EVOSTC FY17-FY21 INVITATION FOR PROPOSALS
FY21 CONTINUING PROJECT PROPOSAL SUMMARY PAGE

Project Number and Title

21170115 Genomic mechanisms that underlie lack of recovery of Prince William Sound herring following the 1990s collapse

Primary Investigator(s) and Affiliation(s)

Andrew Whitehead, Department of Environmental Toxicology, University of California Davis

Date Proposal Submitted

August 14, 2020

Project Abstract

The long-term health of fisheries is of crucial importance for the economic health of our coastal communities and for the food security of our nation. Therefore, the causes and consequences of changes in stock abundance merit careful scientific evaluation. 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 the research proposed here 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 their potential interactive effects on herring health and fitness. We test predictions and hypotheses by reconstructing genome-wide genetic change through time (over the past 30 years) in PWS fish and compare this to population genetic change through time in two reference site populations. Furthermore, a series of laboratory-based experiments will test for population differences in their response to oil exposure in early life and subsequent resilience to pathogen exposures. 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.

EVOSTC Funding Requested* *(must include 9% GA)*

FY17	FY18	FY19	FY20	FY21	TOTAL
$224,600	$492,700	$478,000	$322,700	$242,900	$1,760,800

Non-EVOSTC Funds to be used, please include source and amount per source: (see Section 6C for details)

FY17	FY18	FY19	FY20	FY21	TOTAL
$0	$0	$0	$0	$0	$0

1. PROJECT EXECUTIVE SUMMARY

GOALS and HYPOTHESES:

Genetic attributes unique to the Prince William Sound (PWS) Pacific herring population, that either pre-existed or emerged in the years following the *Exxon Valdez* Oil Spill (EVOS), may help explain the lack of recovery in the PWS stock following the 1993 collapse, and may also illuminate the causes of the collapse. Diseases are key variables that help explain the population dynamics of PWS herring since the 1990s decline (Marty et al. 2010). The contribution of the EVOS to the PWS herring population decline is more controversial (Pearson et al. 1999, Carls et al. 2002). However, recent studies have shown that herring embryos are sensitive to fitness impacts at very low concentrations of oil (Incardona et al. 2015), these low-level exposures can affect fitness in the field (Heintz et al. 2000), and natural selection from pollutants can quickly drive complex genetic change in polycyclic aromatic hydrocarbon (PAH)-exposed populations (Reid et al. 2016). Our **overarching question** is: *Are there functional connections that link the PWS herring collapse and lack of recovery with disease impacts and the EVOS?*

Our hypotheses are:

H1: Natural selection following EVOS exposure came at the cost of compromised immune function.

H1-alternate: No evidence for oil-induced selection, but population collapse resulted in erosion of genetic diversity, especially in immune system genes, which impairs protective innate and adaptive immune responses.

H2: Exposure to oil during development compromises the ability in later life to mount an effective immune response to pathogen exposure.

To test our hypotheses, we propose the following **approaches**: 1) Retrospective population genomics; 2) Experimental/comparative physiology; 3) Experimental/comparative functional genomics.

Expected outcomes:

1. A reference genome sequence and assembly for a single Pacific herring individual will enable 21st century genetics/genomics research for this ecologically and economically important species. Similar reference genomes are available for other genome-enabled species such as humans, mice, fruitflies, and zebrafish.
2. Genetic variation is the raw material that sustains populations over time. Our careful evaluation of how genetic variation differs between populations, and how it changes through time, should serve several purposes. Erosion of genetic variation may provide an early warning signal of stock decline or collapse, allowing for intervention and prevention measures to be quickly enacted. Furthermore, genetic data may aid in stock identification, which is crucial for fisheries management. However, genetic stock identification is sometimes difficult because of lack of resolution and because genetic change through time can complicate assessments. Genome scale data offer the highest level of resolution for stock identification, so data from our studies will be the most useful genetic information for managers. Furthermore, our characterization of genetic change though time will identify regions of the genome that are not only diagnostic of stock, but also that are stable through time and thereby provide reliable diagnostic markers of stock identity.

3. These experiments should offer insight into the mechanisms whereby disease and oil exposures may affect fish health and offer insights into the sustainability of fish stocks through time.

HIGHLIGHTS from FY20 WORK:

We have made much progress this year on all three objectives of the project: 1) animal exposure experiments, 2) genome sequencing and assembly, and 3) population genomics data collection.

