5-2021

Mitochondrial Replacement and Mitonuclear Interactions in the Broadstripe Topminnow (Fundulus euryzonus)

Auburn Ansley

Mitochondrial Replacement and Mitonuclear Interactions in the Broadstripe Topminnow (*Fundulus euryzonus*)

by

Auburn Ansley

A Thesis
Submitted to the Honors College of
The University of Southern Mississippi
in Partial Fulfillment
of Honors Requirements

May 2021

Approved by:

Brian Kreiser, Ph.D., Thesis Advisor,
School of Biological, Environmental and Earth Sciences

Jake Schaefer Ph.D., Director,
School of Biological, Environmental and Earth Sciences

Ellen Weinauer, Ph.D., Dean
Honors College

ABSTRACT

When species hybridize, mitochondrial and nuclear introgression can take place. This is commonly observed in freshwater fish species, such as the *Fundulus notatus* complex, which is composed of *F. notatus*, *F. olivaceus*, and *F. euryzonus*. The broadstripe topminnow, *F. euryzonus*, is only found in the Amite and Tangipahoa Rivers in the Lake Pontchartrain drainage, where it overlaps in distribution with *F. olivaceus*. Previous studies determined that *F. euryzonus* in the Tangipahoa River system possesses mitochondrial DNA (mtDNA) from *F. olivaceus* as a result of hybridization between the two species. The goal of this thesis was to re-examine an existing data set of single nucleotide polymorphisms (SNPs) to determine if this mitochondrial discordance has manifested itself in changes in the nuclear genome between populations of *F. euryzonus* in the Amite and Tangipahoa Rivers. This was conducted using a sliding window analysis to look for genome regions with very high (>0.95) or low (<0.01) weighted fixation index (F_{ST}) values and further analyzing regions that included three or more windows in a series. Sliding window groups were assigned to scaffolds within the genome of *F. heteroclitus*, and the roles of the genes were determined using the UniProt knowledge base. Sixty-nine sliding window groups were identified, which contained a total of 403 annotated genes belonging to 12 functional groups. Ten genes were found to be associated with the mitochondria, and four of these genes (NMES1, UQCRC1, NDUFB4, and SMIM20) were specifically related to the electron transport chain. Given the functional significance of these genes, it is suggested that natural selection could be acting upon *F. euryzonus* in the Tangipahoa River as a consequence of mitonuclear discordance.

Keywords: Fundulus, GBS, mitonuclear discordance, natural selection

ACKNOWLEDGMENTS

I would like to thank my amazing advisor, Dr. Kreiser, for remaining patient with me and understanding throughout this entire process. You helped me tremendously by carefully explaining how to conduct the research and by always being available whenever I needed help. Although I was very busy this semester with my classes, work, leadership positions, and preparation for dental school, you were persistent, yet always kind. Despite my initial delay in the research process, you trusted me to complete the thesis in time, and I really appreciate that.

Thank you to Stephen Flanagan for your research and information in your dissertation from 2016. The samples from your dissertation were instrumental in my research, and I truly could not have done this research without your work. I gained a lot of knowledge about mitonuclear introgression from reading your research, and I applied much of this knowledge when writing my thesis. Also, thank you to everyone that was involved in the process of collecting and analyzing the samples from Flanagan's research.

To my family, thank you for always supporting me in my endeavors and encouraging me to be the best person that I can be. Although you did not have the slightest clue of my research, I greatly appreciate you being there for me to complain about all of my school work. To Brandon, thank you for always motivating me and telling me how proud you are of me. You have no idea how much your kind words meant to me and how it helped me get through this process. To the Honors College staff, thank you for always being a support system and believing in me from the beginning of my college experience at Honors Day to now: completing my thesis, graduating summa cum laude from USM, and being accepted into the UMMC School of Dentistry Class of 2025.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
ATP	Adenosine Triphosphate
DNA	Deoxyribonucleic Acid
F_{ST}	Fixation Index
GBS	Genotype-by-Sequencing
HGT	Horizontal Gene Transfer
mtDNA	Mitochondrial DNA
nDNA	Nuclear DNA
OxPhos	Oxidative Phosphorylation
SNP	Single Nucleotide Polymorphism
USM	The University of Southern Mississippi

CHAPTER I: INTRODUCTION

Mitochondria are membrane-bound organelles that generate the majority of the chemical energy that is needed to power eukaryotic cells. Mitochondria produce adenosine triphosphate (ATP), which cells use as a source of energy for biological functions (Rank et. al., 2020). Each mitochondrion has its own genetic material, which is referred to as mitochondrial DNA (mtDNA) (Nass & Nass, 1963), and there can be multiple copies of the same genetic material (Wiesner et. al., 1992). Because mitochondria have their own DNA, it is suggested through the endosymbiotic theory that mitochondria were originally free-living organisms that developed symbiotic relationships with eukaryotes and became an internal component of the cell (Wallin, 1923; Sagan, 1967).

Once mitochondria integrated into the cell, horizontal gene transfer (HGT) took place, where mitochondrial genes found their way into the host cell's genome. This ultimately led to the mitochondria becoming obligate endosymbionts of their eukaryotic host cell (Blanchard & Lynch, 2000). The extent of HGT from mitochondrial to the nuclear genomes varied between lineages (e.g., plants and animals). The loss of genes from the mitochondria produced a more streamlined genome that allowed it to replicate faster and more efficiently (Cavalier-Smith, 1987).

