ATTACHMENT B. Annual Project Report Form (Revised 11.21.19)

1. Project Number: *See,* Reporting Policy at II (C) (1).

20170115

2. Project Title: *See,* Reporting Policy at II (C) (2).

Genomic mechanisms that underlie lack of recovery of Prince William Sound herring following the 1990's collapse

3. Principal Investigator(s) Names: *See,* Reporting Policy at II (C) (3).

Andrew Whitehead, University of California Davis

4. Time Period Covered by the Report: *See,* Reporting Policy at II (C) (4).

February 1, 2020-January 31, 2021

5. Date of Report: *See,* Reporting Policy at II (C) (5).

March 2021

6. Project Website (if applicable): *See,* Reporting Policy at II (C) (6).

https://pwssc.org/herring/

7. Summary of Work Performed: *See,* Reporting Policy at II (C) (7).

Overview: The causes of the collapse of the Prince William Sound (PWS) Pacific herring stock are controversial, and the reasons for the lack of recovery remain a mystery. In this series of projects we interrogate the genome structure and genome function of PWS fish to test hypotheses about the causes and consequences of the collapse, by revealing ecological, evolutionary, and genetic mechanisms governing the demographic trajectory of PWS fish over the past ~30 years. Conspicuous events that coincided with the dramatic PWS collapse include the *Exxon Valdez* oil spill four years previous, and the emergence of disease. We test hypotheses concerning the effects of oil exposure, the effects of disease challenge, and the potential interactive effects of oil exposure and disease challenge, on herring health and fitness. Physiological measurements and patterns of genome-wide gene expression will serve to reveal similarities and differences in mechanisms of response to these stressors between PWS and reference population fish. These studies should provide novel insights into the causes and consequences of recent dramatic demographic changes in PWS fish, potentially inform novel intervention strategies, and provide modern genomic resources for management and conservation of Pacific herring. We have performed work on three aspects of the project during this third year of the research program.

Animal exposure experiments: This year saw progress in extending work on our large multi-dose oil and pathogen exposure experiments. Results from embryonic oil exposure experiments conducted in Spring 2018 were mostly reported in the FY19 report, which included several later-life virus challenge experiments. Extended data analysis shows that we have detected oil-induced cardiotoxicity in developing herring embryos at the lowest concentrations yet reported in the literature (0.01-3.5 ug/l total PAHs). We also found that the chemical composition of bioaccumulated oil differed between the Puget Sound population and the Alaska populations, indicating population variation in adsorption, metabolism, and excretion. Furthermore, we detected population differences in sensitivity to the cardiotoxic effects of oil exposure; the Prince William Sound population was most sensitive (compared to Sitka Sound and Puget Sound populations) to cardiac developmental impairment caused by oiling, where the lowest observed effect concentration was 0.53 ug/L total PAHs.

An extension of virus challenge experiments included an additional bacterial challenge experiment. Our goal was to test whether embryonic oiling sensitizes animals to bacterial pathogens, and whether responses to bacterial pathogens differs between populations. We found that oil exposure during embryogenesis caused those fish as juveniles to be slightly more sensitive to *Ichthyophonus* challenge than fish not exposed to oil in early life (Fig. 1).

