January 2020

Evolutionary Ecology Of The Native Johnny Darter (etheostoma Nigrum) And The Invasive Round Goby (neogobius Melanostomus): A Genomic Perspective

Abby Joyce Wicks
Wayne State University

EVOLUTIONARY ECOLOGY OF THE NATIVE JOHNNY DARTER (*ETHEOSTOMA NIGRUM*) AND THE INVASIVE ROUND GOBY (*NEOGOBIUS MELANOSTOMUS*): A GENOMIC PERSPECTIVE

by

ABBY J. WICKS

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2020

MAJOR: BIOLOGICAL SCIENCES

Approved by:

_____

Advisor Date

_____

_____

_____

AKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Thomas Dowling for his expertise, support, and enthusiasm that made this dissertation possible. It is truly a privilege to be a part of his lab. I would also like to thank my committee members for their valuable input. Dr. Donna Kashian, Dr. Robert Podolsky, and Dr. Weilong Hao were all extremely helpful with their advice and direction which greatly improved this dissertation.

I am forever grateful to Corey Krabbenhoft. The data she compiled provided context that made this work so much more valuable. I am so fortunate to have teamed up with someone of her experience and expertise; almost everything I now know about fieldwork I owe to Corey. In addition to her knowledge contribution, Corey's company made long drives and long days of field work a lot more fun.

Thank you to current and former Dowling Lab members, Melody Saltzgiver, Trevor Krabbenhoft, Kelly Pollard, and Joel Corush, for their comradery, input, and assistance in the field. Special thanks is owed to Trevor Krabbenhoft. He provided the training I needed in benchwork and bioinformatics to be able to complete this project. He also stepped in and helped with work I was unable to do during my pregnancy and leave. I am forever indebted to him for the expertise and support that make this project possible. A special thank you as well to Kelly Pollard who never said no when I needed her in the field.

Many undergraduate field assistants contributed to this research including Miles Bowman, Brittani Dodd, Kristin Kargol, Robert Roose, and Julie Nitz. I am particularly grateful for Miles Bowman's contributions. His steadfast presence in the field made it possible for me to be home with a newborn baby while the field work continued. Miles was also an integral part of the work for Chapter 1.

Thank you to Friends of the Rouge, in particular Sally Patrella, Robert Mueller, and Phil Kukulski for their local knowledge that informed this project and for their willingness to collaborate.

Lastly, thank you to my family. To my husband, Joseph for his unwavering support of my pursuit down this path. To my children, Penelope and Theodore, for making life so much more fun and for forcing me to have balance. The support of my parents and my husband's parents made it possible for me to complete this project with a young family. I also exploited many family members for fieldwork and their enthusiasm to participate was heartwarming.

This work was made possible with funding from Wayne State University. Fieldwork and collections were undertaken with permission from the Michigan Department of Natural Resources, under scientific collector permits issued to Thomas Dowling and Corey Krabbenhoft. Collection and handling of specimens were in accordance with Wayne State University Institutional Animal Care and Use Committee (IACUC) protocol 18-02-0553.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Invasive species are a major threat to biodiversity, economies, and water quality (Pejchar and Mooney 2009, Rosaen et al. 2012). Aquatic invasive species can cause dramatic changes in aquatic ecosystems including impacts on community composition, food web dynamics, hydrology, and environmental quality (Strayer 2010). For example, Clarkston et al. (2005) noted that predation by non-native fishes in the Colorado river system is the primary barrier to recovery of several endangered species in the system. Along the Atlantic Coast, invasion by Lionfish (*Pterois volitans*) has caused rapid declines in populations of coral reef fish (Green et al. 2012).

As a major thoroughfare for shipping, the Great Lakes region is particularly vulnerable to invasive species introductions. NOAA (2012) notes 180 aquatic invasive species in the Great Lakes, and these have had significant economic and ecological impacts (Pimental 2005, Lovell et al. 2006, Pejchar and Mooney 2009). Establishment of zebra mussels (*Dreissena polymorpha*) in the Great Lakes and beyond provides an excellent example of negative impacts of invasive species. Their spread has led to decreases in phytoplankton and zooplankton, creating a cascade of effects that has altered water chemistry and biota (Strayer 2009). Successful management against invasive species requires understanding how these invaders interact with biotic and abiotic factors in their new system and how these factors impact their success.