Coevolution, where natural selection leads to reciprocal genetic changes (Futuyma, 1998), should be expected between the mitochondrial and nuclear genomes since mitochondrial maintenance and function require components of each genome. Of particular importance is the Oxidative Phosphorylation (OxPhos) pathway. The enzymes of this pathway generate most of the cellular ATP. In animals, this requires interactions

between all 13 mitochondrial proteins and an additional 80 proteins from the nuclear genome (Rank et. al., 2020). The vital role of mitochondrial genes in energy production leads to a strong selective pressure against deleterious mutations, and numerous studies have documented evidence of purifying selection acting on mitochondrial genes (e.g., Morales et al., 2015; Pavlova et al., 2017; Bernardo et. al., 2019). However, in some cases positive selection may reflect adaptation to different environmental or physiological conditions (Pavlova et al., 2017). The types of selection that act on mtDNA are expected to function the same way on the nDNA genes that interact with the mitochondria. This type of co-evolution has been termed mitonuclear functional compensation, and this mechanism is sometimes invoked to explain patterns of mtDNA discordance such as in hybrid zones (Bernardo et. al., 2019).

When closely related species come back into contact and hybridize, their offspring can express proteins that are not coevolved. First generation hybrids would have an equal contribution of nuclear genomes, both maternally and paternally, whereas the mitochondrial genome would only be inherited maternally. Hybrid breakdown refers to reproductive failures that occur in the offspring of crosses between different species or subspecies as a consequence of incompatible genes (Oka et. al., 2004). Impacts on offspring fitness due to this mitonuclear discordance have been broadly documented. One such example of hybrid breakdown is found in monkey flowers (*Mimulus guttatus complex*), where conflicting cytoplasmic and nuclear backgrounds lead to deformities in the male reproductive organs causing sterility in hybrids (Fishman & Willis, 2006). Even the mitonuclear discordance in crosses between genetically distinct populations can have fitness consequences. When evaluating interpopulation crosses of the intertidal

harpacticoid copepod *T. californicus*, F2 hybrids have decreased fitness in terms of survivorship, development, and fecundity when compared to the parental populations (Burton et al. 1999).

Although there can be negative consequences of cytonuclear incompatibility, mitochondrial replacement, where one species contains the mtDNA of another species, is not unheard of in animals. For example, in a phylogenetic analysis of darters, Near et al. (2011) found that over 12% of all darter species examined possessed a form of heterospecific mtDNA. This study and others recognized that mitochondrial introgression in fish tended to be most common in hybrid zones. Replicate hybrid zones across the ranges of broadly distributed species, such as those found between *F. notatus* and *F. olivaceus*, provide natural laboratories to study the ecological and genetic forces at work during contact between reproductively compatible species (Duvernell et. al., 2007, Schaefer et al., 2011). The *Fundulus notatus* species complex, which is composed of *F. notatus*, *F. olivaceus*, and *F. euryzonus*, has proven to be an emerging model system in evolutionary biology since species within the complex have similar ecological niches (Thomerson & Woolridge, 1970; Blanchard, 1996), but the dynamics of their interactions can vary depending on physical differences in the river systems they inhabit (Schaefer et. al., 2011).

In contrast with the two other species, *Fundulus euryzonus* is restricted to only two rivers (the Amite and Tangipahoa Rivers) in the Lake Pontchartrain drainage (Suttkus & Cashner, 1981). *Fundulus euryzonus* typically overlaps with *F. olivaceus* throughout these two rivers. Hybridization has been reported between the two species in the West Fork of the Amite River at a rate of 2.9% (Schaefer et. al., 2009; Schaefer et.

al., 2011). In comparison to the Amite River, there was far less knowledge about the hybridization patterns in the Tangipahoa River until Flanagan (2016) found that 9.3% of his samples were comprised of individuals with mixed ancestry. However, all individuals that identified as *F. euryzonus* possessed the *F. olivaceus* mitochondrial haplotype, which indicated widespread unidirectional mitochondrial introgression in the Tangipahoa River (Flanagan, 2016).

In the past, the majority of studies in ecological and conservation genetics used a small number of molecular markers, such as allozymes, microsatellites, and amplified fragment length polymorphisms (AFLPs), but they only covered a small subset of the genome (Allendorf, 2017). These markers could be used to address a variety of population genetic questions (e.g., population structure, demographic history, and patterns of gene flow), but they had little power to identify loci that may be subject to natural selection. Fortunately, with the help of new genotyping-by-sequencing (GBS) approaches, large numbers of loci can be identified, even for species lacking in extensive information about their genomes (Narum et al., 2013). The observed single nucleotide polymorphisms (SNPs) can be used to find significant groups of loci on the basis of some metric, such as fixation index (F_{ST}), through various approaches such as a sliding window analysis (Hohenlohe & Magalhaes, 2019). Once clusters of SNP loci are detected, genomic resources can be used to identify genes in close proximity to the SNP loci that might be the target of natural selection.

