Fish Out of (Salt) Water:
Adaptation of an Ocean-Going Fish to Freshwater Environment

by
Janna R. Willoughby[1,2], Avril M. Harder[2], Jaqueline M. Doyle[3]

Part I – Natural Selection

As biologists we are interested in how populations change over time. In particular, we often study how a change in one species leads to change in another species. For example, we might study how a loss of a particular prey item changes predator behavior as natural selection favors behaviors suitable for capture of alternate prey. Natural selection is a core concept in evolution, and it provides a mechanism for evolution that is testable and observable. *Evolution by natural selection occurs via differential survival and reproduction of individuals with varying phenotypes. In order for natural selection to bring about phenotypic change in a population, phenotypic differences among individuals must be at least partly due to differences in the alleles they have inherited from their parents. This means that there must also be varying genotypes.*

In our example of changing predator behavior, natural selection might occur when individuals with a trait suited for catching novel prey helps those individuals survive longer and have more offspring compared to individuals that lack that trait. In this case, we would expect that natural selection would result in an increase in the number of individuals with the beneficial trait over time. Over a species' evolutionary history, individuals with a beneficial trait pass that trait to their offspring more often than individuals with a less beneficial trait are able to pass the less beneficial trait to their offspring. As a result, population-wide allele frequencies also shift over time. (In this scenario, "alleles" are simply different forms of the same gene associated with the phenotypic trait.) Although it is possible to observe natural selection by tracking the change in phenotypes, modern genomic approaches often rely on large genetic data sets and bioinformatics to track the changes in alleles (gene forms) over time to identify natural selection in action.

For this case study you will assume the role of a scientist investigating patterns of natural selection in wild populations. In order to prepare for these investigations, you need to develop some background understanding of the processes.

Questions

1. Natural selection acts on phenotypes within a population. If, for example, you want to know if the muscle mass trait is acted upon by natural selection in female lions, what three things must be true about muscle mass in lionesses? (Think about what must be true of a particular gene for natural selection to occur.)

---

[1] Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN.
[2] Department of Biological Sciences, Purdue University, West Lafayette, IN.
[3] Department of Biological Sciences, Towson University, Towson, MD.

2. Suppose that having more muscle mass leads to longer survival time in female lions, relative to lions with less muscle mass. In the plot below, add some point data to illustrate how this trend might occur in a real population.
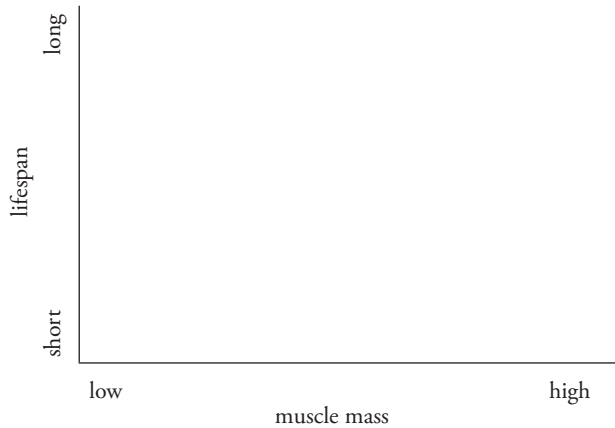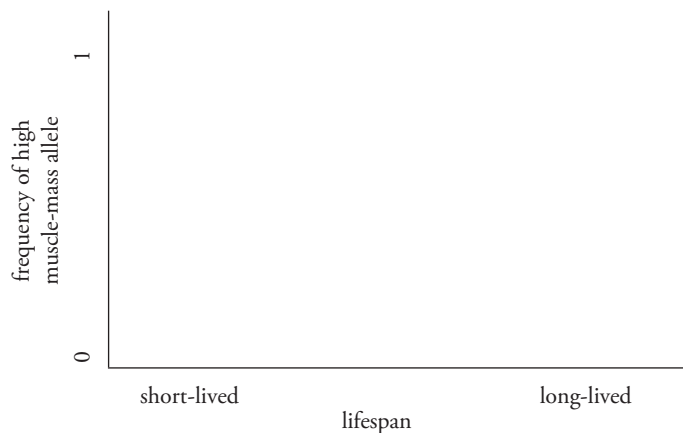

Figure 1. Female lion *(Panthera leo). Credit:* Charles J. Sharp, Sharp Photography. CC BY-SA 4.0, <https://tinyurl.com/y7uwo6ma>.

