	Hypoxia	Dissolved oxygen
Myticin B (MytB)	Innate immunity	Pathogens
Mytilin (Myt)	Innate immunity	Pathogens Ocean acidification
Metallothionein 20 (MT20)	Detoxification	Contaminants - metals
Cytochrome P450, family 3 (Cyp3)	Detoxification	Contaminants
Tumor protein 53 (p53)	Apoptosis	Contaminants - PAHs
18S	Ribosomal reference gene	Low interaction potential

Statistical analysis

Two separate analyses were conducted for the gill and adductor muscle tissue samples. Transcriptomic responses to stress are generally tissue specific, likely related to the specific physiological role of each tissue in the organism (Kadota et al., 2003). Analysis of quantitative PCR data was conducted using normalized values (housekeeping gene threshold crossing subtracted from the gene of interest threshold crossing); whereas lower normalized values are indicative of higher numbers of transcripts, we have inverted the values in the box plots for ease of interpretation. A change in normalized value of 2 is approximately equivalent to a 4-fold change in the amount of the transcript. The measured gene expression variation between samples is the sum of the true biological variation and several confounding factors (i.e. unavoidable differences in pipetting volume, unavoidable differences in sample adhesion to plastic wells, etc.) resulting in non-specific variation. The goal of normalization is to remove the non-biological variation as much as possible.

We first assessed the level of correlation between genes using a Pearson correlation matrix with reported r values (NCSS© Statistical Software, 2007, Kaysville, Utah). Due to the physiological relationship among genes, we expected many strong correlations to exist between genes.

We used conventional mean responses per group (based on location) with data assessed for statistical significance between classification ranks using Kruskal-Wallis with Dunns' Multiple Comparison Tests and Bonferroni correction, with reported Z values (NCSS© Statistical Software, 2007, Kaysville, Utah). One of the statistical analyses compared the 2019 Port Valdez mussel adductor tissue samples to mussel adductor tissue samples collected and analyzed from 2012 through 2015 in PWS (Bowen et al., 2018). The other statistical analysis compared the

2019 Port Valdez gill tissue samples to mussel gill samples collected in 2015 and 2016 at Katmai and Lake Clark National Parks (Counihan et al., 2019).

We conducted two-dimensional non-parametric multidimensional scaling (NMDS) of the Bray-Curtis dissimilarity from gene transcripts using the Vegan package in R version 3.5.0. The graphical representations show individual mussels clustered by similarity in transcription and not by pre-defined groups such as location. We obtained vectors describing the strength of each gene contribution to the two NMDS axes for graphical display. We evaluated goodness of fit for NMDS models using stress plots.

RESULTS

Correlations

The Pearson correlation coefficient, r, can take a range of values from +1 to -1 where a value of 0 indicates that there is no association between the two variables and a value closer to +1 or -1 indicates a stronger association. Values greater than 0 indicate a positive association (as the value of one variable increases, so does the value of the other variable), while values less than 0 indicate a negative association (as the value of one variable increases, the value of the other variable decreases). The relationship between two variables is generally considered strong when their r value is larger than 0.7. As expected, there were many correlations between genes, with more correlations evident in adductor muscle than in gill tissue (Tables 2 and 3).

Table 2. Correlation matrix of genes analyzed for transcript levels in adductor muscle tissues. Bolded values indicate a strong correlation (r > 0.7).

	CNN	CaM	Casp8	CCOIV	CHI	Cyp3	HIFa	HSP70	HSP90	MIF	MT20	MytB	Myt	P53
CNN	1.00	**0.71**	**0.77**	**0.72**	0.55	**0.74**	0.39	0.55	0.49	0.61	0.50	0.52	**0.72**	0.66
CaM	**0.71**	1.00	0.64	0.70	0.62	0.52	0.41	0.51	0.58	0.63	0.45	0.50	**0.77**	0.50
Casp8	**0.77**	0.64	1.00	0.65	0.67	**0.86**	0.50	0.58	0.45	0.61	0.70	0.41	0.57	**0.90**
CCOIV	**0.72**	0.70	0.65	1.00	0.56	0.55	0.27	0.55	0.38	0.51	0.42	0.49	0.61	0.61
CHI	0.55	0.62	0.67	0.56	1.00	0.50	0.36	0.37	0.38	0.38	0.42	0.32	0.48	0.54
Cyp3	**0.74**	0.52	**0.86**	0.55	0.50	1.00	0.49	0.58	0.43	0.61	0.67	0.37	0.55	**0.83**
HIFa	0.39	0.41	0.50	0.27	0.36	0.49	1.00	0.38	0.52	0.43	0.63	0.27	0.37	0.54
HSP70	0.54	0.51	0.58	0.55	0.37	0.58	0.38	1.00	0.37	0.51	0.47	0.24	0.43	0.57
HSP90	0.49	0.58	0.45	0.38	0.38	0.43	0.52	0.37	1.00	0.57	0.48	0.29	0.47	0.45
MIF	0.61	0.63	0.61	0.51	0.38	0.61	0.43	0.51	0.57	1.00	0.47	0.29	0.53	0.55
MT20	0.50	0.45	0.70	0.42	0.42	0.67	0.63	0.47	0.48	0.47	1.00	0.31	0.37	**0.82**
MytB	0.52	0.50	0.41	0.49	0.32	0.37	0.27	0.24	0.29	0.29	0.31	1.00	0.63	0.38
Myt	**0.72**	**0.77**	0.57	0.61	0.48	0.55	0.37	0.43	0.47	0.53	0.37	0.63	1.00	0.48
P53	0.66	0.50	**0.90**	0.61	0.54	**0.83**	0.54	0.57	0.45	0.55	**0.82**	0.38	0.48	1.00

Table 3. Correlation matrix of genes analyzed for transcript levels in gill tissues. Bolded values indicate a strong correlation (r > 0.7).