The goal of this thesis was to build upon Flanagan's (2016) research, which documented the widespread replacement of the *F. euryzonus* mitogenome by *F. olivaceus* mitochondrial DNA in the Tangipahoa River drainage. Given the potential for

mitonuclear functional compensation, I expected to find evidence of purifying and/or positive selection on genes related to mitochondrial function in *F. euryzonus* from the Tangipahoa River. The GBS data collected by Flanagan (2016) for *F. euryzonus* in the Amite and Tangipahoa Rivers was analyzed via a sliding window analysis of F_{ST} values to look for regions of the genome that may be the target of natural selection through interaction between the mitochondrial and nuclear genomes. This work has only recently been made possible through the growing availability of genomic resources for species in the genus *Fundulus* including draft assemblies of the *F. heteroclitus* (Reid et al., 2016; Reid et al., 2017) and *F. olivaceus* (Johnson et al., 2020) genomes.

CHAPTER II: METHODS

Fish Collections

Specimens of *F. euryzonus* in this project were collected as part of S. Flanagan's dissertation (2016). My work focused on specimens collected at two locations, one upstream and one downstream, in both the Tangipahoa and Amite Rivers (n=10 individuals per site). In the Amite River, the upstream site was at the bridge on Graves Road (31.221°, -90.854°) in Amite County, Mississippi. The downstream site was at the LA Highway 10 crossing (30.888°, -90.848°) in St. Helena Parish, Louisiana. In the Tangipahoa River, the upstream site was at Martin Road (31.226°, -90.529°) in Pike County, Mississippi. The southern location was at Chappepeela Creek (30.557°, -90.348°), which is a tributary of the Tangipahoa River in Tangipahoa Parish, Louisiana. Fish were collected with a dip net. A piece of the caudal fin was taken from the specimens and was placed in a salt saturated (SED) preparation buffer (Seutin et. al., 1991). The specimens were stored on ice until they were returned to the laboratory and stored at -20°C.

Molecular Methods & Analyses

The following is a summary of how the data were obtained by Flanagan in 2016. Total genomic DNA was extracted from the fish fin clips using DNeasy Tissue Kit (Qiagen, Valencia, CA) as per the manufacturer's protocol. The Cornell University Genomic Diversity Facility prepared a Genotype-by-Sequencing (GBS) library following protocols modified from Elshire et. al. (2011) with the restriction enzyme *Eco*T22I. This library was then sequenced on an Illumina HiSeq® 2000/2500 instrument. The GBS reads

were processed with TASSEL 5.0 (Glaubitz et al., 2014) using the *Fundulus olivaceus* genome as a reference (Johnson et al., 2020), with the subsequent alignment performed with the default options of bowtie 2.0 (Langmead & Salzberg, 2012). The data were filtered to retain biallelic loci and remove loci with more than 10% missing data or greater than 50% heterozygosity. Filtering conditions were also set to remove any individuals with more than 25% missing data.

The *F. olivaceus* genome is not currently annotated, but is organized into 1,200 contigs. Thus, in order to identify genes associated with SNP loci from *F. euryzonus*, these loci had to be mapped to the annotated genome of *F. heteroclitus* (Waits et al., 2016). Mummer (Marçais et al., 2018) was used to align the *F. olivaceus* contigs from the SNP calling to the *F. heteroclitus* reference genome. I then performed a sliding window analysis with VCFtools (ver. 1.16) using windows 500,000 base pairs wide taken in 100,000 base pair increments. I was interested in genes that were going to reflect the discordance between mitochondrial and nuclear genomes, which could be reflected as either diversifying or stabilizing selection. To do this, I looked for regions of the genome with very high (>0.95) or low (<0.01) weighted fixation index (F_{ST}) values. These values can range from 0 to 1, where a high F_{ST} value represents a large degree of genetic differentiation and a low F_{ST} value implies more genetic similarity among populations. The very high (>0.95) or very low (<0.01) thresholds were chosen to represent the two most extreme levels of genetic differentiation. Regions selected for further analysis included three or more windows in series with F_{ST} values that met these conditions. Descriptive statistics for the window groups within each F_{ST} category were obtained with the *describe.by* function in the R package 'psych.' I compared the sliding window groups

(hereafter called groups) between the two F_{ST} categories in the following metrics: 1) number of SNP loci per group; 2) number of genes per group; and 3) length (in bp) of coverage in the *F. heteroclitus* genome by the group. A Shapiro-Wilk test indicated that the assumption of normality was violated so I statistically compared these measures using a nonparametric Kruskall-Wallace test in R (version 3.6.3; R Core Team, 2020).

Sliding window groups were assigned to scaffolds within the annotated genome of *F. heteroclitus* (http://arthropods.eugenes.org/EvidentialGene/vertebrates/atlantic_killifish/). I determined the putative functional roles of the genes within each scaffold by referencing the UniProt knowledgebase (https://www.uniprot.org/) within the annotation. These genes were assigned to one of the following functional groups: cell cycle, cell development, cell signaling, immune system, intra/intercellular transport, metabolism, miscellaneous, mitochondrial, multiple, transcription, translation and unknown. Some of these groups were based on the categories from Ray (2020), but I added additional ones to better represent the functional classes of genes identified in my study. I used a row-by-columns test of independence to determine if there was an association between the F_{ST} category and the number of genes within the functional groups. Lastly, I explored the influence of length of base pair coverage from the combined area of a contig identified by the sliding window analysis within each F_{ST} category on the number of genes identified.