The subject species of this dissertation, round goby (*Neogobius melanostomus*, family Gobiidae, Order Perciformes), is an invader across multiple continents impacting ecosystems in Europe and the Great Lakes region of North America (Corkum et al. 2004). Likely introduced via ballast water, round goby was first discovered in the St. Clair River in 1990 and has spread rapidly throughout the region (Jude et al. 1991, Kornis et al. 2012); it is now present in all five Great Lakes and is becoming increasingly common in Great Lakes tributaries (Kornis and Vander Zanden 2010, Poos et al 2010).

Round goby is a generalist species and possesses the ability to use a wide range of resources, a potential advantage for invasive taxa (Nurkse et al. 2016). This species is benthic, preferring rocky substrate, but it uses soft substrate as well (Kornis et al. 2012). It feeds at all times of day, consuming mollusks, aquatic insects, fish eggs, and small fish. While other prey are preferred (Coulter et al. 2011), the ability to use mollusks as a food source may provide round goby an advantage over natives due to the recent invasion of the Great Lakes system by zebra and quagga mussels (Ghedotti et al. 1995). Round goby can spawn multiple times a year from April through September, preferring shallow waters with rocky substrate (Kornis et al 2012). Male round gobies establish and aggressively defend territories that includes their nest (Dubs and Corkum 1996, Balshine et al. 2005).

Round goby has become an important member of the lotic ichthyofauna, where it competes with sculpins (*Cottus* sp.), darters (*Etheostoma* and *Percina* sp.), and madtoms (*Noturus* sp., Lauer et al. 2004, Burkett and Jude 2015). Direct interactions between round goby and some of these ecologically similar native species have been studied in the lab, documenting increased aggression with native species as well as egg predation and nest eviction (Dubs and Corkum 1996, Balshine et al. 2005). As an egg predator, round gobies have also been linked to decreases in population sizes of native fishes, including economically important species as well as those of conservation interest (e.g., lake trout, *Salvelinus namaycush*; lake sturgeon, *Acipenser fulvescens*; walleye, *Sander vitreus*; and smallmouth bass, *Micropterus dolomieu*; Kornis et al. 2012); however, round gobies have established at some localities with no apparent impact on the abundance of native fishes (Riley et al. 2008, Kornis et al. 2013). Given variable impacts on native species in invaded systems, round goby invasion presents an opportunity to investigate the relative response to variation in the environment on round goby and native species.

Johnny darter (*Etheostoma nigrum,* Family Percidae, Order Perciformes) is a species common to many river systems of eastern North America. Like round goby, Johnny darter is a benthic species that prefers shallow, slow moving water and sand or gravel substrate where it consumes copepods and aquatic insects (Scott and Crossman 1973). Spawning occurs once a year, typically in May and June. Eggs are laid on the underside of rocks or debris with males guarding these nests (Scott and Crossman 1973). Given overlap in habitat and diet, it is likely that competitive interactions occur between Johnny darter and round goby.

Indirect and direct evidence of negative interactions between these species has been reported. Johnny darter has been shown to decline in abundance as round goby became established (Lauer et al 2004, Burkette and Jude 2015), and in experimental enclosures, Johnny darters have exhibited decreased growth in coexistence with round goby (Kornis et al. 2014). This negative interaction is not absolute, however, as Johnny darter and round goby have been found to coexist without evidence of negative impacts (Riley et al. 2008, Kornis et al. 2013). Johnny darter exhibits niche overlap with round goby and is a relatively common species in Great Lakes streams; therefore, it will make an ideal species for comparison with round goby.