	CNN	CaM	Casp8	CCOIV	CHI	Cyp3	HIFa	HSP70	HSP90	MIF	MT20	MytB	Myt	P53
CNN	1.00	0.69	0.69	0.41	0.18	0.57	0.21	0.33	0.12	0.31	0.37	0.22	0.53	0.48
CaM	0.69	1.00	0.66	0.40	0.32	0.51	0.33	0.48	0.13	0.16	0.47	0.20	0.52	0.49
Casp8	0.69	0.66	1.00	0.40	-0.03	0.77	0.33	0.37	0.15	0.27	0.51	0.24	0.54	**0.74**
CCOIV	0.41	0.40	0.40	1.00	0.11	0.26	0.19	0.28	0.09	0.13	0.15	0.21	0.25	0.32
CHI	0.18	0.32	-0.03	0.11	1.00	-0.04	0.26	0.45	0.13	0.00	-0.04	-0.01	0.04	0.01
Cyp3	0.57	0.51	**0.77**	0.26	-0.04	1.00	0.32	0.31	0.11	0.24	0.53	0.16	0.43	**0.75**
HIFa	0.21	0.33	0.33	0.19	0.26	0.32	1.00	0.41	0.39	0.19	0.44	0.13	0.22	0.51
HSP70	0.33	0.48	0.37	0.28	0.45	0.31	0.41	1.00	0.26	0.18	0.21	0.05	0.23	0.43
HSP90	0.12	0.13	0.15	0.09	0.13	0.11	0.39	0.26	1.00	0.20	0.17	0.03	0.11	0.24
MIF	0.31	0.16	0.27	0.13	0.00	0.24	0.19	0.18	0.20	1.00	0.22	0.05	0.21	0.29
MT20	0.37	0.47	0.51	0.15	-0.04	0.53	0.44	0.21	0.17	0.22	1.00	0.10	0.31	0.48
MytB	0.22	0.20	0.24	0.21	-0.01	0.16	0.13	0.05	0.03	0.05	0.10	1.00	0.38	0.16
Myt	0.53	0.52	0.54	0.25	0.04	0.43	0.22	0.23	0.11	0.21	0.31	0.38	1.00	0.40
P53	0.51	0.43	0.24	0.29	0.48	0.16	0.51	0.43	0.24	0.29	0.48	0.16	0.40	1.00

Conventional mean responses

Geometric mean C_T values (with 95% upper and lower confidence limits) for transcription of genes of interest for mussel adductor and gill tissues were calculated based on location (Valdez Small Boat Harbor, Gold Creek, Jackson Point, Saw Island) and depicted in Table 4. Smaller numbers indicate greater levels of transcription. For comparison, conventional mean responses from mussels sampled in PWS sites (Field, Harbors), as well as Katmai and Lake Clark National Parks, are depicted in Bowen et al. 2018 and Counihan et al. 2019, respectively.

Table 4.

Tissue	Gene	Valdez small boat harbor	Gold Creek	Jackson Point	Saw Island
Adductor	**CNN**	18.05 (16.90, 19.28)	19.75 (18.76, 20.79)	21.42 (19.31, 23.76)	22.81 (22.26, 23.37)
	CaM	17.11 (15.96, 19.57)	19.09 (18.30, 19.92)	20.56 (18.40, 22.98)	22.09 (20.63, 23.64)
	Casp8	12.31 (11.26, 13.45)	14.45 (13.51, 15.46)	15.35 (13.54, 17.40)	16.13 (15.24, 17.06)
	CCOIV	16.88 (14.19, 20.08)	20.47 (19.67, 21.30)	21.57 (19.39, 24.00)	24.75 (23.09, 26.54)
	CHI	18.10 (17.67, 18.54)	19.07 (18.27, 19.91)	20.39 (18.81, 22.11)	20.68 (19.71, 21.71)
	Cyp3	13.55 (12.48, 14.71)	16.12 (15.34, 16.94)	16.87 (14.64, 19.43)	17.80 (16.98, 18.66)
	HIFa	13.54 (13.18, 13.92)	14.70 (13.86, 15.59)	13.97 (12.25, 15.93)	15.01 (14.14, 15.94)
	HSP70	9.87 (9.02, 10.79)	11.39 (10.75, 12.08)	11.75 (9.84, 14.02)	12.86 (11.63, 14.21)
	HSP90	12.37 (11.14, 13.73)	13.41 (12.01, 14.97)	13.01 (11.02, 15.36)	14.44 (13.20, 15.80)
	MIF	14.36 (12.16, 16.95)	20.53 (19.66, 21.42)	18.68 (15.43, 22.62)	20.83 (18.08, 23.99)
	MT20	7.25 (6.12, 8.58)	8.91 (8.32, 9.54)	9.13 (6.64, 12.55)	11.24 (9.79, 12.90)
	MytB	9.77 (7.90, 12.08)	14.54 (11.83, 17.86)	16.25 (13.05, 20.22)	16.48 (13.72, 19.80)
	Myt	14.71 (12.99, 16.66)	17.83 (16.54, 19.22)	19.74 (17.87, 21.82)	21.48 (20.50, 22.51)
	P53	13.27 (12.37, 14.24)	15.64 (14.75, 16.59)	15.58 (13.64, 17.80)	16.85 (15.84, 17.92)
Gill	**CNN**	23.78 (23.04, 24.55)	23.55 (22.72, 24.41)	25.83 (24.88, 26.82)	25.92 (25.09, 26.78)
	CaM	15.12 (14.32, 15.97)	14.75 (14.28, 15.24)	16.96 (16.32, 17.63)	16.74 (15.64, 17.91)
	Casp8	9.56 (8.80, 10.38)	8.59 (8.01, 9.21)	9.37 (8.64, 10.16)	9.61 (8.89, 10.40)
	CCOIV	16.13 (14.24, 18.28)	16.78 (15.80, 17.82)	18.17 (17.10, 19.31)	20.16 (18.49, 21.98)
	CHI	19.29 (17.78, 20.92)	19.42 (18.17, 20.75)	21.40 (20.69, 22.14)	20.96 (19.92, 22.06)
	Cyp3	13.48 (12.79, 14.21)	12.64 (11.66, 13.72)	13.05 (12.17, 14.00)	12.88 (12.06, 13.74)
	HIFa	12.42 (11.12, 13.87)	13.03 (12.34, 13.76)	13.00 (12.26, 13.78)	13.53 (12.95, 14.13)
	HSP70	10.06 (9.03, 11.21)	10.30 (9.93, 10.69)	11.16 (9.70, 12.84)	11.62 (10.47, 12.89)
	HSP90	12.07 (10.39, 14.03)	12.87 (11.45, 14.47)	12.51 (11.15, 14.04)	13.30 (12.21, 14.87)
	MIF	14.44 (12.53, 16.64)	17.48 (16.84, 18.15)	16.28 (13.79, 19.22)	16.85 (14.56, 19.50)
	MT20	5.63 (4.97, 6.37)	5.17 (4.83, 5.53)	5.03 (3.03, 8.33)	6.18 (5.26, 7.25)
	MytB	8.05 (6.29, 10.30)	10.45 (8.17, 13.35)	12.08 (9.06, 16.09)	9.84 (7.20, 13.44)
	Myt	13.42 (12.15, 14.81)	14.43 (13.25, 15.71)	16.37 (15.36, 17.44)	15.48 (14.41, 16.62)
	P53	11.21 (10.73, 11.72)	10.53 (10.09, 10.99)	11.13 (9.93, 12.47)	11.08 (10.58, 11.62)