CHAPTER III: RESULTS

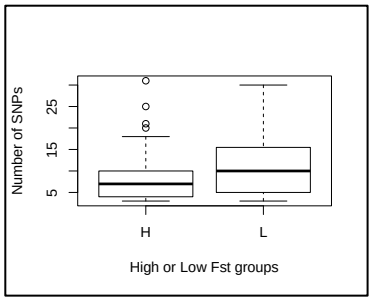
The *F. euryzonus* SNP data from Flanagan (2016) was comprised of a total of 26,744 loci for twenty individuals each from the Tangipahoa and Amite Rivers with 19,453 loci remaining after filtering. The sliding window analysis identified a total of 11,030 windows with an overall mean weighted F_{ST} value of 0.57. The number of windows with F_{ST} values >0.95 were 313, while 254 windows had F_{ST} values < 0.01. From these data, I selected sets of three or more adjacent windows (hereafter referred to as a group) with F_{ST} values > 0.95 (hereafter referred to as High) or < 0.01 (hereafter referred to as Low). A total of 69 sliding window groups were identified, with 45 representing the High category and 24 the Low category. The average weighted F_{ST} value among groups within the High category was 0.99 (SE=0.01) while the average weighted F_{ST} among groups within the Low category was -0.01 (SE = 0.01). The average number of SNPs per group, genes per group and length of coverage by group are reported in Table 1 and graphically represented in Figure 1. None of the differences between High and Low for number of SNPs ($\chi^2(2) = 1.97, p = 0.16$), number of genes ($\chi^2(2) = 0.83, p = 0.36$) or length of coverage ($\chi^2(2) = 1.62, p = 0.20$) were significant.

Table 1. The average number of SNPs and the genes and length of coverage (in kilobases) for groups within the High and Low FST categories. The associated standard errors are reported in parentheses.

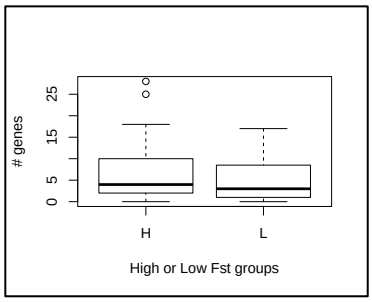
	High	Low
# SNPs	8.89 (0.94)	11.0 (1.44)
# Genes	6.29 (0.94)	5.00 (1.07)
Length of coverage (kb)	251.3 (36.6)	143.0 (32.0)

Figure 1. Box plots for number of SNPs, number of genes, and length of coverage for sliding window groups within the High and Low FST categories represented by panels A-C, respectively.

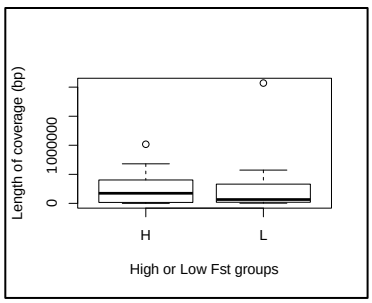
A.



B.



C.



I identified 403 genes within the *F. heteroclitus* genome for the 69 groups (Table 2). This total did not include any transposable elements that may be present within a given scaffold. A linear relationship was seen between the length of coverage in the *F. heteroclitus* genome and the number of genes identified for both F_{ST} categories (Figure 2). The largest three functional groups were cell development (12.9%), metabolism (12.2%) and unknown (12.2%), with mitochondrial genes only representing 2.5%. Within the High F_{ST} category, the top three functional groups were intra/intercellular transport (13.1%), cell development (12.8%), and cell signaling (11.7%), while within the Low F_{ST} category the top three were unknown (19.8%), metabolism (14.9%), and cell development (13.2%). Mitochondrial genes made up the lowest frequency in both F_{ST} categories (2.8% and 1.7%, respectively). Considering only 80 nuclear genes are associated with the Ox-Phos pathway, this proportionally represents a much higher degree of representation, and these genes were of greatest interest in the context of my study (Table 3). The test of independence between F_{ST} categories was significant when all gene function groups were included ($\chi^2(11) = 20.5, p = 0.039$). However, when the unknown genes were removed, the test was no longer significant ($\chi^2(10) = 11.6, p = 0.31$).