Although many studies have investigated invasive species at the ecosystem and organism level (reviewed in Crooks 2002, Kornis et al. 2014, Crane and Einhouse 2016), few studies have compared invasive species and native species using genomic tools. Until recently, gene expression studies were limited to microarray technology, and these were restricted to studies of organisms with available reference genomes (Schena et al. 1995). Next generation sequencing technology has now made it possible to study ecological interactions through characterization of levels and patterns of gene expression (Ekblom and Galindo 2011), providing perspective of how physiological processes can change in varying environments of natural systems. Correlation of

patterns of gene expression with environmental factors in wild populations can validate current knowledge of gene function as well as identify genes of interest for future study.

Here, I use traditional population genetic methods and genomic tools to examine the evolutionary ecology of invasive round goby and native Johnny darter in two river systems, the Clinton and Rouge rivers, from southeastern Michigan. In Chapter 1, I quantify patterns of gene exchange among Michigan populations for each species to test the hypothesis that there are no significant underlying genetic differences among samples of round goby and Johnny darter from these two rivers; therefore, differences in patterns of gene expression among locations and river drainages would result from recent ecological factors, not historical genetic differences. For Chapter 2, I quantify variation in levels of gene expression within and among samples from these rivers for both species. This approach identifies genes that exhibit significantly increased or decreased levels of expression between sites and rivers. In Chapter 3, I identify environmental factors that are correlated with gene networks. Gene expression data generated in Chapter 2 is examined in the context of environmental data collected by C. Krabbenhoft for her dissertation, identifying differences in expression for groups of genes that are associated with specific biotic and/or abiotic environmental variables. This includes the potential impacts of interspecific interactions. Addressing these objectives will allow for better understanding of how native Johnny darter and invasive round goby interact with each other and with their environment.

CHAPTER 1: POPULATION GENETICS OF ROUND GOBY AND JOHNNY DARTER IN LOWER MICHIGAN

INTRODUCTION

Evolutionary history, geologic history, and ecology of a species contribute to variable patterns of genetic diversity across the landscape (e.g., Avise 1992, Heithaus and Laushman 1997, Leclerc et al. 2008). Quantifying these patterns provides information on the structure and dynamics of populations in the context of landscape features. Depending on an organism's dispersal ability, landscape features may function to facilitate gene flow or create barriers that isolate populations. Isolated populations become genetically distinct overtime, proceeding on separate evolutionary tracks due to the processes of mutation, drift, and differential selection. However, if connections exist that facilitate dispersal, populations can experience continuous gene flow, eliminating differences (Slatkin 1987). Species vary in their ability to disperse and the extent to which landscape features facilitate or prevent gene flow depends on the ecology of a species (Hitt and Angermeier 2008). Taxa with more limited habitat preferences may be restricted to suitable areas (Tibbets and Dowling 1996) while a more plastic species may be able to disperse through variable habitats and patterns of genetic diversity are likely to reflect these ecological differences (Bohonak 1999).

For native fishes of the Great Lakes region, the glacial history of the area has influenced population structure and the distribution of genetic diversity. The Laurentian Great Lakes were formed by glacial action during the Pleistocene, approximately 14,000 years ago, created by retreat of the Laurentide ice sheet (reviewed in Bailey and Smith 1981). The Lake Michigan and Lake Erie/Huron basins were formed as a result of separate glacial lobes, each with their own periglacial lakes (Lakes Chicago and Maumee, respectively) and their own outflow systems. Thus, the Eastern

and Western basins of the Great Lakes represent separate glacial refugia for species and were potentially colonized by fishes from distinct source populations (Bailey and Smith 1981, Lewis et al. 2008). Population structure of fishes in rivers of the two basins often reflects this geologic history as some species from these basins are distinct relative to each other (e.g., Gach 1996, Dowling et al. 1997, Stepien et al. 2009). Ecology of a species influences these patterns as well. Coldwater fishes may maintain populations near the glacial front where more recent connections could result in a more homogenous population colonizing the east and west basins of the post-glacial Great Lakes (Bailey and Smith 1981).