(Plot: y-axis labeled "lifespan" ranging from "short" to "long"; x-axis labeled "muscle mass" ranging from "low" to "high".)

3. Although we can track changes in phenotype, natural selection is often detected by changes in genotype. Imagine that allele A is associated with high muscle mass and that high muscle mass results in longer lifespans. (In this example, at least one other allele for this gene would be found at this locus, but we will ignore these other alleles in this example.) In the plot below, draw a bar-graph to predict what you might observe if you measured how often this allele was found in short-lived lionesses compared to long-lived lionesses. Would the frequency of the allele be the same in both categories?

(Plot: y-axis labeled "frequency of high muscle-mass allele" ranging from 0 to 1; x-axis labeled "lifespan" with categories "short-lived" and "long-lived".)

4. Assume that the muscle mass allele is correlated with survival in female lions. If you evaluated the whole genome, should you expect to find alleles at other genes that are also correlated with survival? If yes, where might those alleles be located or what might the function of those genes be? If no, why should we expect no other signals?

Part II – Steelhead Trout Introduction

Steelhead trout, which are native to the North American Pacific coast, have been widely introduced elsewhere in North America and across the globe. Approximately 120 years ago, a small number of steelhead were released into the Great Lakes where they established naturally reproducing populations. In their native range, the maturation phase of the steelhead lifecycle occurs in the ocean (salt-water), whereas in the Great Lakes steelhead must mature in freshwater. Therefore, we might expect that selection will have resulted in some changes to the Great Lakes steelhead populations that allow them to exist in such a different environment. Your charge as a scientist is to investigate how the introduction event changed the steelhead population and how the steelhead have adapted to the new environment.

Figure 2. Steelhead trout *(Oncorhynchus mykiss)* caught near Salmon, Idaho. *Credit:* Rydon Ricks, CC BY 2.0, <https://tinyurl.com/yb7np7h9>.

To demonstrate how introductions influence genetic diversity, you will conduct a simple experiment considering the effects on only one locus. For this experiment, assume that the progenitor population (i.e., the original source population) has ten alleles at the locus of interest and that all alleles are present at the same frequency. We will simulate the release of four individuals (two alleles each) at a new location.

Questions

1. Make a hypothesis about the effects of introduction on genetic diversity. How many of the ten alleles present in the original progenitor population do you think will be represented in the introduced population? (This should be a guess not an experiment yet!)

2. Using a phone, tablet, or computer, navigate to the page <http://www.dave-reed.com/Nifty/randSeq.html> (or scan the QR code pictured to the right) that will generate random letters, which we will use as labels for different alleles. Set the "Number of random letter sequences to generate" to "8," since you will be introducing four individuals with two alleles each. Next, specify that you want the length of each random letter sequence to be "1," and then finally type in ten different letters that will represent each of the ten alleles found in the progenitor population. The program is now ready to simulate which of the ten different alleles will be included in the new population when four individuals (carrying a total of eight alleles) are released.

Random Letter Sequence Generator

Number of random letter sequences to generate: `8` ← # introduced individuals * 2 = total number of alleles in new population

Length of each random letter sequence: `1`

Letters to choose from: `abcdefghij` ← alleles present in progenitor population

On the next page, record the number of different alleles you would have in your introduced population for ten different trials. For example, if the random letters you generate for Trial 1 are a, a, b, c, d, d, e, and f, then you would record having six different alleles in your sample.