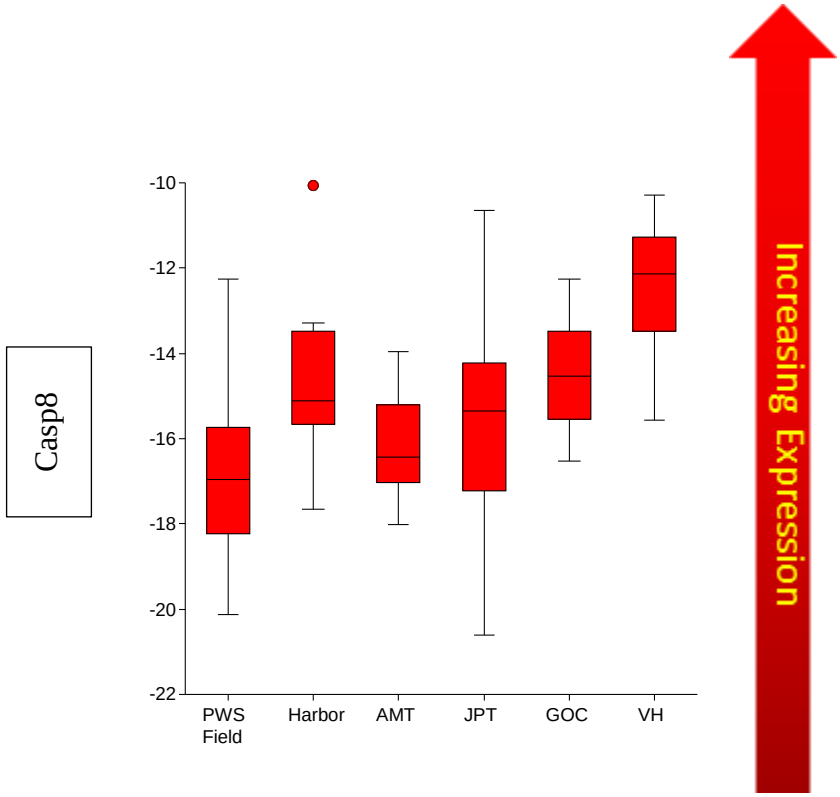
Adductor muscle tissue

The 2019 transcript profiles of adductor tissue sampled in Port Valdez at AMT, JPT, GOC, and VH were statistically compared with transcript profiles of adductor tissue sampled in five PWS field sites (PWS Field) thought to be relatively contaminant-free at the time of sampling (Herring Bay, Hogan Bay, Iktua Bay, Johnson Bay, and Whale Bay) and three positive control harbor sites (Harbor) with known contaminants (Whittier Harbor, Seward Harbor and Cordova Harbor) (data published in Bowen et al. 2018). Genes transcript levels associated with detoxification that were found to differ significantly among locations (i.e., PWS Field, Harbor, and Port Valdez sites) included Casp8, Cyp3, HSP70, MT20, and p53 (Figures 2-6; note: only genes with significantly different transcript levels among locations were depicted in the figures). In general, Port Valdez samples had higher levels of transcription in genes associated with detoxification than field or harbor sites. Genes with potential indirect relationships or no relationship to contaminant presence that were significantly different among locations included CNN, CaM, CCOIV, CHI, HIFa, MIF, MytB, and Myt. HSP90 was not significantly different among sites. Similar patterns existed for each gene, with transcription generally increasing from PWS Field, Harbor, AMT, JPT, GOC, to VH.

Although the transcription of these genes are not generally associated with detoxification, many are influenced by the presence of contaminants. For example, recent research by Banni et al. 2017 shows that controlled exposure to benzo[a]pyrene (B[a]P) results in transcriptional changes of CNN, CaM, Myt, CCOIV, and CHI. These are only a few of the indirect or "downstream" effects of contaminant exposure. The mechanism is as yet unclear, but an initiation of the detoxification pathway in this case results in effects on shell formation, mitochondrial activity, and immune function (Banni et al. 2017).

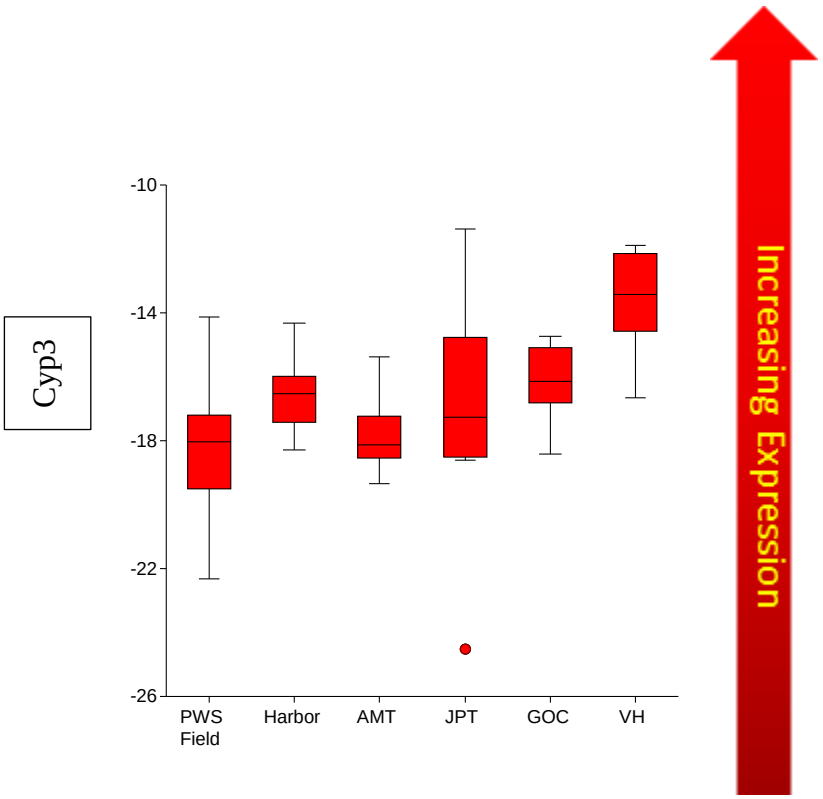
Figures 2 through 6 depict genes in mussel adductor tissue primarily associated with detoxification of contaminants including PAHs, compared among PWS Field, Harbors (i.e., contaminated reference sites), and Port Valdez sites (AMT, JPT, GOC, VH). Bars range from the 10th to the 90th percentile of normalized values for each gene. Circles represent 5th and 95th percentile outliers. Interpretation of gene abbreviations is provided in Table 1.

Figure 2.