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. An extension of those 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. However, preliminary statistical analysis indicates that these survival curves are not statistically significantly different from each other (Fig. 1). Definitive statistical analysis is ongoing.

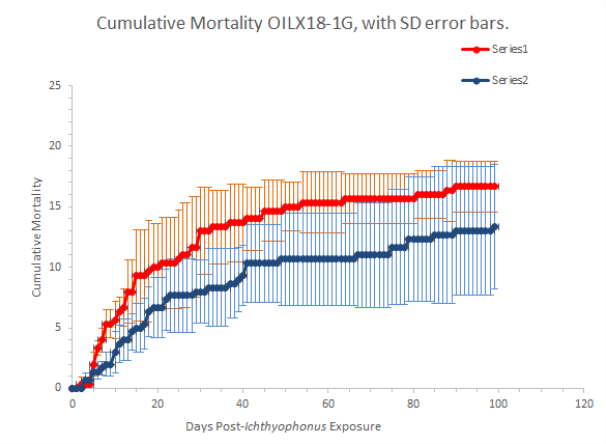

Figure 1. Survival curves for Pacific herring exposed during embryogenesis to oil (red) or control (no oil: blue) then exposed to *Icthyophonus* (error bars are 1 standard deviation).

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 therby 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 are currently generating RNA sequencing libraries 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. Plans are underway to sequence the transcriptomes of 246 individual Pacific herring kidneys that were extracted over the course of this experiment.

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

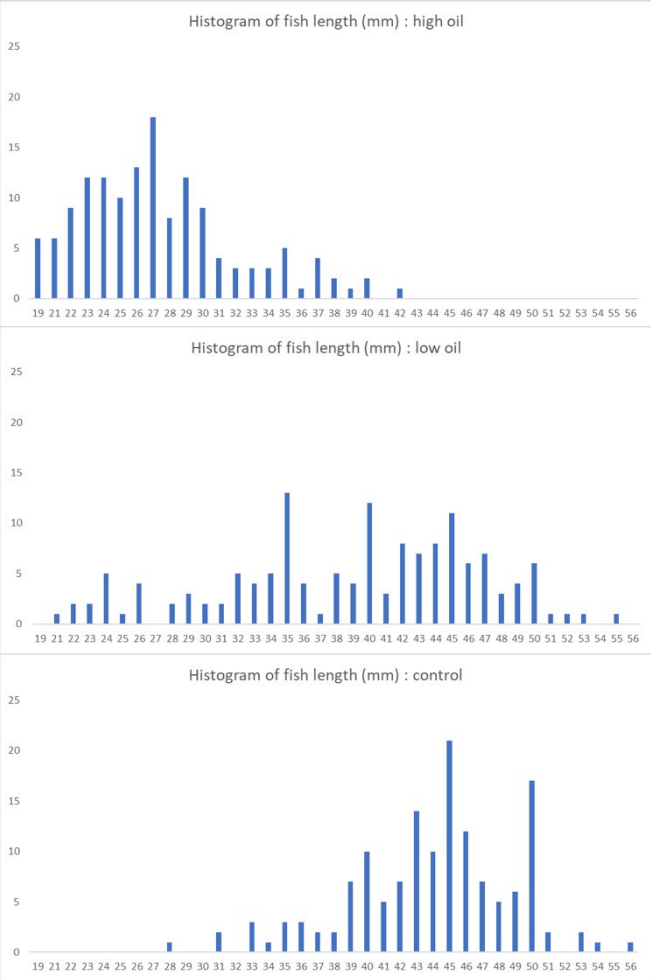

Figure 2. Histograms of fish length as a function of larval exposure to high (5.8 ppb tPAH) and low ((1.5 ppb tPAH) concentrations of crude oil and controls (no oil).

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 were exposing developing animals to oil at the wrong developmental stage. To start to explore hypothesis 3, we chose to conduct 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 perturbation of growth (Fig. 2). Virus challenge experiments are upcoming.

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. 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, and sequence data from those libraries are now in hand. We have recruited the assistance of colleage Dr. Wes Warren (Univeristy of Missouri) to help with the final assembly which integrates the PacBio data with the Hi-C data. This effort is ongoing.

Population genomics: This work was put on pause for several months 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 two 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 several steps of quality control. He is now at the stage of generating summary statistics for each population. The coarse analyses that has been completed includes tests for population structure using genome-wide genetic variation data. This coarse 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. The next level of structure distinguishes southern populations (California and Washington) from the northern populations. The more important fine-scale analyses of geographic and temporal patterns between and within populations are ongoing.