Here, the population genetics of two common Great Lakes fishes, round goby and Johnny darter will be examined to compare their levels of genetic diversity and population structure within each species to better understand factors that contribute to patterns of genetic variation in the two species. Phylogeographic methods will be used to characterize levels and patterns of genetic diversity within and among populations of each of these species using a presumably neutral mitochondrial DNA gene.

As a native species to the Great Lakes, Johnny darter populations may exhibit patterns that reflect their colonization history during glacial retreat. As such, populations from the western basin of the Great Lakes may be significantly different from those of the eastern basin, providing an illustration of how historical processes can influence current patterns of genetic variation. Natural and anthropogenic barriers are likely to influence structure in Johnny darter populations as well. Johnny darters prefer shallow, slow moving water; therefore, large waterways may function as barriers to gene flow and could result in differentiation across study rivers (Leidy 1992). A study of Johnny darter populations in Wisconsin using microsatellites identified significant genetic differences among local drainages (Westbrook 2012).

Unlike Johnny darters, round goby is a recent invader of the Great Lakes, first discovered in the St. Clair River in 1990 (Jude et al. 1992); therefore, this species is not influenced by local geological history. Population genetic studies using microsatellite and mitochondrial DNA markers determined that the Great Lakes were populated by individuals derived from a single source population from a tributary to the Black Sea (Brown and Stepien 2009). Round goby is now present in all five Great Lakes and in many tributaries (Kornis et al. 2012). Analyses of population structure indicate that genetic diversity of the source population is largely maintained throughout the Great Lakes; some limited structure has been identified and is primarily driven by differences in the populations of Saginaw Bay and Lake Ontario (Brown and Stepien 2009). In recently colonized Great Lakes tributaries, round goby populations have generally come to represent their local source populations likely due to consistent propagule pressure, through migration and aided by human activity (Bronnehuber et al. 2011, Saard et al. 2019).

As a recent invader, round goby are expected to exhibit limited differences (e.g., numbers of haplotypes and mutations among them) within and among sampled populations. Additionally, it is expected that the most recently established riverine populations may exhibit only a subset of the variation found in the source lake populations identified (Brown et al. 2008). Johnny darters have resided in the region since glaciation (Bailey and Smith 1981), and given the complex glacial history of the region, there may be differences among regions (e.g., eastern vs western Great Lakes). As a native species, Johnny darter populations are more likely to have diverged since colonization as compared to round goby, thus I expect more genetic differentiation among populations of Johnny darter than round goby. Large waterways also likely represent a barrier to gene flow for Johnny darter and could lead to divergence among populations. These hypotheses

were tested by sequencing the mitochondrial DNA gene NADH dehydrogenase subunit 2 (ND2) for populations of Johnny darter and round goby throughout lower Michigan.

METHODS

Sixteen localities were selected representing the Lake Michigan, Lake Huron, and Lake Erie watersheds (Figure 1). Fishes were collected by seining, with up to 20 individuals of each species collected at each site. In some localities where one or both species was rare, multiple nearby localities and/or sample dates were combined to achieve desired sample size. Individuals were fin clipped, and tissue was stored in 95% ethanol. For DNA isolation, tissue samples were dissolved in a solution of proteinase K and sodium dodecyl sulfate, and purified via one of two methods: phenol/chloroform extraction method or magnetic bead DNA purification (samples collected in 2015-2016 and 2017-2019, respectively). Phenol/chloroform extraction followed Tibbets and Dowling (1996). Magnetic bead purifications followed the manufacturer's (Axygen) protocol with the following changes: 25 microliters of lysate was added to 25 microliters of AxyPrep, 70% ethanol was used for two washes, and DNA was eluted in 40 ul of water. DNA was assessed for quality and quantity with a Nanodrop Spectrophotometer (ThermoFisher).