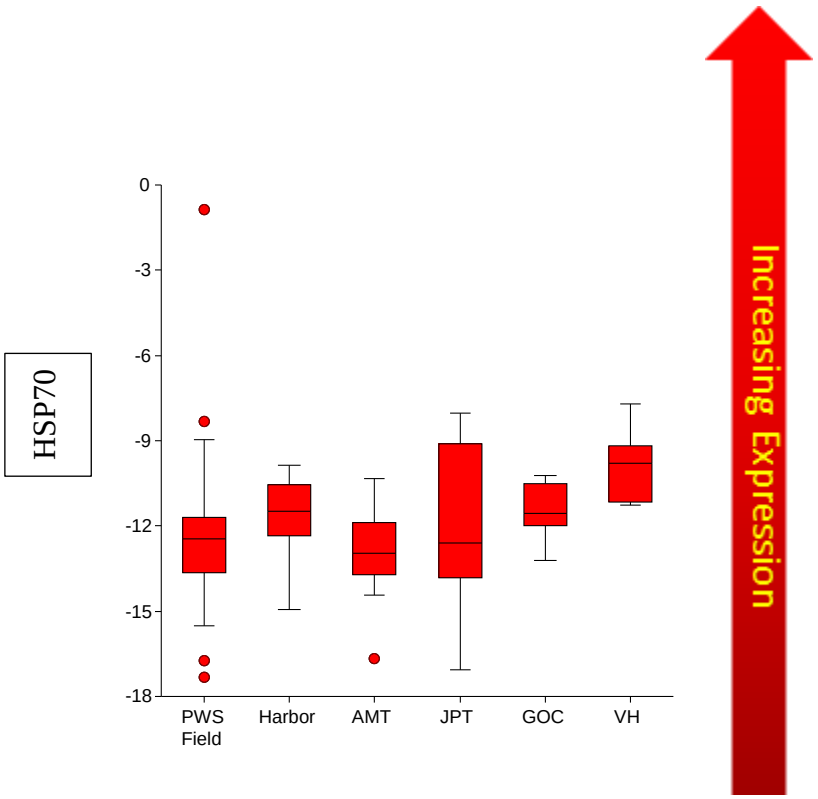
For Casp8, indicative of cell death, tissue death, and inflammation in the presence of pathogens and/or contaminants including PAHs, the lowest transcript levels in adductor muscle tissue were found in mussels sampled in western PWS Field sites, while the highest levels were found at VH. Significant differences: PWS Field < Harbor, GOC, VH; AMT < VH.

Figure 3.



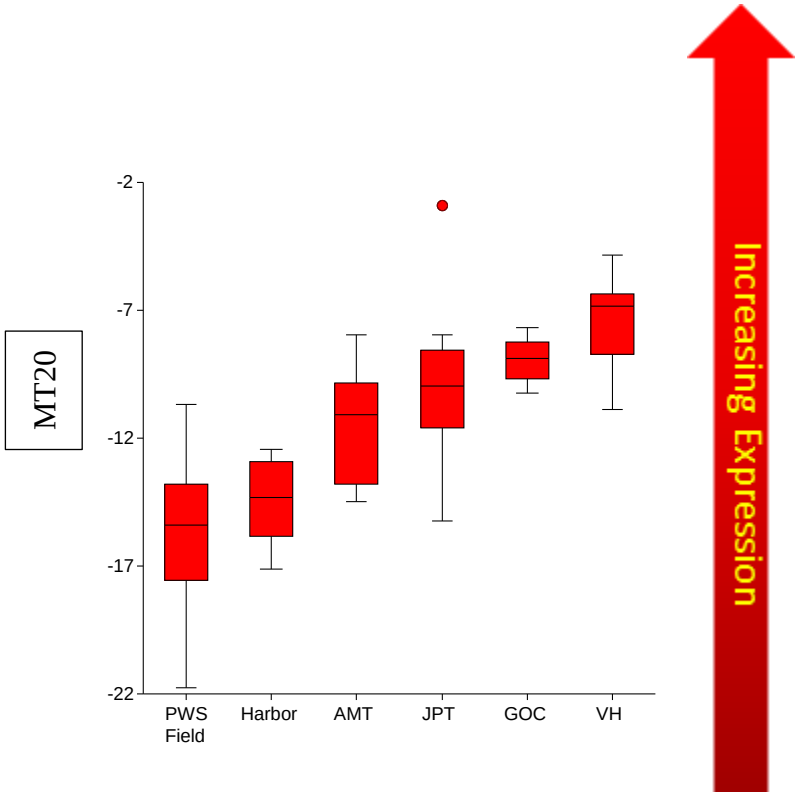
For Cyp3, indicative of contaminant detoxification activities including PAHs, the lowest transcript levels in adductor muscle tissue were found in mussels sampled at western PWS Field sites and AMT, while the highest levels were found at VH. Significant differences: PWS Field < Harbor, GOC, VH; AMT < VH

Figure 4.



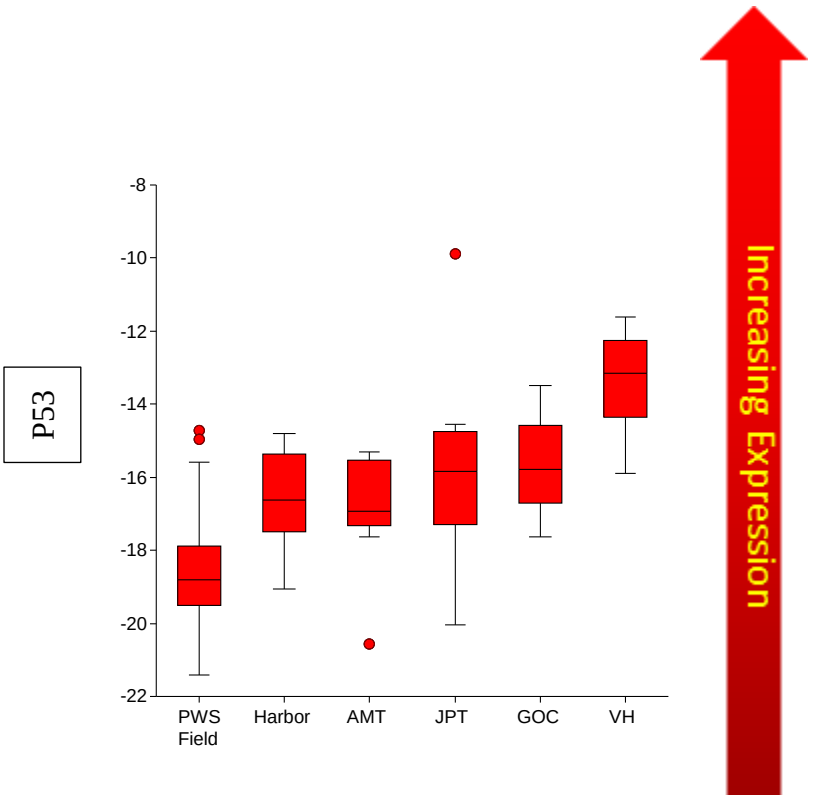
For HSP70, indicative of contaminant exposure (including PAHs) as well as general physiological stress, the lowest transcript levels in adductor muscle tissue were found in mussels sampled at AMT, while the highest levels were found at VH. Significant differences: PWS Field < VH; AMT < VH

Figure 5.