Trial	Number of Different Alleles	Trial	Number of Different Alleles
1		6	
2		7	
3		8	
4		9	
5		10	

3. Looking at the table, what is the average number of alleles present in the introduced population under this scenario? Does this seem like an adequate representation of the variation present in the progenitor population?

4. Using the letter generator, how many individuals (each carrying two alleles) would you need to reliably introduce 90% of the genetic variation (number of alleles) present in the progenitor population? (You should try a few values and replicate the test.)

5. In reality, we do not have very good estimates about the number of steelhead released into the Great Lakes, or the size of the progenitor population at the time. However, we do know that there was an approximate 10% reduction of genetic diversity in the introduced population. Is this reduction larger or smaller than you would have guessed, and why?

6. What effect might loss of genetic diversity have on the effectiveness of natural selection in the steelhead population? Using this idea, make some predictions about the speed of genetic adaptation possible compared to the number of individuals introduced into a novel environment. In other words, how might the number of individuals introduced into a novel environment influence the potential rate of adaptation?

Part III – Identifying Evidence of Selection

We know that steelhead trout successfully adapted to life in freshwater because they have established populations that exist in freshwater. Your next task is to identify what traits were important for survival in freshwater that were not as important in the salt water habitat. First, make hypotheses about what you expect to find and then analyze the genomic data to find interesting regions.

The figure below depicts some of what we know about fish physiology in salt and freshwater habitats. These adaptations are largely focused on how individuals deal with too much or too little salt and involve many organ systems.
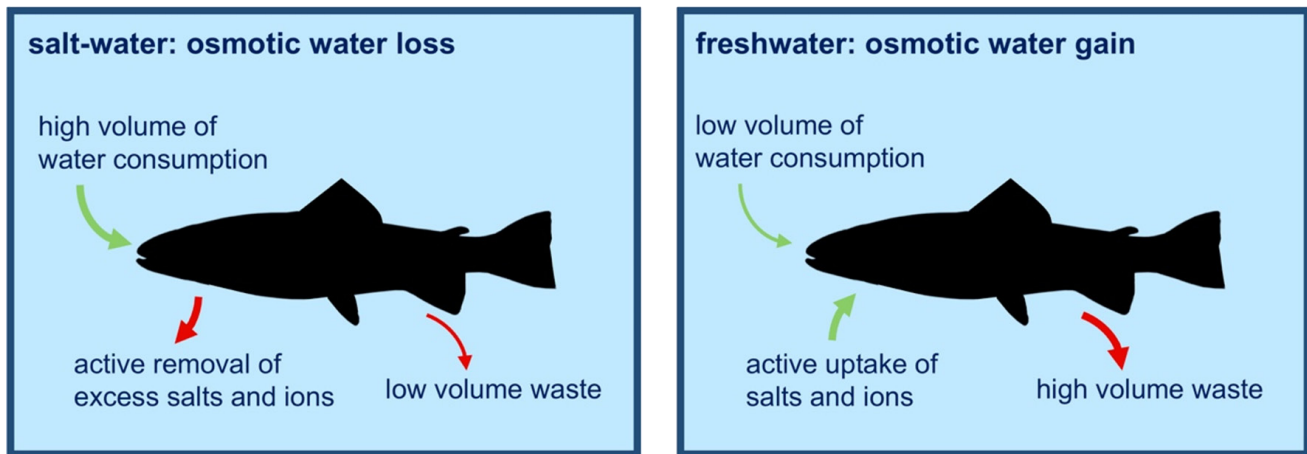

Figure 3. Fish physiological function in oceanic (salt water) and freshwater systems. In general, fish remove more salts and drink more water in salt water compared to freshwater, to compensate for the loss of water that occurs via diffusion. In freshwater, fish absorb water via diffusion and so do not drink as much water. However, because salts are rare in the freshwater environment, fish in that environment will actively take up salts (and other ions) from the water. Modified from Willoughby *et al.* (2018).

