EVOSTC FY17-FY21 INVITATION FOR PROPOSALS
FY19 CONTINUING PROJECT PROPOSAL SUMMARY PAGE

Proposals requesting FY19 funding are due to shiway.wang@alaska.gov and elise.hsieh@alaska.gov by August 17, 2019. Please note that the information in your proposal and budget form will be used for funding review. Late proposals, revisions or corrections may not be accepted.

Project Number and Title

19170115: 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

October 01, 2018

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 (EVOS) 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 will 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
Auth:$224,703.5	Auth:$492,750.4	$477,957.4*§	$322,666.2§	$242,891.2§	$1,760,968.6

Non-EVOSTC Funds to be used, please include source and amount per source:

FY17	FY18	FY19	FY20	FY21	TOTAL

**Total in FY19 includes additional request for $50,346 – see sections 4B and 6B for further details.*

§ *Totals in FY19, FY20, and FY21 include additional requests for travel ($2,588) to the HRM Annual Meeting – see section 6B for details.*

1. PROJECT EXECUTIVE SUMMARY

Provide a summary of the program including key hypotheses and overall goals, as submitted in your original proposal. Please include a summary and highlights since your last annual report: preliminary results with figures and tables. If there are no preliminary results to present, please explain why (i.e., lab analysis is still in progress). List any publications that have been submitted and/or accepted since you submitted your last proposal and other products in *Section 7*. Prior annual reports will be appended to remind reviewers of progress in previous years.

GOALS and HYPOTHESES:

Genetic attributes unique to the PWS population, that either pre-existed or emerged in the years following the 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 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 PAH-exposed populations (Reid et al. in review-b). 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 Pacific herring will enable 21st century genetics/genomics research for this ecologically and economically important species.
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, stock identification is crucial for fisheries management, for which genetic data are one of the most important tools in the kit. 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 useful 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 FY18 WORK:

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

1) Animal exposure experiments: This year saw the completion of a massive and complex experiment, which would not have been possible without the hard and highly collaborative work between Tony Gill (Ph.D. student in the Whitehead lab), all members of collaborator Paul Hershberger's lab, partners at NOAA (Nat Scholz and John Incardona's group), and folks at the Sitka Science Center (including Angie Bowers in particular) and at Alaska DFG (including Eric Coonradt in particular). This experiment involved exposing three different populations of Pacific herring to a broad range of low concentrations of oil during development, including populations from Sitka Sound, Prince William Sound, and Puget Sound. These experiments were difficult because we had to secure gametes from spawning fish from each of these regions, transport them back to exposure facilities at Marrowstone, complete fertilizations, then conduct exposures during development. Since fish mature a different times in each of these regions, these experiments required much coordination. And since experiments have an important component of assessing population differences to oil exposure, we needed exposures to be highly reproducible between time-staggered experiments between populations that were spawning during different weeks in the early spring. Because recovering and shipping gametes live is difficult, and because the Sitka population mysteriously disappeared for several weeks, this required several trips to Alaska. Since experiments with different populations were staggered in time because of spawning time differences, we needed an exposure system that was highly reproducible. This led to our decision to buy and contract the assembly of the SINTEF Exposure System. This is a highly sophisticated state-of-the-art apparatus that can carefully control dose through a computer-controlled solenoid based dosing system. This was expensive and time consuming to order and build, but I think it was a crucial investment that will pay dividends in the reproducibility of experiments, and ultimately enabling our ability to rigorously compare populations in their developmental response to oil exposure. The animals from these exposure experiments have hatched and are currently in their grow-out phase. Pathogen exposure experiments with these animals are planned for September/October 2018.
2) Genome sequencing and assembly: I am very pleased to report that we have completed the first draft of a reference genome assembly for Pacific herring. There is some additional work to be completed before we have a final assembly, but in the current state the assembly is sufficient to serve as a mapping reference to support the population genomics read mapping and RNA-seq read mapping. We prepared libraries compatible with the 10X Genomics technology. We received the sequencing data in early spring. We have been working through iterative assemblies, and the current assembly is of very high quality. N50 scaffold sizes are greater than 2 Gb, which is above average contiguity compared to other fish genomes. We anticipate even higher contiguity once we complete additional tasks this year, including scaffolding from Hi-C libraries, and ordering of scaffolds through recombination mapping. The recombination map will be generated using RAD-seq data from a family of fish that we spawned and archived in the spring of 2018. We will be starting the genome annotation effort this fall 2018. Sequencing of a reference transcriptome will contribute to this annotation. Reference transcriptome sequencing has been completed. We used one of the latest technologies – Pacific Biosciences IsoSeq – to