For MT20, specifically indicative of metal exposure, the lowest transcript levels in adductor muscle tissue were found in mussels sampled in western PWS Field sites, while the highest levels were found at GOC and VH. Significant differences: PWS Field < AMT, JPT, GOC, VH; Harbor < GOC, VH

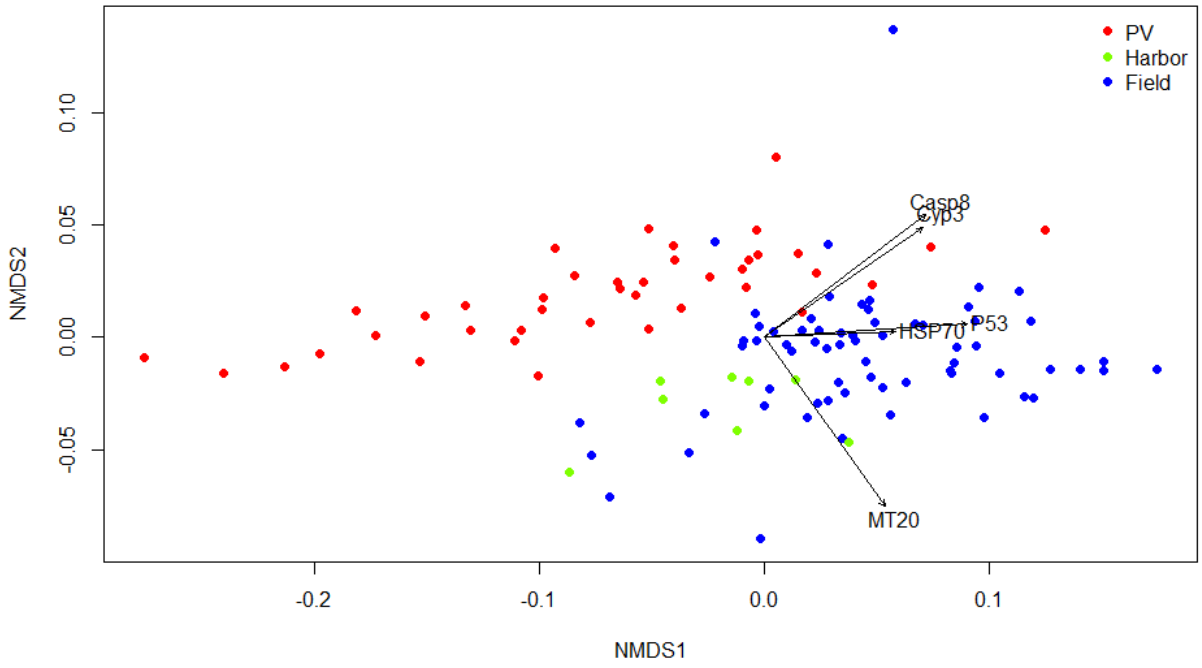
Figure 6.



For P53, primarily indicative of cell death and malignant transformation, as well as PAH exposure, the lowest transcript levels in adductor muscle tissue were found in mussels sampled in western PWS Field sites, while the highest levels were found at VH. Significant differences: PWS Field < Harbor, JPT, GOC, VH

Figure 7 depicts a two-dimensional non-parametric multidimensional scaling (NMDS) plot of the Bray-Curtis dissimilarity from adductor tissue gene transcripts. Gene transcripts were obtained from mussel adductor tissue collected at four sites in Port Valdez (PV), three harbors outside Port Valdez (Cordova, Seward, Whittier), and five field sites (Hogan Bay, Herring Bay, Iktua Bay, Johnson Bay, Whale Bay). The colors correspond to each of the broad categories as described in the figure legend (PV, Harbor, Field). The vector arrows signify the direction of maximum correlation for each gene transcript in the ordination space. The length of the arrow signifies the strength of the relationship of each metric and the two NMDS metrics, with longer arrows signifying greater strength. Our results show that NMDS1 is heavily influenced by P53 and HSP70. NMDS2 is most strongly associated with Casp8, Cyp3, and MT20. The plot clearly shows three clusters divided into PV and Harbor, and Field. Transcript levels of these genes are the most different among sites.

Figure 7.

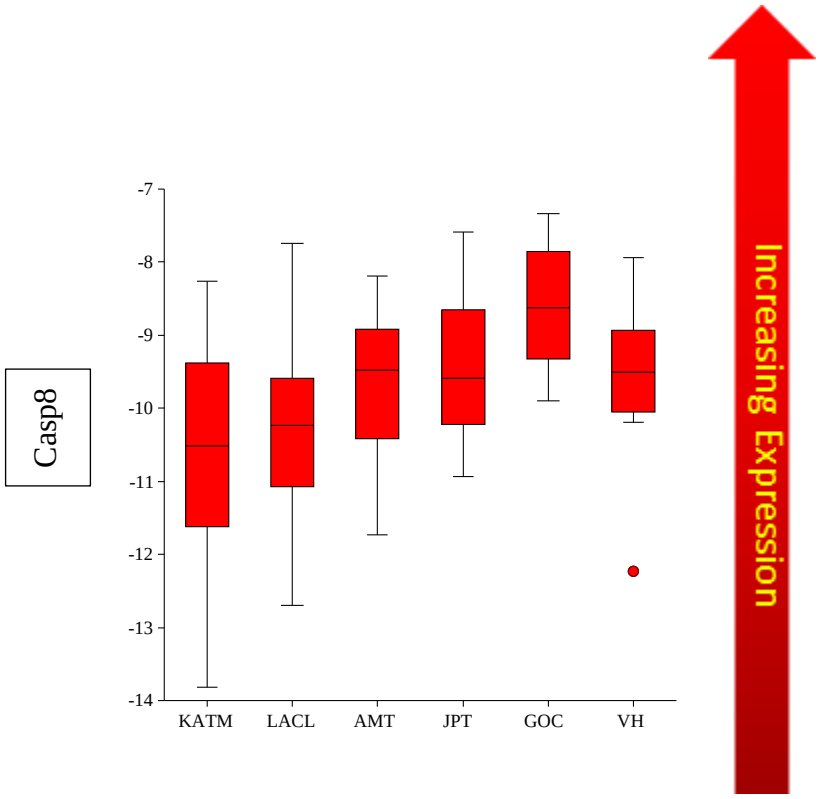


Gill tissue

Transcript profiles of gill tissue from mussels sampled in Port Valdez at AMT, JPT, GOC, and VH were statistically compared with transcript profiles of gill tissue from mussels sampled in Lake Clark (LACL) and Katmai (KATM) National Parks (collected in 2015 and 2016). Gene transcript levels associated with contaminant presence that were significantly different among locations included Casp8, Cyp3, MT20, and p53 (Figures 8-11; note: only genes with significantly different transcript levels among locations were depicted in the figures). Gene transcript levels that were not significantly different among sites included CNN, CaM, CCOIV, CHI, HIFa, MytB, and Myt. HSP70, HSP90, and MIF – for these genes there was either no relationship to contaminant presence or potential indirect relationships