The mitochondrial gene, NADH dehydrogenase subunit 2 (ND2) was selected because of its relatively high rate of mutation and strict maternal inheritance, allowing for better characterization of more recent events. Sequences were amplified using the primers B2Gila (5' CTCTTAGTGCTTCCTCACA 3') and ASN (5' CGCGTTTAGCTGTTAACTAA 3') for Johnny darter and RGND2F (5' AGCATGCCGGTTAAAATCC 3') and RGND2R (5' GGATCCGAGGCCTTCCTGTCT 3') for round goby. The following PCR conditions were used for both species: 94°C for 15 mins, 25 cycles of denaturation at 94°C for 1 min, annealing at 58°C

for 1 min, and polymerization at 72°C for 2 min with a final annealing step of 72°C for 10 min, and holding at 4°C until long term storage at -20°C. Amplification products were checked for quantity and quality using agarose gel electrophoresis prior to being sent to the Wayne State University's Applied Genomics Technology Center for 2015-2018 samples and Eton Bioscience for 2019 samples. Products were sequenced with the same primers using Applied Biosystems DNA Analyzer 3730 sequencer, producing two approximately 700 bp overlapping fragments.

Some individuals included in this study were also included in the RNA-seq studies in the subsequent chapters. For these individuals, the ND2 gene was obtained from the RNA-seq data (see Chapter 2 for methods). From the assembly and aligned read files, the SAMtools package (Li 2011) was used to call SNPs and output a consensus sequence for each individual. Variants were quality filtered for phred score >30 and individuals with ambiguous variant calls were excluded from the analysis.

Sequences were aligned, assembled, and trimmed in Bioedit (Version 7.0.5.3) (Hall et al. 2011). MEGA (Version 5.2.2) (Kumar et al. 2016) was used to produce a maximum likelihood. tree and assign haplotypes. Haplotype counts for each location were used in ARLEQUIN (version 3.5.1.2) (Excoffier and Lischer 2010) to obtain estimates of genetic diversity (e.g., number of haplotypes, gene diversity) within populations. Variation among populations was assessed and tested for significance using a molecular analysis of variance (AMOVA), also in ARLEQUIN. This approach was used to partition levels of genetic variation within and among river drainages and regions (e.g., eastern vs western basins of the Great Lakes). Haplotype networks were created as median-joining networks with PopArt (Version 1.7) (Leigh and Bryant 2015). Neighbor joining trees of populations were generated with pairwise F_{ST} values in MEGA. For Johnny darter, additional ND2 sequences were obtained from NCBI GenBank and included in the maximum

likelihood tree of haplotypes, also generated with MEGA with 1000 bootstrap replicates. These included closely related species *E. olmstedi* (EF027210), *E. podostemone* (JQ088571), *E. perlongum* (JQ088568), *E. susanae* (JQ088589), *E. vitreum* (FJ381264), and *E. longimanum* (JQ088552) and co-occurring darter species *E. blennioides* (JQ088546), *E. exile* (EF027194), *and E. caereleum* (JQ088546). Additionally, one Johnny darter ND2 sequence (JQ088561) from an individual collected from the Embarras River, a tributary of the Wabash River (N. Lang, pers. comm.) was obtained from GenBank and included in the analysis.

RESULTS

Round Goby

A total of 247 Round Goby samples were collected from 12 localities and sequenced for ND2. Five haplotypes were identified (Table 1, Figure 2). Each was separated by one substitution from its closest neighbor with the exception of haplotype C which differed by two changes. Four variants resulted from changes to third codon positions (Haplotypes A-D) and did not result in amino acid changes. One variant resulted from a mutation in the first codon position and resulted in an amino acid change (Haplotype E).

Haplotype A was found in 96% of all individuals sampled, resulting in very low levels of gene and nucleotide diversity within samples (Table 2, Figure 3). AMOVA was used to quantify levels of genetic variance within and among samples (F_{ST}). The distribution of variation was further partitioned to test for differences between eastern (Lakes Huron, St. Clair, and Erie) and western lower Great Lakes drainages (Lake Michigan) (F_{CT}) and among samples within those two groups (F_{SC}). Statistical assessment failed to identify significant differences among samples (F_{ST}

= 0.033, P = 0.11), among the two sets of drainages (F_{CT} = 0.017, P = 0.05), or among samples within these two sets of drainages (F_{SC} = 0.016, P = 0.25). This lack of differentiation was also supported by examination of pairwise estimates of F_{ST} as only three of the comparisons were significant (Table 3). Similarity among samples was examined by clustering samples by F_{ST} using the neighbor-joining method (Figure 4). There was no overarching genotypic structure of round goby populations as samples from different lake basins were intermingled.