generate full-length RNA sequences from a library of mixed adult and early life stage tissues. This sequencing and assembly is complete. This has resulted in ~900,000 high-quality full-length transcript sequences. This will be a crucial resource for genome annotation as a mapping reference for the quantitative transcriptomics that we will perform during FY19 on samples generated from FY18 animal experiments outlined in the previous section. A description of the reference genome, including assembly strategy with the reference transcriptome, and annotation, will be packaged into a manuscript. We anticipate that this will be submitted for publication within one year from now.

3) Population genomics: Last year we chose a different sequencing strategy than we had originally proposed. Rather than pooled-sample sequencing, we decided to individually index and sequence samples. This is because the cost-per-sequence versus sequence data coverage math had come out in favor of individually-indexed whole genome sequencing with the publication of a new library preparation protocol. We decided that this strategy would offer higher quality data for the same cost as originally proposed. This required us to get a new protocol up and running in the lab. This required three months of troubleshooting. We finally got this up and running reliably in June 2018. We then created libraries for all ~1,200 samples within one month. The libraries were submitted to the core facility for sequencing in July 2018. QC was quickly completed, and the samples progressed into the queue for sequencing. We anticipate that sequencing will be completed within the next few weeks, likely before the end of August. This will result in a huge dataset; including 20 lanes of Illumina Hi-Seq 4000 paired-end 150 base sequence. This should yield over 2,000 gigabases of new sequence! Once sequence data have arrived, we are prepared to initiate sequence QC, read mapping, and sequence variant calling. Sequence data will then be made public by uploading to NCBI. The population genomics data analysis will then be started.

2. PROJECT STATUS OF SCHEDULED ACCOMPLISHMENTS

A. Project Milestones and Tasks

Milestones are annual steps to meet overall project objectives. For each milestone listed, specify the status (completed, not completed) when each was completed and if they are on schedule, as submitted in your most current proposal.

Tasks are annual steps to meet milestones. Specify, by each quarter of each fiscal year, when critical tasks (for example, sample collection, data analysis, manuscript submittal, etc.) were and will be completed.

Please identify any substantive changes and the reason for the changes. *Reviewers will use this information in conjunction with annual program reports to assess whether the program is meeting its objectives and is suitable for continued funding.*

B. Explanation for not completing any planned milestones and tasks

Please identify any substantive changes and the reason for the changes. If tasks were not completed as scheduled or delayed, please explain why and the anticipated completion date.

C. Justification for new milestones and tasks

Please identify any new milestones and tasks and the reason why they have been added.

A. Project Milestones and Tasks

Project milestone and task progress by fiscal year and quarter, beginning February 1, 2017. Yellow highlight indicates proposed fiscal year Work Plan. Additional milestones and tasks may be added. C = completed, X = not completed or planned. Fiscal Year Quarters: 1= Feb. 1-April 30; 2= May 1-July 31; 3= Aug. 1-Oct. 31; 4= Nov. 1-Jan 31.

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
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															
Laboratory exposure experiments						C														
Extract RNA from laboratory exposure experiments							C													
Prepare libraries for functional genomics (RNA-seq)							C													
Finish reference genome assembly and validation								X												
Sequence RNA-seq libraries										X										
Data																				
QA/QC of population genomics data								X												
Analyze developmental physiology and disease data from exposure experiments								X												
Upload reference genome to NCBI								X												
Read mapping and variant calling for population genomics									X											
Annotation and comparative analysis of reference genome									X											
RNA-seq and population genomics data analysis													X	X	X	X	X			
Prepare and upload all data to online repositories (NCBI)																			X	
Reporting																				
Annual reports					C				X				X							
FY work plan (DPD)			C				X				X				X					
Draft FY17-21 Final Report																		X		
Conferences and Meetings																				
Annual PI meeting				C				X				X				X				X
Publications																				
Draft and submit reference												X								

genome release manuscript																			
Draft and submit population genomics and exposure manuscript																	X		
Draft and submit population RNA-seq manuscript																	X		
New Milestone																			
New task																			

B. Explanation for not completing any planned milestones and tasks

We have nothing to report here, as the project is on schedule.