Over the past two years we developed an economical and high-throughput protocol for generating hundreds of libraries for RNA sequencing. With higher throughput and lower cost we are able to increase sample size thereby increasing statistical power, and we are able to include more treatments and levels within treatments thereby expanding the scope of our studies. We can now prepare RNA-seq libraries for 96 samples within two days, at a materials cost of approximately $10 per sample. This past year we made progress in sequencing samples from the multi-population crude oil exposure experiments. This year whole transcriptomes were sequenced from 432 samples gathered from a multi-population oil exposure experiment comprising: 1) three populations of Pacific herring; 2) five levels of oil concentrations; and 3) five developmental time-points during embryogenesis. We seek to discover detailed mechanisms underlying oil spill toxicity at very low doses, and mechanisms of oil spill exposure response that are common among, or differ between, populations. Our group has developed a custom bioinformatics pipeline to facilitate analyzing the multifactorial experimental design and large sampling effort for this oil exposure experiment. We have recently collected RNA-seq data for an additional 264 samples that include later developmental time-points from the same experimental cohort to pin-point the onset of immune system development in Pacific herring larvae. Finally, survivors from the three population oil exposure were challenged with viral hemorrhagic septicemia virus (VHSV) as juveniles. Activities are underway to sequence the transcriptomes of 246 individual Pacific herring kidneys that were extracted over the course of this experiment. Progress on these experiments have been significantly impaired by the COVID-19 pandemic. Progress requires much time at the wet lab bench, which was prohibited for much of 2020. It has slowly ramped up in the last months of 2020, but this has been slow because securing supplies like pipette tips and 96-well plates has been difficult because much of this global supply chain has been recruited for coronavirus testing. We have been forced to adjust some of our protocols to accommodate these challenges, and the work is slowly progressing.

Animal experiments over the past three years have surprised us, insofar as results have indicated that exposure to oil during embryogenesis does not seem to sensitize animals to pathogen challenge during later life. This could be because 1) oil exposure during early development does not perturb immune development/function, 2) early-life exposure to oil does perturb immune development/function but our doses were too low, or 3) early-life exposure to oil does perturb immune development/function but we have not yet exposed developing animals to oil at the effective developmental stage. To start to explore hypothesis 3, we chose to set up a set of experiments this past year that went beyond what we originally proposed. We chose to expose animals during larval development, post-hatch, for 10 days. During this period important components of fish immune systems are developing, and may be perturbed by exposures. Oil exposure has been completed, and animals have been subsequently reared in clean water. Early results indicate that the very low exposure levels (low ppb range total PAHs [tPAH]) caused significant reduction of growth (Fig. 2). Subsequent virus challenge experiments showed a surprising result, where oil-exposed animals had reduced sensitivity to virus-induced mortality (Fig. 3). Though the effect oil early life oil exposure was opposite to that expected, these data show that early life exposure to oil does affect the developing immune system, and that this effect persists until post-metamorphosis.

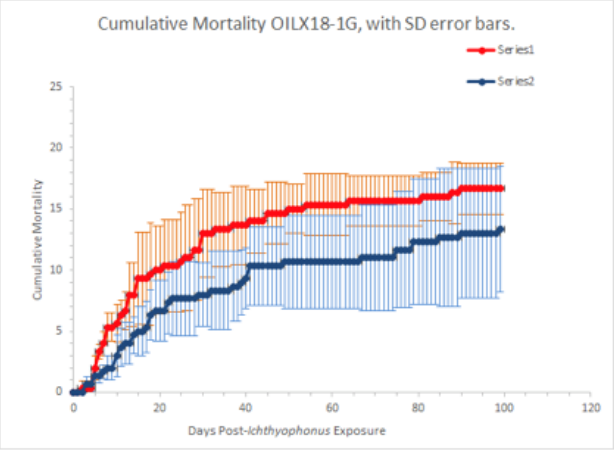

Figure 1. Survival curves for Pacific herring exposed during embryogenesis to oil (red) or control (no oil: blue) then exposed to Icthyophonus (error bars are 1 standard deviation).

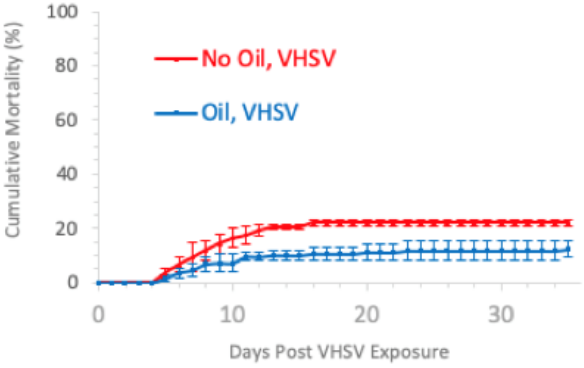

Figure 3. Oil exposure during early larval development (blue) leads to decreased sensitivity to VHSV virus-induced mortality as post-metamorphosis juveniles.

