CPB talk

Pacific herring is a coastal schooling species found on the eastern and western sides of the Pacific. Schools of mature adults migrate into estuaries to breed once a year where spawning occurs on kelp and eelgrass. A single female can produce up to 20,000 eggs. After those eggs hatch, it takes about three to four years for the herring to mature and join up with other adults in large schools. At each stage of their life cycle, herring are an important food source, and because many species in a complex food web depend on Pacific herring and their eggs, it is considered keystone species.

One part of that complex food web includes humans. This is artwork by April White, a member of the Haida indigenous group from the coast of British Colombia where herring have been culturally important for centuries. The prints show many of the species that depend on herring inside a herring’s stomach, including a human down here.

Another way that herring have been important to humans is in studies of population genetics. There has been a lot of this type of research on Atlantic herring, which are about 3 million years diverged from Pacific herring. This is because they come close to meeting some assumptions that population geneticists like to make regarding infinite population size and random mating. They are one of the most abundant vertebrates on earth and they reproduce in these really large spawning groups. As a result of all of this research, there are very nice resources for Atlantic herring, including a chromosome level assembly and a known mutation rate. Despite their huge census size, estimated to be around 1 trillion, herring show nucleotide diversity that is pretty middle of the road for vertebrates and they have a pretty small effective population size.

This makes them a extreme example related to Lewontin’s Paradox, which suggests that, under neutral theory, larger populations should contain more genetic variation than what is observed. For those of you that have seen Vince Buffalo’s new preprint, this is where herring fall on the spectrum with other organisms, with really moderate genetic variation despite a huge population size. Lucky for me, this is part of what’s made Vince interested in the system and we’ve been collaborating on a lot of the work I’ll present today.

OUTLINE

Herring are a forage fish, and forage fish typically go through boom and bust population cycles, even in unstressed populations. These plots show changes in herring biomass for two populations off the coast of Alaska. You can see that population sized decreased in the early 90s and then recovered shortly after. However, this population in Prince William Sound collapsed and numbers have remained very low for the past 28 years. This collapse is exceptional across all herring populations on record, with PWS showing the most extreme collapse here. And there was a major event that likely contributed to this collapse.

In the spring of 1989 the Exon Valdez oil tanker spilled 11 million gallons of crude oil into Prince William Sound. Oil covered the coastline at the height of herring spawning season, which is really bad news for herring eggs that are laid right near the shore. Like I mentioned earlier, it takes about 3-4 years for newly hatched herring to mature and join the large schools of adults in the open ocean. We can see here that just a few years after the oil spill, herring biomass dropped by 75%.

Even though it is still hotly debated to this day, it seems pretty clear that the oil spill contributed to the collapse. However, what is still a mystery is why the population hasn’t recovered after almost three decades. Every other fish population that initially declined has since recovered, so why not herring?

Well, one thing we know is that herring have been struggling with disease outbreaks. There are two main pathogens causing issues, Viral hemorrhagic septicemia (VHSV) and Ichthyophonus hoferi (fittingly referred to as Ick). These pathogens have plagued the population since just after the oil spill. But these pathogens are not unique to Prince William Sound. Other nearby populations have likely been dealing with them for just as long. Given that the oil spill *was* unique to Prince William sound, we think there’s an interaction here between disease and oil exposure.

One of the ways we are thinking about this interaction comes from previous work with Atlantic killifish populations that have adapted to highly polluted waters in the Northeast United States. In this figure, populations are labeled T if they are pollution tolerant and labeled S if they are sensitive to pollution, measured by survival across a range of experimental exposures to polychlorinated biphenyl (PCB) concentrations, a compound that is often formed as waste in industrial processes. Tolerant populations were found to have large deletions spanning genes that are part of the aryl hydrocarbon receptor pathway. These deletions likely desensitize these populations to pollution exposure, but the fact that these deletions are rarely found in sensitive populations suggest there might be a trade-off. AHR signaling has diverse functions and interacts with multiple pathways, including estrogen and hypoxia signaling, regulation of cell cycle, and, importantly, immune system function. Overall this work shows that rapid adaptation to pollutants can favor loss-of-function variants that impair the function of signaling pathways important for healthy immune systems.

Another way we are thinking about population recovery is related to genetic load, which we can think about as a measure largely determined by a reduction in fitness contributed by deleterious mutations across the entire genome. Bottlenecks can increase genetic load because drift is stronger, and bottlenecks cause larger increases in genetic load when population size recovers slowly, like we see in Prince William Sound.

So keeping all of this in mind, I want to test a number of different hypotheses that are all related to this main question: Are there genetic factors limiting population recovery in Prince William Sound?

To get at this question, we have low coverage sequencing for 1100 Pacific herring. This is our pipeline to identify genetic variation limiting population recovery. I want to give a couple of quick acknowledgements. All libraries were prepared by Jenn Roach and Elias Oziolor before I joined the lab. This was a ton of work, with over 20 plates of DNA sequenced. I also want to thank Vince Buffalo for analyses he performed using his cvtk workflow.

So we have low coverage sequencing about 1x coverage. I ran sequences through this variant calling pipeline to calculate genotype likelihoods, and used these filters to generate a final data set including 200000 SNPs.

Here are the sample sizes we are working with. In red is Prince William Sound where the oil spill occurred, and for Togiak Bay in blue, we have samples from 1991,1996,2006, and 2017. All samples come from adults, meaning that the 1991 fish were born before the oil spill, giving us the ability to compare allele frequencies pre, and post oil spill. We can think of Togiak bay as our healthy reference population that was never exposed to oil and maintains a healthy population size. We also have samples that span the Pacific coast from Alaska to California, although most populations were sampled at a single time point in 2017.