C. Justification for new milestones and tasks

There are no new milestones and tasks planned.

3. PROJECT COORDINATION AND COLLABORATION

A. Within an EVOTC-Funded Program

Provide a list and clearly describe the functional and operational relationships with any EVOSTC-funded Program (Herring Research and Monitoring, Long-Term Research and Monitoring or Data Management Programs). This includes any coordination that has taken or will take place and what form the coordination will take (shared field sites or researchers, research platforms, sample collection, data management, equipment purchases, etc.).

B. With Other EVOSTC-funded Projects

Indicate how your proposed project relates to, complements or includes collaborative efforts with other proposed or existing projects funded by the EVOSTC that are not part of a EVOSTC-funded program.

C. With Trustee or Management Agencies

Please discuss if there are any areas which may support EVOSTC trust or other agency work or which have received EVOSTC trust or other agency feedback or direction, including the contact name of the agency staff. Please include specific information as to how the subject area may assist EVOSTC trust or other agency work. If the proposed project requires or includes collaboration with other agencies, organizations or scientists to accomplish the work, such arrangements should be fully explained and the names of agency or organization representatives involved in the project should be provided. If your proposal is in conflict with another project, note this and explain why.

3.A. Starting in FY19, this project will be part of the HRM program to help facilitate increased interaction and collaboration with the HRM PIs and also the GWA program. This project is a formal collaboration with my research group at UC Davis and that of Dr. Paul Hershberger at USGS Marrowstone. Animal experiments described above were conducted by his group at the Marrowstone facility. Furthermore, a Ph.D. student from my research group travelled to Marrowstone to participate in pathogen challenge experiments in September 2017, and will again this year. Members of Dr. Hershberger's group also collected and shipped gametes from Prince William Sound in March 2018 for animal exposure experiments. They also spawned and grew out one family of herring to be used for genetic mapping. This collaboration has been immensely fruitful.

3.B. N/A

3.C. 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. These measurements are easily added to tasks associated with our (Whitehead and Hershberger) planned and ongoing exposure experiments, so it made sense to join forces and leverage the collective experience, personnel, and equipment of our groups – the most important of these being the outstanding aquatics systems for rearing live herring at Marrowstone. Personnel from their groups were commuting on a regular basis from Seattle to the Marrowstone facility during 2018 animal exposure experiments, and achieved a great deal of work in collaboration with us. In particular, they contributed to detailing developmental defects caused by oil exposures at multiple stages of development. This requires much expertise and time at the microscope. Gamete collections this year, in preparation for animal experiments, were fraught with difficulties, especially from Sitka Sound since the population mysteriously disappeared for several weeks. Successful collections required three trips to Alaska. During the last trip, my student Tony Gill was hosted at the Sitka Science Center for two weeks. Several key people selflessly assisted with daily reconnaissance trying to find spawning adults. Particularly helpful were Angie Bowers (Sikta Science Center) and Eric Coonradt (Alaska DFG). We are deeply grateful for their energy and commitment.

4. PROJECT DESIGN

A. Overall Project Objectives

Identify the overall project objectives for your project as submitted in your original proposal.

B. Changes to Project Design and Objectives

If the project design and objectives have changed from your original proposal, please identify any substantive changes and the reason for the changes. Please include the revised objectives in this section. Include any information on problems encountered with the research or methods, if any. This may include logistic or weather challenges, budget problems, personnel issues, etc. Please also include information as to how any problem has been or will be resolved. This may also include new insights or hypotheses that develop and prompt adjustment to the project.

4.A.

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 Prince William Sound 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 *Exxon Valdez* oil spill, subsequent epizootic, and lack of recovery.