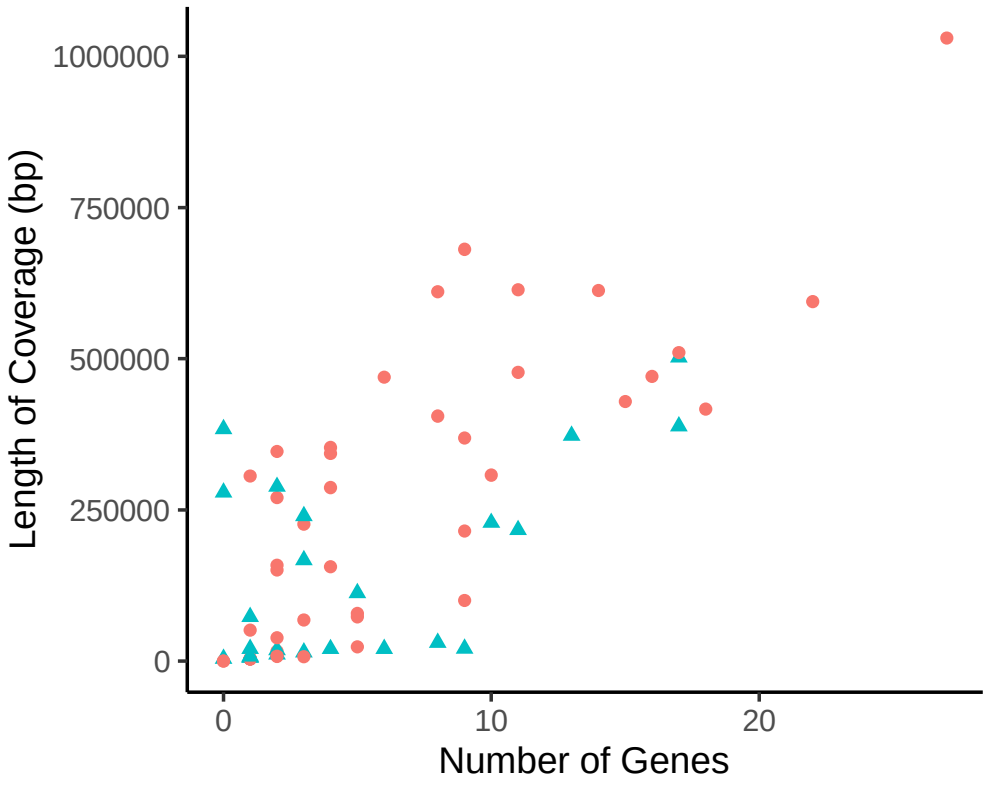
Table 2. Number of genes with the twelve functional groups identified within the High and Low F_{ST} categories.

Gene Function Group	# Genes Total	High F_{ST}	Low F_{ST}
Cell Cycle	27	16	11
Cell Development	52	36	16
Cell Signaling	43	33	10
Immune System	15	12	3
Intra/Intercellular Transport	42	37	5
Metabolism	49	31	18
Miscellaneous	40	28	12
Mitochondrial	10	8	2
Multiple	17	12	5
Transcription	39	30	9
Translation	20	14	6
Unknown	49	25	24

Table 3. Mitochondrial genes identified in the sliding window analysis. The UniProt gene name and scaffold location within the genome of *Fundulus heteroclitus* are provided for each. The first eight genes were from the High F_{ST} category, and the last two were in the Low F_{ST} category.

Gene	UniProt	*F. heteroclitus* Scaffold Location
Funhe2EKm004550t1	CHCHD1	Scaffold10014:180647-180799
Funhe2EKm003242t1	RETSAT	Scaffold161:58176-60793
Funhe2EKm035653t1	NMES1	Scaffold103:89170-103203
Funhe2EKm035656t1	NNT	Scaffold10051:493758-494536
Funhe2EKm033836t1	NQO1	Scaffold10051:504276-526065
Funhe2EKm033833t1	GABPA	Scaffold10014:127187-133402
Funhe2EKm014973t1	PDZD8	Scaffold1046:82593-86984
Funhe2EKm022436t1	UQCRC1	Scaffold9855:689903-689969
Funhe2EKm001633t1	NDUFB4	Scaffold9912:421826-425518
Funhe2EKm027614t1	SMIM20	Scaffold40:270414-273916

Figure 2. Scatter plot of the length of coverage in the Fundulus heteroclitus genome and the number of genes identified in the High F_{ST} (circles) and Low F_{ST} (triangles) categories.



CHAPTER IV: DISCUSSION

This study was an attempt to explore the genetic differences between the genomes of *F. euryzonus* from the Amite and Tangipahoa Rivers in light of the mitonuclear discordance in the latter population. Specifically, I was interested in determining if there were parts of the genome of *Fundulus euryzonus* in the Tangipahoa River that were under purifying or positive selection as a consequence of mitonuclear interactions. The fact that *F. euryzonus* has persisted in this river despite mitochondrial replacement suggests that some sort of natural selection has been at work on this population. I was able to use Flangan's existing SNP data set (2016) to explore this question with a sliding window analysis of F_{ST} values. While some level of genetic differentiation is expected between populations from the Amite and Tangipahoa Rivers, loci demonstrating very high or low F_{ST} values presumably would be linked to genes that are non-neutral with respect to natural selection (Lewontin and Krakeuer, 1973).

I identified 69 regions with what I quantified as High or Low F_{ST} values. After assigning these groups to scaffolds within the *F. heteroclitus* genome, I was able to locate genes annotated in these regions and assign them to a functional group. The test of independence was significant when comparing all twelve gene classes between the Low and High F_{ST} categories. However, that was not the case when genes of unknown function were removed from the analysis, suggesting there was not a pattern of association between gene class groups and the F_{ST} categories. This is perhaps not surprising. Given the number of genes (354 with a known function) identified in the sliding window analysis, not all of these are necessarily potential targets of natural selection in the context of mitonuclear discordance.