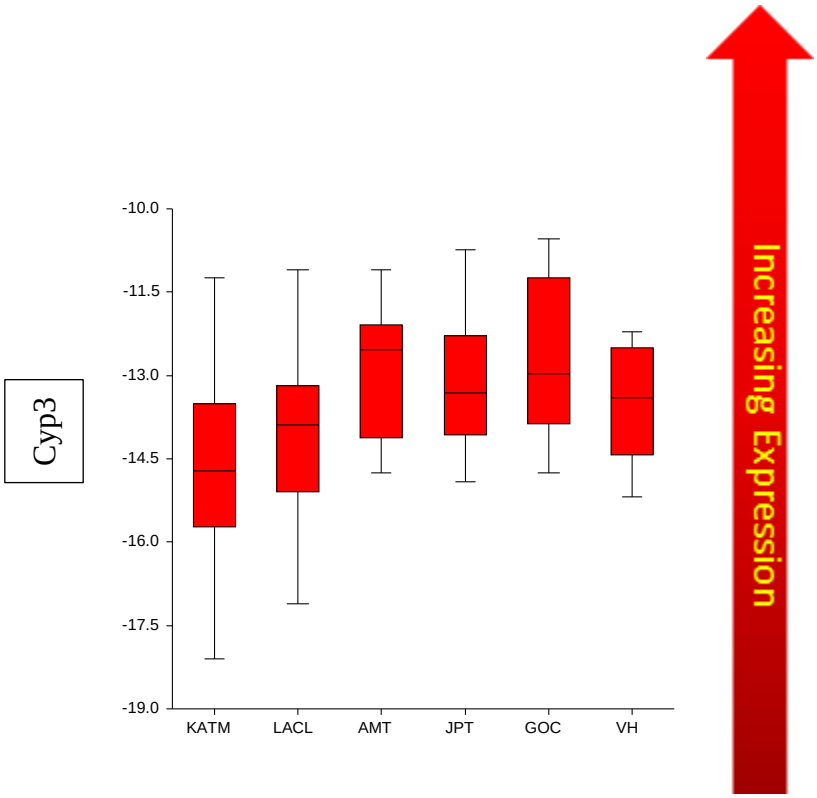
Figures 8 through 11 depict genes in mussel gill tissue primarily associated with detoxification of contaminants including PAHs, analyzed between KATM, LACL, and Port Valdez sites (AMT, JPT, GOC, VH). Bars range from the 10th to the 90th percentile of normalized values for each gene. Circles represent 5th and 95th percentile outliers. Interpretation of gene abbreviations is provided in Table 1.

Figure 8.



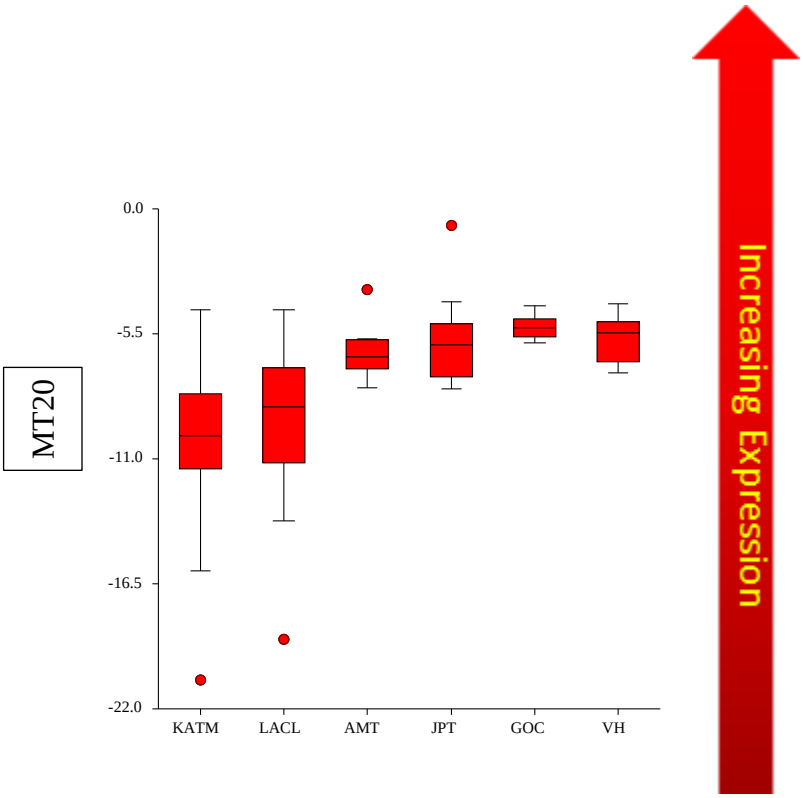
For Casp8, indicative of cell death, tissue death, and inflammation in the presence of pathogens and/or contaminants including PAHs, the lowest transcript levels in gill tissue were found in mussels sampled KATM and LACL sites, while the highest levels were found at GOC. Significant differences: KATM < GOC; LACL < GOC

Figure 9.



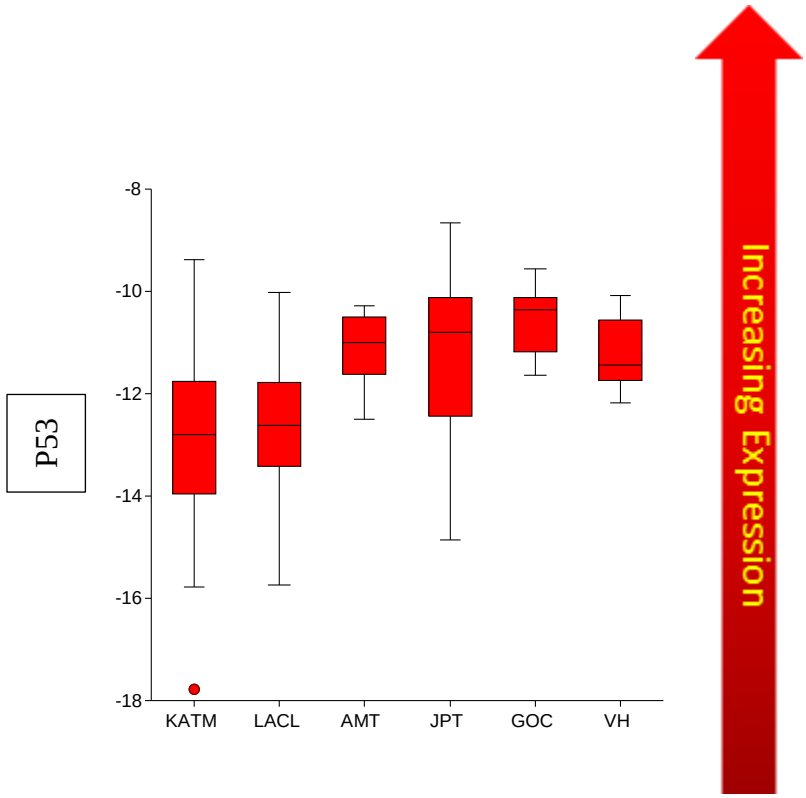
For Cyp3, indicative of contaminant detoxification activities including PAHs, the lowest transcript levels in gill tissue were found in mussels sampled at KATM sites, while the highest levels were found at AMT. Significant differences: KATM < AMT, GOC

Figure 10.



