Possible Discovery of a New Species of *Fundulus* (Lacépède, 1803) in the New York Harbor

Max Feldman[1], Mauricio Gonzalez[1], Louise Bodt[2]

[1] New York Harbor School, Marine Biology Research Program; [2] Urban Barcode Project and American Museum of Natural History, Educator

Introduction

This study was motivated by intriguing results found in the mitochondrial non-coding control region in a prior completed project focusing on Killifish. The samples were collected around the New York Harbor in the summers of 2011,2012,2018, and 2019. In 2018 and 2019, the mitochondrial non-coding control region was analyzed from these samples providing insight into the population and maternal genetics of the fish (Lee, 1995). Population genetics in fish provide insight into how specific populations are affecting the general ecosystemic dynamic (Allendorf, 1979; Okumuş, 2003). Results showed around 4-10 base pair mismatches in certain samples in the mitochondrial non-coding control region. Results also showed BLAST hits on *Fundulus* species that are native to freshwater environments in the midwest led to a new genetic region being sequenced to identify these samples of *Fundulus*. The mitochondrial COI barcoding region was analyzed for this project to determine whether or not a new species had been described based on base pair mismatches in the sequence (Ames, 2006). In this project, the barcoding COI region was sequenced from already extracted DNA that was being stored in the American Museum of Natural History. It is hypothesized that the barcoding COI region sequences of the unidentified samples will show enough base pair mismatches to describe a new species of *Fundulus*.

Background Information

As a generalist predator in the Hudson-Raritan Estuary, the Killifish plays an important role in an ecosystem as prey for certain organisms such as White Perch and Blue crab, and predators for organisms such as grass shrimp and snails (Mugue, 1995). This makes the Killifish a good candidate for analysis since it has a large impact on an ecosystem's general dynamic. Each site where samples were collected (Fig. 1) (Jones Beach, Inwood Hill Hark, Jamaica Bay, and Pelham Bay) had varied habitats (shores, bays/oyster reefs, and estuaries) and an abundance of Killifish. The present author focused on Killifish for this project because it's an abundant tidal fish that is found along the east coast, and in the Hudson River estuarine (Mugue, 1995).

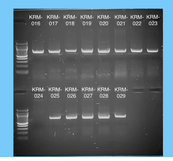
Scientific Problem

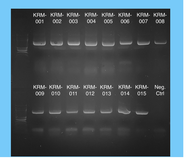
Are there enough base pair mismatches in the Barcoding COI region to describe a new species of *Fundulus*?

Methods

01) Fish samples collected using seine nets, umbrella traps, and minnow traps from 4 sites around NY harbor (Fig. 1), 02) placed into buckets on ice (5 min.), 03) placed in conical tubes (99% ethanol), 04) DNA extraction in lab using Qiagen Blood and Tissue kit (Qiagen, 2013-21), 05) ran electrophoresis to verify extraction (Figs. 2 & 3), 06) ran PCR protocol using fish primer (DNA Learning Center, 2021), 07) amplified DNA targeting barcoding COI region, 07) sample sequenced by GENEWIZ, 08) sequence analysis conducted in DNA Subway (30 total samples were sequenced for this project), 09) dedrogram created.

Figure 2 and 3: Gel electrophoresis





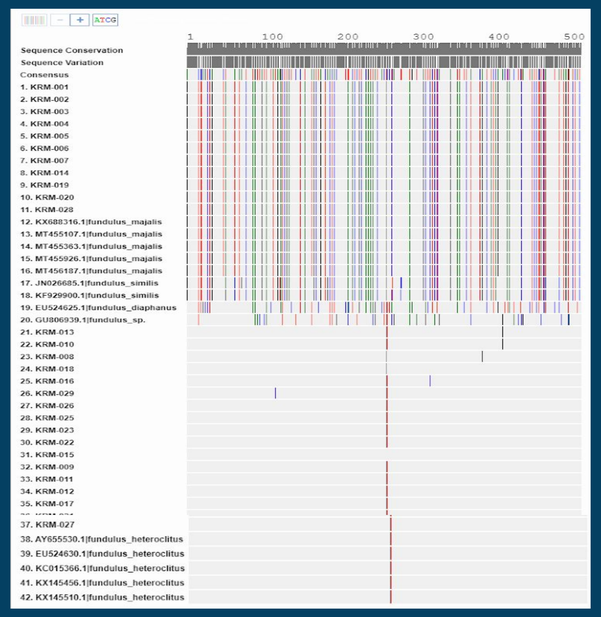


Figure 4: Multiple Sequence Comparison by Log-Expectation.
In this figure all the sequences are aligned with *Fundulus majalis* sequences from GenBank.