Only ten of the genes identified appeared to have mitochondrial function. This may seem like a relatively small number, but by way of comparison, Baris et al. (2017) found that none of their 349 outlier loci in *F. heteroclitus* were directly associated with the 79 nuclear OxPhos proteins. They suggested that the impact of mitonuclear discordance in their study population could instead reflect the regulatory regions of genes with mitochondrial function or genes with indirect influences on mitochondrial function. With this in mind, the fact that cell signaling and intra/intercellular transport genes were in the three highest classes for the High F_{ST} category suggests that other genes within the 403 total could be significant with regard to mitonuclear discordance in *F. euryzonus*. For the purposes of this study, however, I have elected to focus on genes directly related to mitochondrial function.

Out of the ten genes that were associated with the mitochondria, eight genes had High F_{ST} values (CHCHD1, RETSAT, NMES1, NNT, GABPA, NQO1, PDZD8, and UQCRC1), and two genes had Low F_{ST} values (NDUFB4 and SMIM20). These ten mitochondrial genes represent a variety of functional roles. I found that four of the genes are directly related to the electron transport chain. These genes are NMES1, UQCRC1, NDUFB4, and SMIM20. The NMES1 gene encodes a novel accessory protein of Complex IV of the mitochondrial electron transport chain (Endou et. al., 2020). UQCRC1 is a component of Complex III of the mitochondrial electron transport chain and increases oxidative phosphorylation and ATP production at elevated expression levels (Wang et. al., 2020). NDUFB4 is a component of Complex I of the electron transport chain that directly interacts with complex III (Jang & Javadov, 2018) and encodes a non-catalytic subunit of the NADH dehydrogenase (ubiquinone) (Chang et. al., 2020).

SMIM20 forms part of the mitochondrial translation regulation assembly intermediate of cytochrome c oxidase (MITRAC) and stabilizes a subunit of cytochrome c oxidase (COX), which is an essential part of the electron transport chain (Mcilwraith & Belsham, 2018).

The remaining genes have other roles in mitochondrial function. I found that CHCHD1 is mainly associated with the small subunit of mitochondrial ribosomes (Koc et. al., 2013) and was found to be under positive selection in *Nothobranchius fruzeri*, *N. kadleci* and *N. kuhntae.* (FKK-branch; Sahm et. al., 2017). RETSAT is an oxidoreductase that is important in retinoid metabolism and is a widely shared component of mechanisms that are involved in insulin resistance (Rhee & Plutzky, 2012). NNT is an important enzyme that regulates mitochondrial NADPH levels and mitochondrial redox balance (Rao et. al., 2020). GABPA controls the expression of nuclear genes encoding mitochondrial proteins of the respiratory chain (Kropotov et. al., 2007). The NQO1 gene codes for an enzyme that can enhance mitochondrial activity at elevated levels without causing an increase in production of reactive oxygen species (ROS), and it can also protect cells against mitochondrial toxins (Li et. al., 2014; Kim et. al., 2013). Lastly, PDZD8 connects the endoplasmic reticulum and mitochondrial membranes and mediates key physiological processes, such as Ca^{2+} transfer (Hirabayashi et. al., 2017).

When comparing the functions of the genes, NMES1, UQCRC1, NDUFB4 are all components of the electron transport chain, and SMIM20 is also associated with the electron transport chain due to the stabilization of a key subunit. Given the critical role of the electron transport chain in cellular respiration, it should not be a surprise that these genes might have experienced selection as a consequence of mitonuclear discordance.

Although the other genes appear to be independent with regard to their function in the mitochondria, several of these have been previously identified in the literature as having important fitness consequences, such as a longer lifespan (Sahm et. al., 2017).

This work presents evidence for selection operating on mitonuclear interactions through the identification of outlier loci associated with nuclear genes with mitochondrial function and adds to a growing literature on non-model organisms (i.e., not *Drosophila spp*.). While I focused on nuclear genes associated with mitochondrial function, we should not neglect to consider that mitonuclear discordance could influence selection operating on other classes of genes, such as ones that interact with regulatory regions of mitonuclear genes. In addition to future genomics work on this system, another way to address the issue of mitonuclear discordance in *F. euryzonus* would be to pursue studies that measure fitness (e.g., metabolic rates, swimming performance) of individuals from the two river systems. If the coevolution between the nuclear and mitochondrial genomes in the Tangipahoa *F. euryzonus* is incomplete, then one might expect to see lower mitochondrial function that could manifest itself in some sort of reduction in performance.

REFERENCES

Allendorf, F.W. (2017). Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology*. 26, 420-430.

Bernardo, P.H., Sanchez-Ramirez, S., Sanchez-Pacheco, S.J., Alvarez-Castaneda, S.T., Aguilera-Miller, E.F., Mendez-de la Cruz, F.R., & Murphy, R.W. (2019). Extreme mito-nuclear discordance in a peninsular lizard: the role of drift, selection, and climate. *Heredity*. 123, 359-370.

Blanchard, J. L., & Lynch, M. (2000). Organellar genes: why do they end up in the nucleus?. *Trends in Genetics*, *16*(7), 315-320.

Blanchard, T. A. (1996). Ovarian cycles and microhabitat use in two species of topminnow, *Fundulus olivaceus* and *F. euryzonus*, from the southeastern United States. *Environmental Biology of Fishes*, *47*(2), 155-163.