For MT20, specifically indicative of metal exposure, the lowest transcript levels in gill tissue were found in mussels sampled at KATM sites, while the highest levels were found at GOC and VH. Significant differences: KATM < AMT, JPT, GOC, VH; LACL < AMT, JPT, GOC, VH

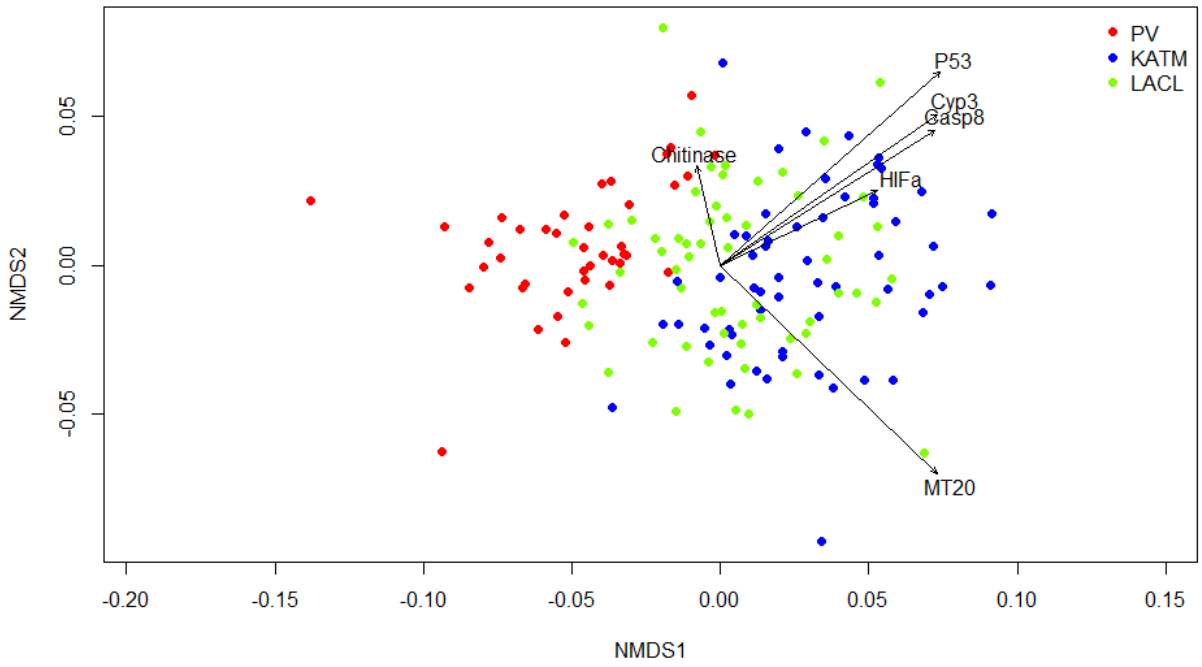
Figure 11.



For P53, primarily indicative of cell death and malignant transformation, as well as PAH exposure, the lowest transcript levels in gill tissue were found in mussels sampled at KATM and LACL sites, while the highest levels were found at GOC, JPT, and AMT. Significant differences: KATM < AMT, JPT, GOC, VH; LACL < AMT, GOC, VH

Figure 12 depicts two-dimensional non-parametric multidimensional scaling plot of the Bray-Curtis dissimilarity from gene transcripts. Gene transcripts were obtained from mussel gill tissue collected at four sites in Port Valdez (PV), and Lake Clark (LACL) and Katmai (KATM) National Parks. The colors correspond to each of the broad categories as described in the figure legend (PV, LACL, KATM). The vector arrows signify the direction of maximum correlation for each gene transcript in the ordination space. The length of the arrow signifies the strength of the relationship of each metric and the two NMDS metrics, with longer arrows signifying greater strength. Our results show that NMDS1 is influenced by HIFa, Casp8, Cyp3, and P53. NMDS2 is most strongly associated with CHI and MT20. The plot clearly shows two clusters almost entirely divided into PV and KATM/LACL. Transcript levels of these genes are the most different among sites.

Figure 12.

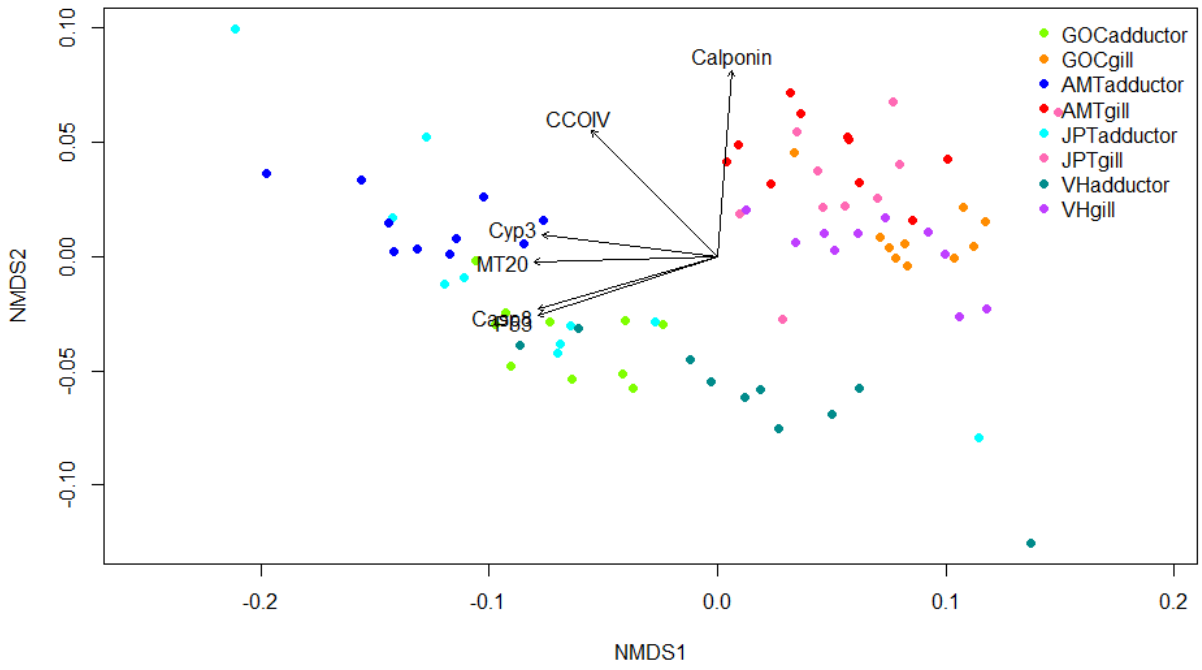


Tissue comparison

Transcript profiles of gill and adductor tissue from mussels sampled at AMT, JPT, GOC, and the VH were compared using NMDS (Figure 6). NMDS yielded two distinct clusters separating samples by tissue type. NMDS1 is influenced by MT20, Casp8, Cyp3, and P53, while NMDS2 is most strongly associated with Calponin and less strongly with CCOIV.