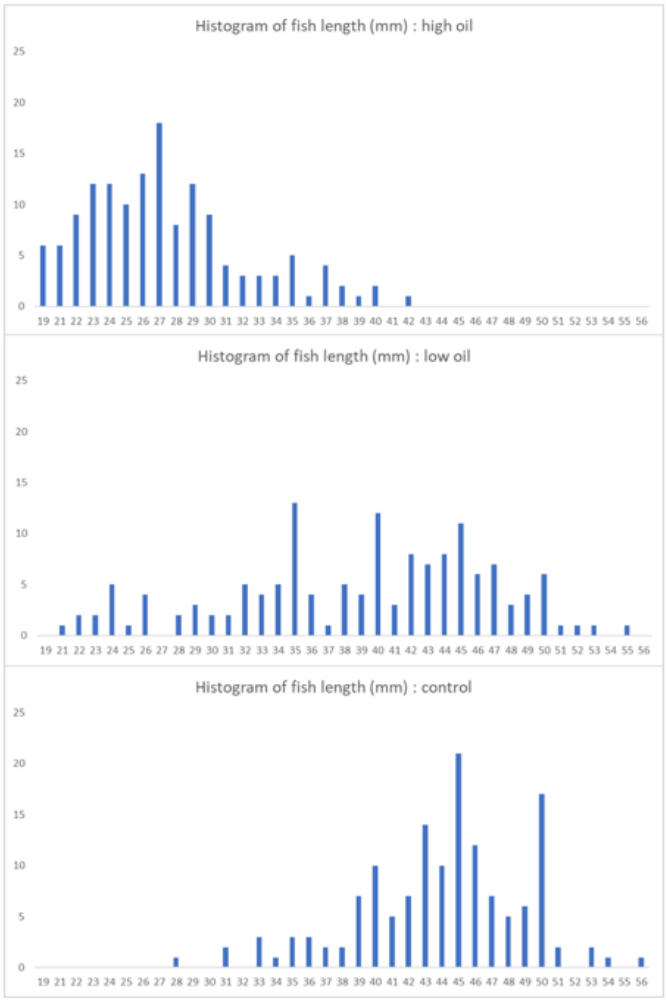

Figure 2. Histograms of fish length as a function of larval exposure to high (5.8 ppb tPAH) and low ((1.5 ppb tPAH) concentrations of crude oil and controls (no oil).

Genome sequencing and assembly: Last year we reported that a fragmented draft assembly using 10x Genomics technology has been completed, but that the long-range assembly was on hold because Hi-C libraries failed. This past year we made the difficult decision to totally re-do the genome sequence. This is because longer-read technologies had become affordable enough to justify a new approach that should lead to a much better final product. Longer reads, that enable greater assembly contiguity, were achieved by using Pacific Biosciences (PacBio) technology. PacBio reads libraries and reads were collected last year. After preliminary analysis we determined that additional PacBio read data were required for assembly, but this effort has been slowed because of the COVID-19 pandemic, since the UC Davis Genome Center is occupied with virus testing/surveillance. We also re-did the Hi-C libraries which will be used for chromosome-scale scaffolding. Preliminary results indicate that Hi-C library preparation worked this time. We have recruited the assistance of colleague Dr. Wes Warren (University of Missouri) to help with the final assembly which integrates the PacBio data with the Hi-C data. We This effort is ongoing.