4.B. There have been no substantive changes to our core project objectives. There have been a few changes in some protocols. For example, our population genomics sequencing strategy has changed. We originally proposed pooling samples from each population for sequencing. We since decided to individually index samples in each population. This was previously out of reach financially, but new protocols for library preparation, with high efficiency and low-cost reagents, made this feasible. Data collection was set back approximately three months because the new protocol required some considerable troubleshooting. However, since this portion of the project was previously ahead of schedule, adoption of this protocol has not set us back.

Sitka Sound fish mysteriously disappeared for several weeks during the spring spawning season which caused some problems with collecting live gametes from Sitka Sound for animal experiments this year. Successful collections required three trips to Alaska. During the last trip, my student Tony Gill was hosted at the Sitka Science Center for two weeks. Several key people selflessly assisted with daily reconnaissance trying to find spawning adults. Particularly helpful were Angie Bowers (Sikta Science Center) and Eric Coonradt (Alaska DFG). We are deeply grateful for their energy and commitment.

We initially planned to use oiled-gravel columns for animal exposure experiments, since they are one of the standard exposure systems used in oil exposure research, especially for many previous herring studies. These were available for use through our project partners at NOAA. However, since population contrasts were a core part of our research objectives, and since the availability of gametes from each population was staggered with differences in seasonal timing of spawning, we needed a highly reproducible dosing system. If we could not precisely reproduce dose ranges for staggered experiments, then robust population contrasts would have been difficult, and would have posed challenges for meeting core project objectives. In our recent collective experience, and upon consultation with NOAA collaborators, we were concerned about the precise reproducibility of doses using the oiled-gravel column system. Additionally, we had evidence that responses to the dose range that we were using were non-linear. That is, FY18 experiments indicated that very low dose exposures actually enhanced later life immune responses, perhaps by priming the immune system, whereas higher doses are expected to be immunotoxic. When dose responses are non-linear, then comparisons across dose response curves (e.g., among populations) are statistically difficult, making the nature of population differences difficult to discern, thereby affecting our ability to interpret our data. As such, our population contrasts required precisely reproducible doses across the range. For these reasons we made the difficult decision to purchase a state-of-the-art electronically controlled dosing system. Only one company in the world makes these systems: SINTEF (Norway). Expert colleagues in oil spill research confirmed that these are highly reliable systems. Upon very short notice SINTEF agreed to custom build and assemble this system in time for our definitive spring 2018 multi-population exposure experiments. Initial quality control tests demonstrated that the system was capable of extremely precise reproducibility. With this system we were able to complete a definitive set of exposure experiments with all populations during the 2018 spawning season, meeting project objectives. This system is now core research infrastructure that will support additional carefully controlled dosing experiments. Without this new state-of-the-art controlled dosing system, it is possible that project objectives would not have been met.

5. PROJECT PERSONNEL – CHANGES AND UPDATES

If there are any staffing changes to Primary Investigators or other senior personnel please provide CV's for any new personnel and describe their role on the project.

There are no staffing changes to senior personnel.

6. PROJECT BUDGET FOR FY19

A. Budget Forms (Attached)

Provide completed budget forms.

B. Changes from Original Proposal

If your FY19 funding request differs from your original proposal, provide a detailed list of the changes and discuss the reason for each change.

C. Sources of Additional Funding

Identify non-EVOSTC funds or in-kind contributions used as cost-share for the work in this proposal. List the amount of funds, the source of funds, and the purpose for which the funds will be used. Do not include funds that are not directly and specifically related to the work being proposed in this proposal.

6A. Attached

6B. Revised October 2, 2018.

i. To facilitate collaboration with the HRM Program and as per discussions with the HRM program and PI this project will be part of the HRM program starting in FY19; this proposal is revised to include travel costs to the annual HRM PI meeting.

ii. The Science Panel noted that PI had identified unexpected differences in the seasonal timing of spawning from each population, thus requiring the use of electronically-controlled oil dosing equipment for these experiments to be highly reproducible. The Panel understood and agreed with the PIs decision to purchase the needed equipment ($50,346). The PAC discussed that prior similar studies would have been strengthened by use of this equipment. Noting the need for high tech equipment in genetics work, the PAC recommended the additional funding.

6C. N/A

7. FY18 PUBLICATIONS AND PRODUCTS

Products include publications (include *in prep* and *in review*), published and updated datasets, presentations, and outreach during FY18.

None to report yet.