Figure 13 depicts two-dimensional non-parametric multidimensional scaling plot of the Bray-Curtis dissimilarity from gene transcripts. Gene transcripts were obtained from mussel gill (pink hue points) and adductor (blue hue points) tissue collected at four sites in Port Valdez, GOC, AMT, JPT, and VH. The vector arrows signify the direction of maximum correlation for each gene transcript in the ordination space. The length of the arrow signifies the strength of the relationship of each metric and the two NMDS metrics, with longer arrows signifying greater strength. Our results show that NMDS1 is influenced by MT20, Casp8, Cyp3, and P53. NMDS2 is most strongly associated with Calponin and less strongly with CCOIV. The plot clearly shows two clusters entirely separated by tissue type. Transcript levels of these genes are the most different among sites.

Figure 13.



CONCLUSIONS

We have conducted preliminary analyses of gene expression in two tissues, adductor muscle and gill, collected from mussels at four sites near Valdez in June 2019. For comparison with these samples, we have gene expression data for mussels from sites considered to be relatively clean (for adductor muscle: western PWS Field sites, collected 2012-2015, and for gill: KATM and LACL sites, collected 2015-2016). Additionally, we have gene expression data for adductor tissue in mussels collected from harbors (Cordova, Whittier and Seward) in 2014-2015, representing a group exposed to relatively high levels of contaminants.

In general, for both adductor muscle and gill, our analyses indicate the expression of genes associated with contaminant exposure was elevated in the samples from Valdez in 2019 when compared to samples from "clean" sites, either in western PWS Field or in KATM and LACL. The expression of these genes in the samples from the four Valdez sites was similar, in

most cases, to samples previously collected from other harbors in the PWS area. However, two exceptions were the 2019 adductor samples from the Valdez Small Boat Harbor (VH), which had higher expression for several genes associated with contaminant exposure than samples from any other sites, and the 2019 Gold Creek (GOC) gill samples, which tended to have higher transcription levels than all other sites.

These findings suggest that all four sites sampled in the Valdez area in 2019 have levels of contaminants higher than the background levels found at more remote sites in PWS, and the findings support the use of gene expression analyses in mussels as a method to monitor the presence of contaminants in Port Valdez.

Additionally, although overall patterns were similar, transcript profiles were tissue specific, to gill and adductor muscle tissue (see Figure 13). This leads us to recommend continuation of dual tissue sampling.

RECOMMENDATIONS

1. Add at least one additional site outside Port Valdez to act as a control.
2. Assess gene transcript levels from all sites at least once per year during the same season; transcript patterns in mussels have been shown to fluctuate seasonally. With longitudinal sampling (sampling from the same location at regular intervals over a period of time), comparisons among studies are not necessarily needed; trends can be identified through comparisons over time.
3. Continue to assess hydrocarbon chemistry in mussel tissues. Current levels of hydrocarbons are below the limits of detection. As per William B. Driskell, consider adding to the existing panel of hydrocarbon chemistry analyses. It is possible that the mussels are responding physiologically to a hydrocarbon not being measured.

4. Add approximately three genes to the gene panel. Choose from genes identified in Banni et al. 2017 as having transcriptional responses to B[a]P exposure:

Associated with	**Gene**	**Gene ID**	**Log2 fold change**
Mitochondrial activity	cox3_ES cytochrome c oxidase subunit 3	KF220400.1	6.38
Adhesion to substrate	fibrinogen-related protein (FREP_G3)	HQ236407.1	2.49
Phase 1 metabolism	cytochrome P450	HQ234335.1	2.46
Immune response	NF-kappaB transcription factor	HQ127223.2	-1.47
Oxidative stress response	superoxide dismutase (SOD)	FM177867.1	-1.13
Phase 1 metabolism	glutathione S-transferase (GST)	AF527010.1	-2.62
Carbohydrate metabolism	phosphofructokinase	AY580261.1	-1.96

5. Include simple mussel morphometric measurements such as length and height in future sampling.
6. Potentially include sites surrounding the Valdez Small Boat Harbor to identify extent of area impacted by Valdez Small Boat Harbor contaminants.

REFERENCES

Banni M, Sforzini S, Arlt VM, Barranger A, Dallas LJ, Oliveri C, Aminot Y, Pacchioni B, Millino C, Lanfranchi G, Readman JW, Moore MN, Viarengo A, Jha AN. 2017. Assessing the impact of Benzo [a] pyrene on marine Mussels: Application of a novel targeted low density microarray complementing classical biomarker responses. PloS ONE 12: e0178460.

Bowen L, Miles AK, Ballachey B, Waters S, Bodkin J, Lindeberg M, Esler D. 2018. 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, 147: 27-35.

Counihan K, Bowen L, Ballachey B, Coletti H, Hollmen T, Pister B. 2019. Physiological and gene transcription assays to assess responses of mussels to environmental changes. *PeerJ* 7: e7800.

Farr S, Dunn RT. 1999. Concise review: gene expression applied to toxicology. *Toxicological Sciences* 50: 1-9.

Heid, CA, Stevens J, Livak KJ, Williams PM. 1996. Real time quantitative PCR. *Genome Research* 6: 986-994.

Kadota K, Nishimura SI, Bono H, Nakamura S, Hayashizaki Y, Okazaki Y, Takahashi K. 2003. Detection of genes with tissue-specific expression patterns using Akaike's information criterion procedure. *Physiological Genomics* 12: 251-259.

McLoughlin K, Turteltaub K, Bankaitis-Davis D, Gerren R, Siconolfi L, Storm K, Cheronis J, Trollinger D, Macejak D, Tryon V, Bevilacqua M. 2006. Limited dynamic range of immune response gene expression observed in healthy blood donors using RT-PCR. *Molecular Medicine* 12: 185-195.

Poynton HC, Vulpe CD. 2009. Ecotoxicogenomics: emerging technologies for emerging contaminants. *Journal of the American Water Resources Association* 45: 83-96.