2. PROJECT STATUS OF SCHEDULED ACCOMPLISHMENTS

A. Project Milestones and Tasks

Table 1. Project milestones and task progress by fiscal year and quarter, beginning February 1, 2017. C = completed, X = planned or not completed. Fiscal year quarters: 1 = Feb 1 – April 30; 2 = May 1 – July 31; 3 = Aug. 1 – Oct. 31; 4 = Nov. 1 – Jan. 31.

	FY17				FY18				FY19				FY20				FY21			
Milestone/Task	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Sampling																				
Acquire tissue samples from ADFG for genomics and reference genome	C																			
Collect herring eggs for laboratory exposure experiments					C															
Lab Analysis																				
Extract genomic DNA	C																			
Prepare libraries for reference genome sequencing		C																		
Prepare libraries for population genomics			C	C																
Sequence reference genome			C																	
Assemble reference genome				C																
Sequence population genomics libraries					C															
Prepare libraries for functional genomics (RNA-seq) for genome annotation							C													
Laboratory exposure experiment 1: oil exposure during embryogenesis						C														
Laboratory exposure experiment 2: first pathogen (virus) challenge								C												
Laboratory exposure experiment 3: second pathogen (virus) challenge, cold water									C											
Laboratory exposure experiment 4: third pathogen challenge, cold water bacterial												C								
Prepare RNA-seq libraries for Lab experiment 1												C								
Sequence RNA-seq libraries for Lab experiment 1												C								

Milestone/Task	FY17				FY18				FY19				FY20				FY21			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Prepare RNA-seq libraries for Lab experiment 3															X					
Sequence RNA-seq libraries for Lab experiment 3															X					
Sequence RNA-seq libraries for Lab experiment 4																X				
Finish reference genome assembly and validation																X				
Data																				
QA/QC of population genomics data								C												
Exposure experiment 1: analyze chemistry data									C											
Exposure experiment 1: analyze embryo phenotype (cardiac morphology, physiology) data													C							
Exposure experiment 2: analyze survival from pathogen challenge								C												
Exposure experiment 3: analyze survival from second (cold) pathogen challenge									C											
Exposure experiment 4: analyze survival from third (cold, bacterial) pathogen challenge												C								
Upload reference genome to NCBI																	X			
Read mapping for population genomics								C												
variant calling for population genomics													C							
Annotation and comparative analysis of reference genome																X				
RNA-seq and population genomics data analysis														X	X	X	X			
Prepare and upload all data to online repositories (NCBI)																			X	
Reporting																				
Annual reports					C				C				C							
FY work plan (DPD)			C				C				C				C					
Draft FY17-21 Final Report																		X		
Conferences and Meetings																				
Annual PI meeting				C				C				C				X				X
Publications																				
Draft and submit reference genome release manuscript																	X			
Draft and submit population genomics and exposure manuscript																		X		
Draft and submit population RNA-seq manuscript																		X		

B. Explanation for not completing any planned milestones and tasks

No missed milestones.

C. Justification for new milestones/tasks

No new milestones or tasks are being proposed.

3. PROJECT COORDINATION AND COLLABORATION

A. Within an EVOSTC-funded Program

Herring Research and Monitoring

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. Furthermore, a Ph.D. student and a research technician from the Whitehead research group travelled to Marrowstone to participate in pathogen challenge experiments in March 2019. This collaboration continues to be immensely fruitful.

Gulf Watch Alaska

N/A

Data Management

Population genomics data have been uploaded to a durable public database design (European Nucleotide Archive) for the long-term archiving and public availability of these types of datasets. Although data are archived, they are not scheduled for public release until our first publication on this work is imminent.

B. With Other EVOSTC-funded Projects

None.

C. With Trustee or Management Agencies

National Oceanic and Atmospheric Administration (NOAA) Fisheries scientists John Incardona and Nathaniel Scholz (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 were commuting on a regular basis from Seattle to the USGS Marrowstone facility during 2018/2019 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 developmental defects caused by oil exposures at juvenile stages of development. They have contributed to collection of body burden and water chemistry data (e.g., Fig. 1 and 2 above). 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.