Questions

1. What are some biological systems or traits that you expect to be different in the freshwater steelhead compared to the oceanic steelhead?

2. Thinking about the system or trait identified above, and assuming we know which genes control that trait in the genome, what is one way you could compare the salt and freshwater steelhead populations that would help us understand if that trait was different in the two populations?

One value used to understand how populations differ across the genome is F_{ST}. This statistic compares the allele frequency of a locus in one population to the allele frequency of that locus in another population. Generally speaking, if two populations have similar allele frequencies then the F_{ST} value for that locus will be small. If allele frequencies differ between the populations, F_{ST} will be larger. Below is a figure of F_{ST} estimates made for 20,000 loci across all 29 chromosomes in the steelhead genome, comparing an oceanic steelhead population to a freshwater steelhead population. In this instance, F_{ST} has been statistically transformed to $Z(F_{ST})$, but this does not change how you interpret the trends.
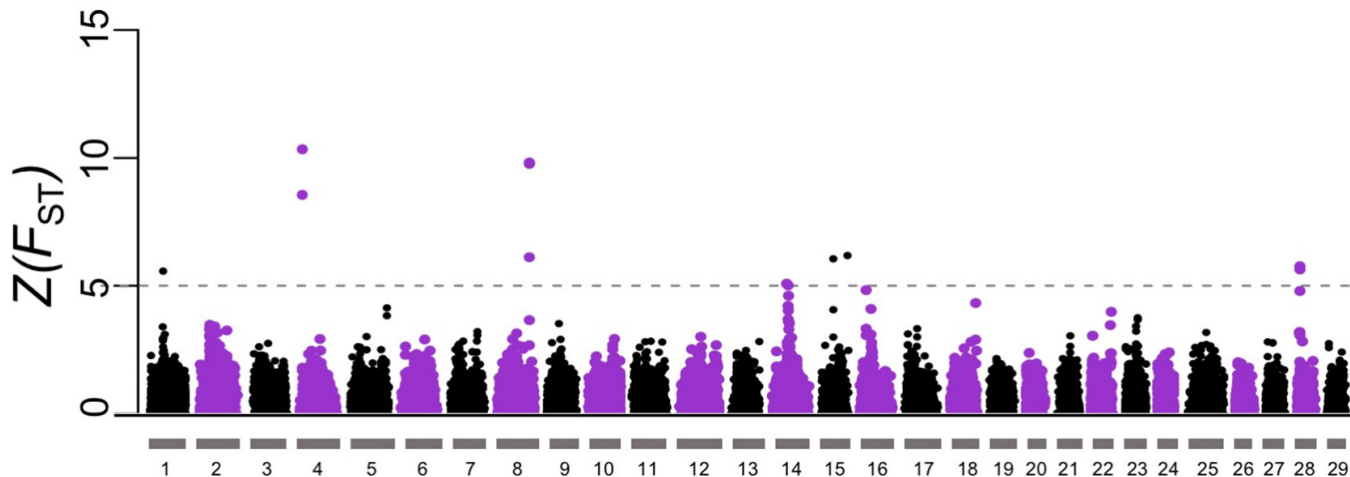

Figure 4. Estimates of $Z(F_{ST})$ for all 29 chromosomes (numbered across the x-axis), comparing oceanic steelhead to freshwater steelhead. For these estimates, F_{ST} was averaged across 100 kilobase-pair regions (1 kilobase-pair = 1000 base-pairs), in order to find regions of the genome that were very different between the two populations. Each point represents the mean $Z(F_{ST})$ across a 100 kilobase-pair region, and the dotted line notes 5 standard deviations from the mean $Z(F_{ST})$. Modified from Willoughby *et al.* (2018).

Questions

3. The figure above contains nearly 20,000 individual points. Look at the figure and think about which points are different from the others. Which points stand out? In reference to the allele frequencies of the two populations, what does it mean to have an F_{ST} that stands out from the average value?

4. What do you think the outlier points mean in terms of adaptation to the freshwater environment for the steelhead population?