Population genomics: This work was put on pause for nearly a year because postdoctoral research associate Elias Oziolor accepted a career position at Pfizer and left the University of California (UC) Davis. Elias was disappointed to leave the project, but the Pfizer opportunity was outstanding. It took some time to recruit a new postdoc to take over the population genomics work. I am very pleased to report that we have recruited Dr. Joseph McGirr who started working in the Whitehead lab in June 2020. In the past few months he has familiarized himself with the massive population genomics dataset collected by Elias. He completed variant calling started by Elias, and has progressed through all steps of quality control. He has generated summary statistics for each population, and characterized population structure using genome-wide genetic variation data. This analysis so far shows structure across geographic regions. The greatest genetic differentiation among populations distinguishes the Bering Sea population (Togiak Bay) from all other populations (Figures 4, 5). The next level of structure distinguishes southern populations (particularly California) from the northern populations. We have detected striking signatures of strong natural selection between population, including what appear to be large chromosomal inversions that appear to be segregating particularly in Southern populations (Figure 6). The important fine-scale analyses of geographic and temporal patterns between and within populations, especially Alaskan populations, are ongoing. We have algorithms working that are extracting signals of temporal covariance in allele frequency change across time. These analyses are sensitive for detecting signatures of recent natural selection among the within-site temporal samples, testing for parallel or divergent temporal patterns of selection between populations, and are sensitive for measuring the breeding population size for each population and how that has changed over very recnet timescales (e.g., over the past three decades). We are modifying our code to incorporate estimates of generation time and mutation rate in order to finalize analyses.

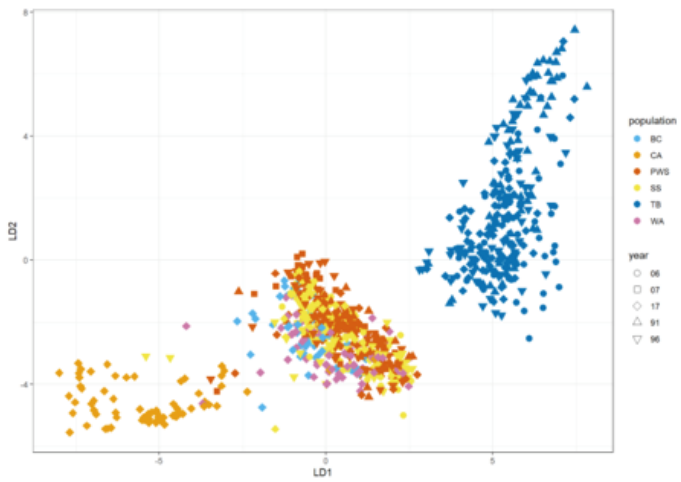

Figure 4. Two-dimensional plot of genome-wide differentiation among individuals and populations. Each point represents an individual. Colors indicate population (British Columbia, California, Prince William Sound, Sitka Sound, Togiak Bay, and Washington), and symbol shape indicate year of collection

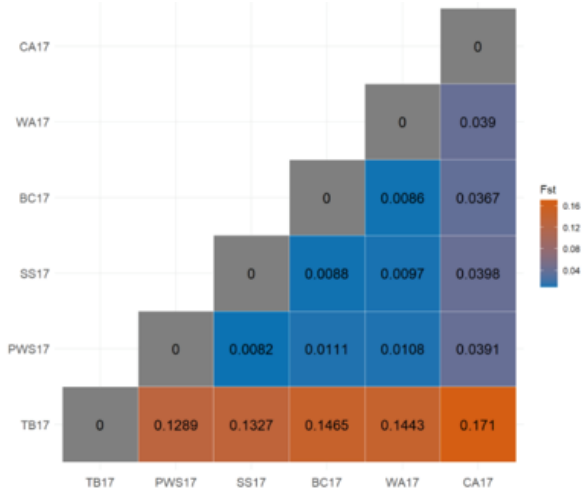

Figure 5. Pairwise measures of genetic differentiation (F_{ST}) between contemporary (2017) populations.