4. PROJECT DESIGN

A. Overall Project Objectives

Objective 1: Animal exposure experiments will test whether early life exposure to oil affects abilities to defend against pathogen exposure in later life. We will also test whether the PWS population differs in oil and pathogen responses compared to two other populations (Sitka Sound and Puget Sound).

Objective 2: Sequence, assemble, and annotate a reference genome and transcriptome for Pacific herring.

Objective 3: Collect retrospective population genomics data in multiple populations of Pacific herring to test hypotheses about the causes and population-levels consequences of the EVOS, subsequent epizootic, and lack of recovery.

B. Changes to Project Design and Objectives

There have been no substantive changes to our core project objectives.

5. PROJECT PERSONNEL – CHANGES AND UPDATES

Dr. Joseph McGirr has been added as a postdoc.

Joseph A. McGirr

Department of Biology
University of North Carolina at Chapel Hill
250 Bell Tower Road
2244 Genome Science Building
Chapel Hill NC, 27514
email: jmcgirr@email.unc.edu
website: fishfacegenetics.web.unc.edu
phone: (303)-257-4466

Education

2015-present	**Ph.D. Biology**, University of North Carolina, Chapel Hill
2010-2014	**B. S. Biology**, University of Colorado, Colorado Springs

Publications

In prep	**McGirr JA** and Martin CH. Ecological divergence in sympatry causes gene misregulation in hybrids.
2019	**McGirr JA** and Martin CH. Hybrid gene misregulation in multiple developing tissues within a recent adaptive radiation of *Cyprinodon* pupfishes. ***PLoS One***.
2019	St. John ME, **McGirr JA**, and Martin CH. The behavioral origins of novelty: did increased aggression lead to scale-eating in pupfishes? ***Behavioral Ecology***.
2018	**McGirr JA** and Martin CH. Parallel evolution of gene expression between trophic specialists despite divergent genotypes and morphologies. ***Evolution Letters***.
2017	Turissini DA, **McGirr JA**, Patel SS, Matute DR. The rate of evolution of postmating-prezygotic reproductive isolation in *Drosophila*. ***Molecular Biology and Evolution***.
2017	**McGirr JA** and Martin CH. Novel candidate genes underlying extreme trophic specialization in Caribbean pupfishes. ***Molecular Biology and Evolution***.
2017	**McGirr JA**, Johnson LM, Kelly W, Markow TA, Bono JM. Reproductive isolation among *Drosophila arizonae* from geographically isolated regions of North America. ***Evolutionary Biology***.

Fellowships and Awards

2018	Triangle Center for Evolutionary Medicine Graduate Fellowship.
2017	Rosemary Grant Travel Award, Society for the Study of Evolution.
2017	Best Graduate Student Presentation, SouthEastern Population Ecology and Evolutionary Genetics Conference.
2017	L.I. Gilbert Travel Award, University of North Carolina Chapel Hill.
2015	NSF Graduate Research Fellowship Program: Honorable Mention.
2014	College of Letters, Arts, and Sciences Research Award, University of Colorado Colorado Springs.

Presentations and Invited Seminars

2019 *Contributed talk.* Society for the Study of Evolution meeting. Providence, RI.

2018 *Invited speaker*. Research in Progress Seminar Series. East Carolina University. Greensville, NC.

2018 *Contributed talk.* Society for Integrative and Comparative Biology meeting. San Francisco, CA.

2017 *Contributed talk.* SouthEastern Population Ecology and Evolutionary Genetics Conference. Laurel Hill NC.

2017 *Contributed talk.* Society for the Study of Evolution meeting. Portland, OR.

Technical Skills

Computational Experience

- Proficient in R and Python.
- Comfortable with cloud computing in a Unix/Linux environment.
- Analyzed next generation sequencing datasets (whole genome, RNAseq, and CHIPseq) with industry standard pipelines and software including GATK, Samtools, VCFtools, PLINK, BWA, PICARD, and R-Bioconductor packages.
- Applied machine learning statistical methods and bioinformatic approaches to study the genetic basis of complex traits (genome-wide association mapping, differential expression, allele specific expression, and population genomic analyses).

Laboratory Experience

- Designed and performed CRISPR/Cas9 gene editing experiments in a non-model organism.
- Maintained a breeding population of tropical fishes.
- Used standard genetic protocols for DNA/RNA extractions, PCR, gel electrophoresis, and Sanger sequencing.
- Designed *in situ* hybridization probes.
- Performed immunohistochemistry experiments and visualized results with fluorescence microscopy.