Part IV – Annotating Adaptive Regions

To get an idea of the function of the regions of the genome that are different between the oceanic and freshwater steelhead, you are next going to consider the regions with high $Z(F_{ST})$ on chromosomes 8 and 28. In Figure 4, you evaluated the mean $Z(F_{ST})$ across a very large region of the genome; Figure 5 zooms in on two smaller regions. Consider each base position (called a SNP, or single nucleotide polymorphism) along the chromosome whose allele frequencies are different between the two populations. In the plot below, each point represents a single SNP, and F_{ST} compares the allele frequencies of that SNP in the oceanic and freshwater populations.
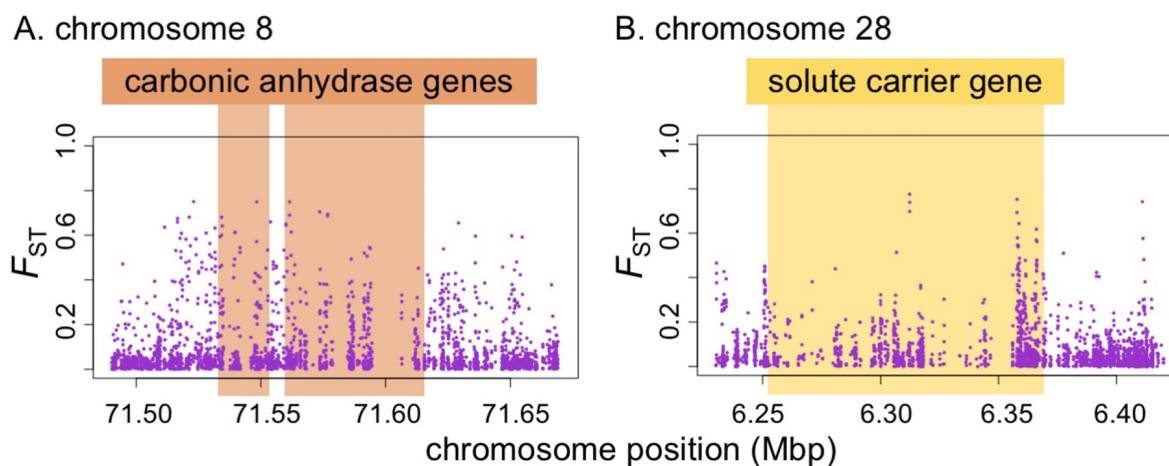

Figure 5. Comparison of oceanic and freshwater steelhead using F_{ST} estimates. Panels *A* (left) and *B* (right) correspond to two of the regions with high $Z(F_{ST})$ identified in Figure 4. The location of genes in these regions are noted by the shaded rectangles, and gene functions are noted at the top of each gene region. Modified from Willoughby *et al.* (2018).

Questions

1. What do the high F_{ST} SNPs within and near the carbonic anhydrase and solute carrier genes mean about these genes in the two steelhead populations?

2. Can you come up with any hypotheses about the function of these genes? Even if you don't know anything about these genes in reality, think about how you expect the populations to be different due to the different conditions in which the populations exist.