Burton, R. S., Rawson, P. D., & Edmands, S. (1999). Genetic architecture of physiological phenotypes: empirical evidence for coadapted gene complexes. *American Zoologist*, *39*(2), 451-462.

Cavalier-Smith, T. (1987). The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. *Annals of the New York Academy of Sciences*, *503*(1), 55-71.

Chang, Z., Miao, X., & Zhao, W. (2020). Identification of prognostic dosage-sensitive genes in colorectal cancer based on multi-omics. *Frontiers in Genet*ics. 10:1310. doi: 10.3389/fgene.2019.01310.

Duvernell, D. D., Schaefer, J. F., Hancks, D. C., Fonoti, J. A., & Ravanelli, A. M. (2007). Hybridization and reproductive isolation among syntopic populations of the topminnows *Fundulus notatus* and *F. olivaceus*. *Journal of Evolutionary Biology*, *20*(1), 152-164.

Elshire, R.J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLOS ONE*, *6*(5), e19379.

Endou, M., Yoshida, K., Hirota, M., Nakajima, C., Sakaguchi, A., Komatsubara, N, & Kurihara, Y. (2020). Coxfa4l3, a novel mitochondrial electron transport chain complex 4 subunit protein, switches from Coxfa4 during spermatogenesis. *Mitochondrion*. 52, 1-7.

Etterson, J. R., Keller, S. R., & Galloway, L. F. (2007). Epistatic and cytonuclear interactions govern outbreeding depression in the autotetraploid *Campanulastrum americanum*. *Evolution*, *61*(11), 2671-2683.

Fishman, L., & Willis, J. H. (2006). A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution*, *60*(7), 1372-1381.

Flanagan, S. D. (2016). Evolution in Light of Mitonuclear Landscapes: an Examination of Mitochondrial Replacement in Killifish (Fundulus Spp.) The University of Southern Mississippi. 1-15.

Futuyma, D. J. (1998). Evolutionary Biology. 3rd edition. *Sinauer Associates*, Sunderland, Massachusetts.

Glaubitz, J.C., Casstevens, T.M., Lu, F., Harriman, J., Elshire, R.J., Sun, Q., & Buckler, E.S. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS one*, 9(2), e90346.

Hirabayashi, Y., Kwon, S., Paek, H., Pernice, W., Paul, M.A., Lee, J., Erfani, P., Raczkowski, A., Petrey, D.S., Pon, L.A., & Polleux, F. (2017). ER-mitochondria tethering by PDZD8 regulates Ca2+ dynamics in mammalian neurons. *Science*. 358(6363), 623-630.

Hohenlohe, P.A. & Magalhaes, I.S. (2019). The population genomics of parallel adaptation: lessons from threespine stickleback. *Population Genomics: Marine Organism*, 249-276.

Jang, S. & Javadov, S. (2018). Current challenges in elucidating respiratory supercomplexes in mitochondria: methodological obstacles. *Frontiers in Physiology.* 9:238. doi: 10.3389/fphys.2018.00238.

Johnson, L.K., Sahasrabudhe, R., Gill, J.A., Roach, J. L., Froenicke, L., Brown, C.T., & Whitehead, A. (2020). Draft genome assemblies using sequencing reads from Oxford Nanopore Technology and Illumina platforms for four species of North American Fundulus killifish. *GigaScience*, 9(6), giaa067.

Kim, J., Kim, S., Mattson, M.P., & Hyun, D. (2013). Mitochondrial function in human neuroblastoma cells is upregulated and protected by NQO1, a plasma membrane redox enzyme. *PLOS*. doi:10.1371/journal.pone.0069030.

Koc, E.C., Cimen, H., Kumcuoglu, B., Abu, N., Akpinar, G., Haque, M.E., Spremulli, L.L., & Koc, H. (2013). Identification and characterization of CHCHD1,

AURKAIP1, and CRIF1 as new members of the mammalian mitochondrial ribosome. *Front. Physiol*. 4:183.

Kropotov, A., Usmanova, N., Serikov, V., Zhivotovsky, B., & Tomilin, N. (2007). Mitochondrial targeting of human peroxiredoxin V protein and regulation of PRDX5 gene expression by nuclear transcription factors controlling biogenesis of mitochondria. *The FEBS Journal.* 274(22), 5804-5814.

Langmead, B., & Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357-359.

Lewontin, R., & Krakauer, J. (1973). Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. *Genetics*, 74, 175-195.

Li, J.Z., Ke, Y., Misra, H.P., Trush, M.A., Li, Y.R., Zhu, H., & Jia, Z. (2014). Mechanistic studies of cancer cell mitochondria- and NQO1-mediated redox activation of beta-lapachone, a potentially novel anticancer agent. *Toxicology and Applied Pharmacology*, 281(3), 285-293.

Marcais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., & Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. PLOS Computational Biology, 14(1), e1005944.

Mcilwraith, E.K. & Belsham, D.D. (2018). Phoenixin: uncovering its receptor, signaling and functions. *Acta Pharmacologica Sinica*, 39, 774-778.

Morales H.E., Pavlova, A., Joseph, L., and Sunnucks, P. (2015). Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance.

Molecular Ecology, 24, 2820–2837

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R., & Hohenlohe, P.A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22(11), 2841.