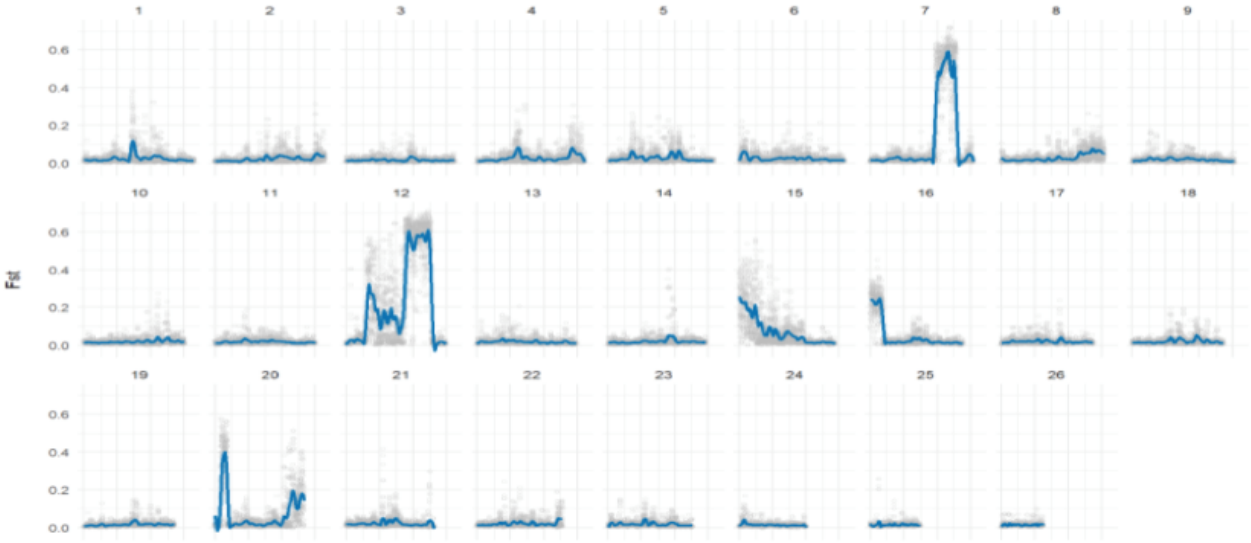

Figure 6. Genome-wide pattern of genetic differentiation (F_{ST}) between the California and Washington population, mapped for each of 26 chromosomes. Divergence tends to be low (average FST = 0.039), but with several striking regions of very high divergence (e.g., on chromosomes 7 and 12). These regions of very high divergence are consistent with chromosomal inversions that may be adaptive. Intriguingly, an inversion located in the same region of chromosome 12 is associated with temperature adaptation in Atlantic herring.

8. Coordination/Collaboration: *See*, Reporting Policy at II (C) (8).

A. Long-term Monitoring and Research Program Projects

1. Within the Program

This project is a formal collaboration with the Whitehead research group at UC Davis and that of Dr. Paul Hershberger at U.S. Geological Survey (USGS) Marrowstone (project 21120111-E). Animal

experiments described above were conducted by his group at the Marrowstone facility. Travel restrictions because of COVID-19 prevented Whitehead lab personnel from traveling to Marrowstone for meetings and experiments as usual in 2020, but this collaboration continues to be immensely fruitful

2. Across Programs

a. Gulf Watch Alaska

N/A

b. Data Management

We plan to make our data publicly available once quality control is completed. The reference genome sequence will be uploaded to NCBI. Population genomics data have been uploaded to NCBI for long-term archival, but have not yet been made public – we will make these data public once analyses are completed. We are finishing data curation for RNA-seq data and plan to upload to NCBI soon. We will then create a project page at EBI BioStudies that will include links to raw data at NCBI, and will also house variant call files for the population genomics data and matrices of read counts for the RNA-seq data. The EBI site will also house data from animal challenge experiments. All custom bioinformatics scripts will be archived at GitHub, and will be linked through the EBI BioStudies project site. So far, we have started archiving all of our scripting for population genomics work on a dedicated GitHub page which is a durable archive, and will be made publicly available once manuscript are in review. We will expand on descriptions and annotations as the project matures. Publications will also eventually be linked through the BioStudies project site. The databases described above are designed to accommodate the types of data that we need to make public, and they are durable, and they are standard practice for our research field. Once these data archives become public, we will create links to them through the Gulf of Alaska Data Portal.

B. Individual Projects

None.