Johnny Darter

A total of 207 Johnny darter samples were collected from 13 localities and sequenced for ND2. Twenty-eight haplotypes were identified (Table 3, Figure 5). There were 53 polymorphic sites with 13 of those variants in the first codon position and 40 in the third codon position. Of these, 10 resulted in amino acid changes. Most of variants resulted from single base pair changes from the most common haplotype, A. Two haplotypes, AB and AC, differed from haplotype A by 22 single base pair changes.

A maximum likelihood tree was generated to examine variation in Johnny darter from the lower peninsula of Michigan in phylogenetic context relative to other samples obtained from GenBank (Figure 6). Most haplotypes sampled here cluster closely to the most common haplotype, A, and formed a monophyletic lineage with limited phylogenetic structure within. Haplotypes U, V, W, and X form a lineage in the minimum spanning network (Figure 5), however, this group is not significantly supported by the ML analysis, consistent with the relatively small number of changes relative to A. A distinct haplotype lineage Y-AA occurred only in the Oqueoc and Little Manistee rivers (Lakes Huron and Michigan drainages, respectively); however, this lineage was very similar to the large group of Johnny darter haplotypes. *Etheostoma podostemone* was the sister taxon to this large group of haplotypes, and the divergent haplotypes (AB and AC) from

Stoney Creek (Lake Erie drainage) were sister to all haplotypes from Michigan. Lineage AB-AC was most similar to Johnny darter from the Embarras River (Wabash River drainage) in central Illinois.

Geographic patterns of variation were used to assess population structure within and among samples of Johnny darter (Figure 7). Haplotype A was the most common, representing over 40% of all individuals sampled, occurring at every locality except the Kalamazoo River. All other haplotypes were more localized, occurring at a maximum of two localities. Private alleles were also common, with 22 alleles occurring at only single localities; however, every locality harbored more than one haplotype.

Levels of sequence diversity were highly variable among samples (Table 4), with gene diversity ranging from 0.13 – 0.76. Exceptional levels of gene and nucleotide diversity were noted in Stoney Creek and Little Manistee samples due to the presence of divergent haplotypes discussed above. Stoney Creek samples exhibited five haplotypes with a mean number of pairwise differences of 9.67 and the Little Manistee River contained five haplotypes with a mean number of pairwise differences of 5.05.

AMOVA was used to partition genetic variance into within and among sample components, identifying significant differences among samples (F_{ST}= 0.406, P < 0.0001). Further subdivision to assess levels of variation within and among two sets of drainages (Lake Michigan vs Lakes Huron, St Clair, and Erie) indicated that variation was largely attributable to differences among samples within those two drainage groups (F_{SC} = 0.392, P < 0.0001), not differences among them (F_{CT} = 0.024, P = 0.19).

Pairwise estimates of F_{ST} (Table 6) were used to construct a neighbor-joining tree for Johnny darter samples, revealing similarities among the samples within the major basins (Figure

8). Samples from Lake Huron River drainages were similar to each other while those of the other basins had more divergence among populations. Samples from the Muskegon River population, a Lake Michigan drainage, and the Raisin River, a Lake Erie drainage, clustered with samples from the Lake Huron basin instead of more geographically proximate samples because of the high frequency of haplotype A (Table 4, Figure 7). Drainages of Lake St. Clair and Lake Erie were generally intermediate to those from Lakes Huron and Michigan; however, samples from these basins and Lake Michigan have long terminal branches reflecting the high frequency of private alleles at these locations.

DISCUSSION

Using a variable mitochondrial DNA marker, I explored the patterns of genetic variation in round goby and Johnny darter. I predicted that as a recently introduced species round goby would show limited divergence among populations and low genetic diversity. As a native species with more limited dispersal, I predicted greater diversity and more divergence among Johnny darter populations. Results were consistent with these expectations. I also predicted that Johnny darter populations of east and west basins would be significantly different due to distinct colonization sources, but eastern and western lake basins were not significantly different. Despite this result, the distribution of genetic diversity provided insight into the geologic and ecological factors that may have shaped the population structure in Johnny darter.