Nass, M. M., & Nass, S. (1963). Intramitochondrial fibers with DNA characteristics I. Fixation and electron staining reactions. *The Journal of Cell Biology*, *19*(3), 593-611.

Near, T. J., Bossu, C. M., Bradburd, G. S., Carlson, R. L., Harrington, R. C., Hollingsworth, P. R., ... & Etnier, D. A. (2011). Phylogeny and temporal diversification of darters (Percidae: Etheostomatinae). *Systematic Biology*, *60*(5), 565-595.

Newton, K. J. (1988). Plant mitochondrial genomes: organization, expression and variation. *Annual Review of Plant Physiology and Plant Molecular Biology*, *39*(1), 503-532.

Oka, A., Mita, A., Sakurai-Yamatani, N., Yamamota, H., Takagi, N., Takano-Shimizu, T., Toshimori, K., Moriwaki, K., & Shiroshi, T. (2004). Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics*, 166(2), 913-924.

Pavlova A., Gan, H.M., Lee, Y.P., Austin, C.M., Gilligan, D.M.,Lintermans, M., & Sunnucks, P. (2017). Purifying selection and genetic drift shaped Pleistocene evolution of the mitochondrial genome in an endangered Australian freshwater fish. *Heredity*, 118, 466–476

R Core Team. (2020). R: a language and environment for statistical computing. Vienna, Austria: *R Foundation for Statistical Computing.* Retrieved from https://www.R-project.org/.

Rank, N.E., Mardulyn, P., Heidl, S.J., Roberts, K.T., Zavala, N.A., Smiley, J.T., & Dahlhoff, E.P. (2020). Mitonuclear mismatch alters performance and reproductive success in naturally introgressed populations of a montane leaf beetle. *Evolution*, *74*(8), 1724-1740.

Rao, K.N. S., Shen, X., Pardue, S., & Krzywanski, D. (2020). Nicotinamide nucleotide transhydrogenase (NNT) regulates mitochondrial ROS and endothelial dysfunction in response to angiotensin II. *Redox Biology*, 36, 101650.

Ray, A.F. (2020). Patterns of Genomic Introgression in Topminnow Hybrid Zones. Honors Thesis. The University of Southern Mississippi, Hattiesburg.

Reid, N. M., Jackson, C. E., Gilbert, D., Minx, P., Montague, M. J., Hampton, T. H., ... & Whitehead, A. (2017). The landscape of extreme genomic variation in the highly adaptable Atlantic killifish. *Genome Biology and Evolution*, *9*(3), 659-676.

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., ... & Whitehead, A. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, *354*(6317), 1305-1308.

Rhee, E.J., & Plutzky, J. (2012). Retinoid metabolism and diabetes mellitus. *Korean Diabetes Association*, 36, 167-180.

Sagan, L. (1967). On the origin of mitosing cells. *Journal of Theoretical Biology*, *14*(3), 225-274.

Sahm, A., Bens, M., Platzer, M., & Cellerno, A. (2017). Parallel evolution of genes controlling mitonuclear balance in short-lived annual fishes. *Aging Cell*, 16(3), 488-496.

Schaefer, J. F., Duvernell, D. D., & Kreiser, B. R. (2011). Ecological and genetic assessment of spatial structure among replicate contact zones between two topminnow species. *Evolutionary Ecology*, *25*(5), 1145-1161.

Schaefer, J., Kreiser, B. R., Champagne, C., Mickle, P. M., & Duvernell, D. D. (2009). Patterns of co-existence and hybridisation between narrowly endemic (*Fundulus euryzonus*) and broadly distributed (*F. olivaceus*) topminnows in a riverine contact zone. *Ecology of Freshwater Fish*, *18*(3), 360-368.

Seutin, G., White, B.N., & Boag, P.T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, 69(1), 82-90.

Suttkus, R. D., & Cashner, R. C. (1981). A new species of cyprinodontid fish, genus *Fundulus* (Zygonectes), from Lake Pontchartrain tributaries in Louisiana and Mississippi. *Bulletin of the Alabama Museum of Natural History*, (6).

Thomerson, J. E., & Wooldridge, D. P. (1970). Food habits of allotopic and syntopic populations of the topminnows *Fundulus olivaceus* and *Fundulus notatus*. *American Midland Naturalist*, 1970(1), 573-576.

Waits, E., Martinson, J., Rinner, B., Morris, S., Proestou, D., Champlin, D., & Nacci, D. (2016). Genetic Linkage Map and Comparative Genome Analysis for the Atlantic Killifish (*Fundulus heteroclitus*). *Open Journal of Genetics*, 6(1), 28-38.

Wallin, I. E. (1923). The mitochondria problem. *The American Naturalist*, *57*(650), 255-261.

Wang, Q., Li, M., Gan, Y., Jiang, S., Qiao, J., Zhang, W., Fan, Y., Shen, Y., Song, Y., Meng, Z., Yao, M., Gu, J., Zhang, Z., & Tu, H. (2020). Mitochondrial protein UQCRC1 is oncogenic and a potential therapeutic target for pancreatic cancer. *Theranostics*, 10(5), 2141-2157.

Wiesner, R. J., Rüegg, J. C., & Morano, I. (1992). Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissues. *Biochemical and Biophysical Research Communications*, *183*(2), 553-559.