Here is what population structure looks like, with the Togiak bay population separating from the NorthEast Pacific populations on PC1 and the California population separating out on PC2.

So I want to use this genomic data set to test some hypotheses. This first one is related to population collapse. READ.

If we measure effective population size using theta = 4Nemu we see that Ne is most variable and a little higher in PWS in red compared to the nearby healthy populations in blue and yellow. However, this is a long-term measure of effective population size that is related to coalescence time in generations. If we expect that the decline in population size happened within the past few decades, the then number of generations between samples is very low, and much lower than Ne. So we shouldn’t expect to be able to see signal of a decline in Ne using this long term measure.

But, We can also look at a short-term measure of Ne that can be estimated when you have temporal sampling across generations. We can measure the variance in temporal changes in allele frequency across sampling periods, and this measure is inversely proportional to effective population size. Using this approach, we see a relatively stable population sizes in the healthy populations in blue and yellow, but an overall decline in population size in PWS between 1991 and 2017. The contemporary effective population size in PWS is similar to the other nearby healthy populations.

So relating this back to our hypothesis, the short-term estimates of Ne suggest a decline in population size in Prince William Sound while the reference populations remained stable, and this is consistent with surveys measuring a decline in census population size in all three populations. I want to keep this in mind as I move onto hypotheses related to slow population recovery.

Given what we know about the atlantic killifish populations, rapid adaptation to pollutants can favor genetic variation that impairs the function of signaling pathways important for healthy immune systems. Specifically the AHR signaling pathway. In Pacific herring it is possible that genetic variants that conferred resistance to oil exposure immediately after the spill, came at the cost of compromised immune function. In other words, variants that helped individuals survive the oil exposure are now deleterious because they disrupt the ability to mount an effective immune response. Finally, given that those deleterious alleles increased in frequency after the oil spill.

To test this I characterized loss of function alleles within each population using an Ensembl tool called variant effect predictor. I am assuming that particular loss-of-function alleles are deleterious because they are predicted to disrupt gene function. I input a bed file with reference and alternate alleles, and the tool searches across 24000 genes that are annotated for the atlantic herring reference genome, and reports how many of my alleles would cause different categories of coding mutations. Here I’m showing the number of stop gained, stop lost, and start lost mutations that would likely disrupt gene function.

Here I’m showing the allele frequencies for each of these predicted mutations across all populations. As you can see, there wasn’t much difference between populations, particularly between PWS and nearby healthy populations in blue and yellow. Furthermore, there wasn’t a big difference between pre- and post- oil spill samples in prince William sound.

So these results do not suggest that population recovery is limited by an overall increase in genetic load. We don’t see a general increase in the frequency of deleterious alleles in Prince William Sound. However, what if there are specific mutations that are important, like we see in the killifish system? If a gene is important for surviving oil exposure *and* disease exposure as we predict, then that gene region should show signs of fluctuating selection.

The key to this hypothesis is that selection pressures in PWS changed from oil exposure to disease exposure, while selection pressures were presumably more stable in nearby populations. So next I wanted to look for regions of the genome that showed shifts in allele frequencies consistent with this change in selection pressures.

In other words, I want to look for regions of the genome where allele frequency shifts changed direction, and I’m going to do that by measuring this delta Z statistic across genomic windows. So for example, if there was this minor allele that increased in frequency between 1991 and 1996, and then decreased again between 1996 and 2006, we would see a positive delta z in this comparison and a negative delta z in the later comparison.

Since AHR signaling genes were important in the Killifish system, I first looked for frequency shifts at those genes. These are two genes in that pathway, aip and arnt, and we can see that minor allele frequencies increased between 1991 and 1996 around the arnt gene in Prince William Sound, but decreased in the healthy Togiak Bay population. Then between 1996 and 2006 we see that allele frequencies changed less drastically in both populations.

So these are just two genes on chromosome 20. If we zoom out and look at allele frequency shifts genome wide between 1991 and 1996, we see these two huuuuge dips on chr 7 and 12. Then if we look between 1996 and 2006, we see a reversal in the direction of allele frequency change. Now I’ll zoom in on this chromosome, chromosome 7. We see significant shifts in the healthy Togiak bay population in blue too, but we see that the reversal only occurs in PWS.

So these results do show genomic evidence consistent with fluctuating selection pressures in Prince William Sound. In the future we want to pair this analysis with measures of temporal covariance in allele frequency change to identify regions under selection. Under drift, we expect random changes in allele frequencies across generations, so that the covariance in allele frequency change is around zero. Under directional selection we expect a positive covariance. This is still very much a work in progress, but Vince just sent me some results on Friday that look very interesting. We see…

What’s really cool is that inversions were recently discovered in southern populations of Atlantic herring too. Atlantic herring and pacific herring are about 3 million years diverged. One of the inversions occurs at the same region on chromosome 12! In both systems, inversions are present in populations in warmer waters. As you move south, herring tend to be

READ

So its possible that you have a bunch of genes that influence some of these traits, and they are all clumped together on an inversion, the whole cluster can be inherited as a large supergene. The inversion suppresses recombination so that linkage between those adaptive genes in maintained over time. Its possible that this is a common genetic mechanism of adaptation in marine species where gene flow is frequent across long ranges.