Part V – Interpreting Adaptation Mechanisms

The functions of these particular genes are complementary; both are related to osmoregulatory (salt-balancing) and pH balancing (within the body) functions and operate in specific cells called mitochondria-rich cells. These cells are typically concentrated in the gills of fish. Within these cells, carbonic anhydrases work to break down CO_2, which builds up in the bloodstream during respiration. The resulting ions are excreted as waste into the environment directly from the gills. Using some of the ions created by carbonic anhydrases, solute carrier proteins can retrieve needed salts from the environment and move them into the gills. Working in tandem, these structures balance salt concentrations and pH within the body.
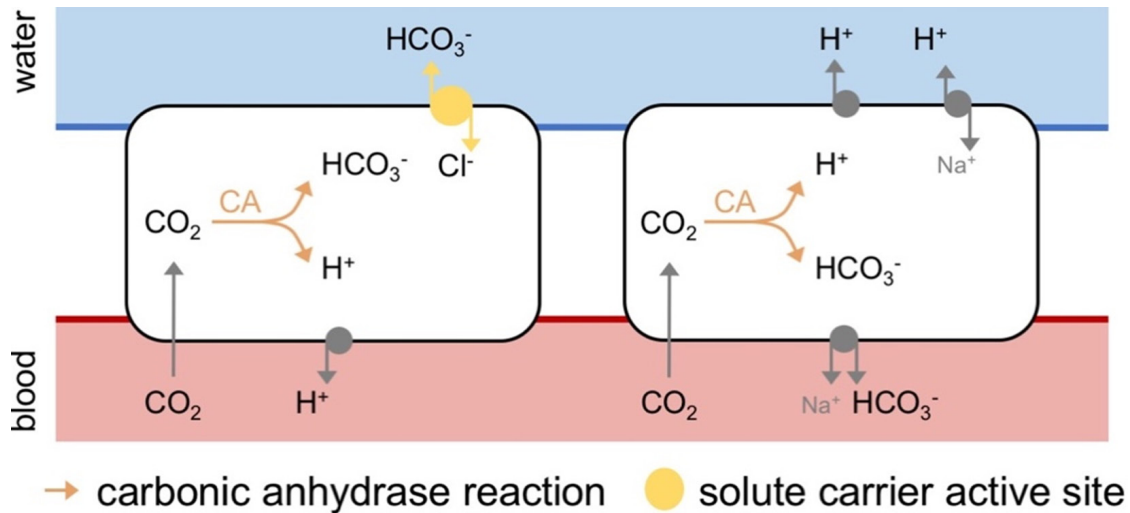

Figure 6. Illustration of carbonic anhydrase and solute carrier protein function, within the mitochondria-rich cells found in fish gills. Modified from Willoughby *et al.* (2018).

Questions

1. Now think about what you would expect to happen to the high F_{ST} SNPs if the freshwater-adapted steelhead were moved to the salt-water environment. How might the high F_{ST} SNPs change under these conditions?

2. Now that you've formed a hypothesis for what will happen if you release freshwater-adapted steelhead into a salt-water environment, think about how likely this is to occur. Would you expect the freshwater-adapted fish that were released into salt water to be able to cope with the new, salty environment? Why or why not?

Further Information

Natural selection occurs at all times. Read the following post for additional interesting examples of natural selection:

- BEACON. 2012. Evolution 101: natural selection [blog]. <https://www3.beacon-center.org/blog/2012/10/01/evolution-101-natural-selection/>.

Shane Campbell-Staton is a biologist that studies evolution. He took advantage of a severe cold event that occurred in Texas to understand how green anole lizard populations react to extreme temperature fluctuations. You can see him talk about his work on YouTube:

- *Shane Campbell-Staton: Harvard Horizons Symposium* [video]. Running time: 6:19 min. <http://youtu.be/FLcLHyk-18k>.

References and Additional Resources

Axelsson, E., A. Ratnakumar, M.-L. Arendt, K. Maqbool, M.T. Webster, M. Perloski, …, and K. Lindblad-Toh. 2013. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495: 360–4.

Carlson, S.M., C.J. Cunningham, and P.A.H. Westley. 2014. Evolutionary rescue in a changing world. *Trends in Ecology and Evolution* 29: 521–30.

Campbell-Staton, S.C., Z.A. Cheviron, N. Rochette, J. Catche, J.B. Losos, and S.B. Edwards. 2017. Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. *Science* 357: 495–8.

Willoughby, J.R., A.M. Harder, J.A. Tennessen, K.T. Scribner, and M.R. Christie. 2018. Rapid genetic adaptation to a novel environment despite a genome-wide reduction in genetic diversity. *Molecular Ecology* 27(20): 4041–51. DOI: 10.1111/mec.14726.