C. With Trustee or Management Agencies

National Oceanic and Atmospheric Administration (NOAA) Fisheries scientists John Incardona, Nathaniel Scholz, and Alysha Cypher (NOAA Northwest Fisheries Science Center, Seattle) have been collaborating in animal exposure experiments, since they have research goals that include oil exposure impacts on growth and development. Personnel from their groups commuted on a regular basis from Seattle to the USGS Marrowstone facility during animal exposure experiments and achieved a great deal of work in collaboration with us. This collaboration has extended into activities conducted since last year's report, including coordination of tasks so that they could conduct measurements that were beyond the scope of work proposed by the Whitehead group. In particular, they contributed to detailing early-life oil exposure impacts on larval growth. They have also contributed to collection of body burden and water chemistry data. Hundreds of Pacific herring tissue samples for population genomics analysis were sent to us from colleagues at Alaska

Department of Fish and Game during the first year of the project, as recorded in previous year's reports.

9. Information and Data Transfer: *See,* Reporting Policy at II (C) (9).

A. Publications Produced During the Reporting Period

1. Peer-reviewed Publications

Oziolor, E.M., N.M. Reid, S. Yair, K.M. Lee, S. Guberman VerPloeg, P.C. Bruns, J.R. Shaw, A. Whitehead*, and C.W. Matson*. 2019. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. Science. 364: 455-457. (* co-corresponding authors)

Note: though research reported in the above publication does not include Pacific herring data, it did use methods and analyses that our group has developed over the past two years for our EVOSTC funded Pacific herring research. We therefore cite the EVOSTC funding in the acknowledgements for this paper.

2. Reports

None yet

3. Popular articles

Ph.D. student Tony Gill, with me and Paul Hershberger, wrote an article describing our research projects for the Delta Sound Connections periodical published by the Prince William Sound Science Center.

B. Dates and Locations of any Conference or Workshop Presentations where EVOSTC-funded Work was Presented

1. Conferences and Workshops

Gill, T., T. Linbo, P. Hershberger, J. Incardona, and A. Whitehead (2019) Interactions between oil exposure and immune function relevant for Pacific herring population collapse. Annual Meeting of the Society of Environmental Toxicology and Chemistry. Toronto, ON, Canada. November 2019.

2. Public presentations

None yet

C. Data and/or Information Products Developed During the Reporting Period, if Applicable

Population genomics data collection is complete. Raw sequence read data has been uploaded for long-term archive at NCBI, as per common practice in our field. Data have not yet been publicly released – we will keep data under embargo until our analyses are complete, at which time we will provide links to the data to be posted on the Gulf of Alaska Data Portal. We have started archiving our population genomics bioinformatics analysis scripts in GitHub. When script archives are completed following our analysis, we will provide links through the Gulf of Alaska Data Portal.

D. Data Sets and Associated Metadata that have been Uploaded to the Program's Data Portal

None yet

10. Response to EVOSTC Review, Recommendations and Comments: *See,* Reporting Policy at II (C) (10).

N/A

11. Budget: *See,* Reporting Policy at II (C) (11).

Budget Category:	Proposed	Proposed FY 18	Proposed FY 19	Proposed FY 20	Proposed FY 21	TOTAL PROPOSED	ACTUAL CUMULATIVE
Personnel	$77.0	$186.7	180.2	175.2	127.4	$746.5	$ 585.1
Travel	$0.0	$1.2	2.6	2.6	2.6	$9.0	$ 21.6
Contractual	$0.0	$0.0	0	0.0	0.0	$0.0	
Commodities	$54.4	$100.7	71.5	18.3	20.2	$265.1	$ 176.2
Equipment	$0.0	$0.0	50.3	0.0	0.0	$50.3	$ 51.4
Indirect Costs (*will vary by proposer*)	$74.8	$163.5	133.9	99.9	72.6	$544.7	$ 427.6
SUBTOTAL	$206.1	$452.0	$438.5	$296.0	$222.8	$1,615.6	$ 1,261.9
General Administration (9% of	$18.5	$40.7	$39.5	$26.6	$20.1	$145.4	N/A
PROJECT TOTAL	$224.6	$492.7	$478.0	$322.7	$242.9	$1,760.9	
Other Resources (Cost Share Funds)	$0.0	$0.0	$0.0	$0.0	$0.0	$0.0	