Results

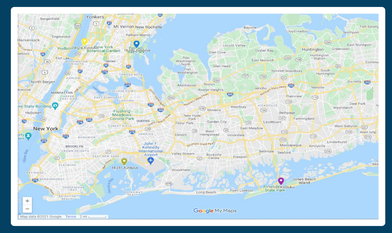


Figure 1: Locality map where samples were collected around the New York Harbor.
See Google Maps (2021) pin:

Figure 5: Maximum Likelihood Phylogenetic Tree.
Samples are grouped based on their BLAST hits. The group in the top left is *F. majalis* and the other is *F. heteroclitus*. The other groups are outliers. The orange arrow represents the *F. majalis* group and the purple arrow represents the *F. heteroclitus* group. The red arrow represents the outlier groups at the bottom of the tree

Results

In the Multiple Sequence Comparison by Log-Expectation (MUSCLE) (Fig. 4), all of the sample sequences were aligned with *Fundulus majalis* GenBank sequences. 11 out of the 28 successful sequences were identical to previous COI barcoding sequences of *F. majalis* (Group A) from GenBank. The other 17 successful sequences were either identical or had 1-4 base pair mismatches with previous *F. heteroclitus* COI barcoding sequences (Group B). In the maximum likelihood phylogenetic tree above (Fig. 5), the samples were grouped based on what species of *Fundulus* each sample's COI barcoding sequence aligned with. The majority of the samples had BLAST hits with *F. heteroclitus*.

Analysis

Based on the results, all the samples have either been identified as either *F. heteroclitus* or *F. majalis*. This rejects the hypothesis that there would be enough base pair mismatches in the barcoding COI region to describe a new species of *Fundulus*. It's important to note that the 1-4 base pair mismatches (Fig. 5) for the *F.* heteroclitus sequences wasn't significant enough to not identify the samples as *F. heteroclitus*. Also, in the phylogenetic tree (Fig. 5) there were outlier groups that further supported the idea that each sample was either *F. heteroclitus* or *F. majalis*. In addition, both *F. heteroclitus* and *F. majalis* appeared in the 2011-2012 and in 2018-2019 samples. This is important because if there were only *F. heteroclitus* in 2011-2012 and only *F. majalis* in 2018-2019 then many questions would arise about what happened to the *Fundulus* population between 2012-2018. Also, the base pair mismatches in the mitochondrial non-coding control region for certain samples may have been due to there being possibly no non-coding control regions in *F. majalis* sequences published (Louise Bodt, pers. comm., 2021).

Conclusions

These data were valuable since they helped identify the unknown samples of *Fundulus* from the 2018-2019 project. In addition, these data explain why there were base pair mismatches in the mitochondrial non-coding control region in certain samples from the 2018-2019 project. In the future, if a similar issue occurs (base pair mismatches in the mitochondrial control region in *F. heteroclitus* or other species of *Fundulus*), the sequences from the present project will help fill in the missing genetic database sequences (based on the 2018-2019 project's results). Finally, these mitochondrial non-coding control sequences might be some of the first, if not the first, sequences of their kind for *F. majalis* samples.

References

Allendorf, F. W., & Utter, F. M. (1979). Population genetics. *Fish physiology*, *8*, 407-454.
Ames, C., Turner, B., & Daniel, B. (2006). The use of mitochondrial cytochrome oxidase I gene (COI) to differentiate two UK blowfly species–Calliphora vicina and Calliphora vomitoria.
DNA Learning Center Barcoding 101. (n.d.). https://dnabarcoding101.org/lab/protocol-3.html.
Forensic science international, 164(2-3), 179-182.
(Lee, W. J., Conroy, J., Howell, W. H., & Kocher, T. D. (1995). Structure and evolution of teleost mitochondrial control regions. Journal of Molecular Evolution, 41(1), 54-66.)
Nikolai Mugue, JS Weis, et al. (Dissertation), Population Genetics of Fundulus Heteroclitus in the Hudson River and North New Jersey Estuaries: Evaluation of Subspecies Boundary and Hybridization with F. diaphanus
Okumuş, İ., & Çiftci, Y. (2003). Fish population genetics and molecular markers: II-molecular markers and their applications in fisheries and aquaculture. *Turkish Journal of Fisheries and Aquatic Sciences*, *3*(1), 51-79.
QIAGEN. (n.d.). https://www.qiagen.com/us/product-categories/top-sellers/?cmpid=PC_DA_NON_qiagen-shop_1220_SEA_GA&gclid=Cj0KCQjwvYSEBhDjARIsAJMnolhg8dLqntvygdTckEdy3TXNXzid9Pt6SKXR1TceazR6dlYkmQi92a8aAm-IEALw_wcB.

Acknowledgements

Thank you to Mauricio Gonzalez, Louise Bodt, Noah Burg, Bobby Habig, Brian Levine, and everyone else who reviewed my work. Also, thank you to the American Museum of Natural History and the Urban Barcode Project for making this project possible. Finally, a big thank you to The Pinkerton Foundation, Cold Spring Harbor Laboratory, and any other organization that helped fund these projects.