Teaching and Outreach

Teaching Assistant:

- Ecology and Evolution BIOL 201
- Animal Behavior BIOL 278
- Course-based Undergraduate Research Experience BIOL 102L
- Comparative Vertebrate Anatomy BIOL 315

Journal reviewer: *Molecular Ecology, G3: Genes, Genomes, Genetics.*

Society memberships: Society for the Study of Evolution, Society for Molecular Biology and Evolution

Graduate student peer mentor

- Mentored first year graduate students entering the Biological and Biomedical Sciences Program at UNC. Helped develop departmental talks, write manuscripts, and navigate their first year at UNC as they choose thesis labs.

Contributor to the Scientific Research and Education Network

- Developed and distributed a high school lesson plan based on my research with the help of educators from North Carolina K-12 schools.

DNA Day volunteer

- Taught DNA related lesson plans in three classes (ranging 9-11 grade) at a rural high school in NC.

Darwin Day volunteer

- Annual science fair open to the public held at the North Carolina museum of natural history.

6. PROJECT BUDGET

A. Budget Forms (See HRM FY21 Budget Workbook)

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	
Travel	$0.0	$1.2	2.6	2.6	2.6	$9.0	
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	
Equipment	$0.0	$0.0	50.3	0.0	0.0	$50.3	
Indirect Costs (*will vary by proposer*)	$74.8	$163.5	133.9	99.9	72.6	$544.7	
SUBTOTAL	$206.1	$452.0	$438.5	$296.0	$222.8	$1,615.6	
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	$1,760.8	
Other Resources (Cost Share Funds)	$0.0	$0.0	$0.0	$0.0	$0.0	$0.0	

B. Changes from Original Project Proposal

N/A

C. Sources of Additional Project Funding

N/A

7. FY17-20 PROJECT PUBLICATIONS AND PRODUCTS

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 *Exxon Valdez* Oil Spill Trustee Council (EVOSTC) funded Pacific herring research. We therefore cite the EVOSTC funding in the acknowledgements for this paper.

Published and updated datasets

Sequence reads for all 1,237 Pacific herring genomes have been uploaded to the European Nucleotide Archive (Study ID PRJEB27171 (ERP109223) – data not yet released to the public)

Presentations

None to report yet

Outreach

None to report yet

8. LITERATURE CITED

Carls, M.G., G.D. Marty, and J.E. Hose. 2002. Synthesis of the toxicological impacts of the Exxon Valdez Oil Spill on Pacific herring (*Clupea pallasi*) in Prince William Sound, Alaska, USA. Canadian Journal of Fisheries and Aquatic Sciences 59:153-72. https://doi.org/Doi 10.1139/F01-200.

Heintz, R.A., S.D. Rice, A.C. Wertheimer, R.F. Bradshaw, F.P. Thrower, J.E. Joyce, and J.W. Short. 2000. Delayed effects on growth and marine survival of pink salmon *Oncorhynchus gorbuscha* after exposure to crude oil during embryonic development. Marine Ecology Progress Series 208:205-16. https://doi.org/10.3354/meps208205.

Incardona, J.P., M.G. Carls, L. Holland, T.L. Linbo, D.H. Baldwin, M.S. Myers, K.A. Peck, M. Tagal, S.D. Rice, and N.L. Scholz. 2015. Very low embryonic crude oil exposures cause lasting cardiac defects in salmon and herring. Sci Rep 5:13499. https://doi.org/10.1038/srep13499.

Marty, G.D., P.-J.F. Hulson, S.E. Miller, T.J. Quinn, S.D. Moffitt, and R.A. Merizon. 2010. Failure of population recovery in relation to disease in Pacific herring. Diseases of Aquatic Organisms 90:1-14. https://doi.org/10.3354/dao02210.

Pearson, W.H., R.A. Elston, R.W. Bienert, A.S. Drum, and L.D. Antrim. 1999. Why did the Prince William Sound, Alaska, Pacific herring (*Clupea pallasi*) fisheries collapse in 1993 and 1994? Review of hypotheses. Canadian Journal of Fisheries and Aquatic Sciences 56:711-37. https://doi.org/10.1139/f98-207.

Reid, N.M., D.A. Proestou, B.W. Clark, W.C. Warren, J.K. Colbourne, J.R. Shaw, S.I. Karchner, et al. 2016. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. Science 354:1305-1308. https://doi.org/10.1126/science.aah4993.