Round goby

Round goby ND2 sequences showed substantially less diversity than Johnny darter. Invasion of Michigan's streams by round goby is recent, and unlike invasion of the Great Lakes, is not aided by ballast water movement of propagules. Movement through the Great Lakes by ballast has allowed round goby to retain levels of genetic variation on par with source populations

(Brown and Stepien 2009). Studies of population genetics of round goby in more recently established stream populations show that round goby is expanding its range without founder effects (Bronnenhuber et al. 2011). The small number of ND2 haplotypes and dominance of single haplotype identified in round goby is consistent with previous work examining levels of mtDNA variation in round goby (Stepien and Tumeo 2006, Brown and Stepien 2008).

I found few significant differences among populations of round goby and no significant geographic structure. Brown and Stepien (2008) found limited structure among round goby populations in the Great Lakes, with this result primarily driven by differences in Lake Ontario and Saginaw Bay which were not included in this study. Additional statistically significant structure in these previous studies may also have been driven by low frequency private alleles in lake populations which may have been missed here due to the small number of samples and their sizes, or that these rare alleles may not yet be present in these more recently established stream populations.

Round goby has dispersed rapidly throughout the Great Lakes, resulting in high levels of gene flow, and the homogeneity and low diversity of samples around the Great Lakes is consistent with expectations. Established populations of round goby are less prone to dispersal and favor residents (Thorlacius et al. 2015); however, even very low levels of migration can provide sufficient gene flow to prevent population divergence at neutral loci (Slatkin 1985, Newman and Tallmon 2001). As a high dispersing species not restricted by habitat, migration will likely be a persistent, homogenizing force in round goby. This situation has been aided by human activity as secondary spread of round goby to Michigan's inland waterways as anglers often use round gobies as live bait and release them. Populations founded exclusively by bait release where barriers such as dams prevent natural migration will likely be distinct from connected populations (Saard et al.

2018). Where migration and human activity are a constant source of gene flow, the lack of structure observed here is likely to persist.

Johnny darter

As a native species, the distribution of Johnny darter has been shaped, in part, by the action of glaciers, wherein ancestors likely colonized the early Great Lakes more than 10,000 years ago as the glaciers receded, and the modern lake basins were formed (Bailey and Smith 1981). In addition, Johnny darter tends to be habitat restricted, occupying shallow, slow-moving stretches of streams and rivers; therefore, they are less likely to move among rivers and more likely to diverge (Tibbets and Dowling 1996). Johnny darter also have relatively low fecundity with males guarding nests, characteristics which may limit gene flow (Turner and Trexler 1998, Stepien et al. 2007). This is reflected in the AMOVA, where significant differences were mainly driven by variation among individual river drainages within lakes basins, but not between east and west basins as expected. Westbrook (2012) identified a similar pattern in Wisconsin drainages of the Great Lakes, in which geographic isolation resulted in limited gene flow among Johnny darter populations.

For species limited in their ability to disperse by suitable habitat, isolation of individual populations can lead to divergence. Dowling et al. (2015) highlighted the importance of local forces in shaping population structure among populations of three species of roundtail chub (*Gila robusta*), noting high levels of population level divergence that would obscure broader historical factors. It is possible that the east and west basins of the Great Lakes were colonized by distinct Johnny darter populations, but that local forces have over time driven divergence among populations such that differences among broader hierarchical categories cannot be detected (Hedrick 1999). Similar patterns have been observed for other species including smallmouth bass

(*Micropterus dolomieu*, Stepien et al. 2017), white sucker (*Catostomus commersoni*, Lafontaine and Dodson 1997), and mottled sculpin (*Cottus bairdii,* Homola et al. 2016).

Despite isolation by limited dispersal ability, shifts in hydrology and geologic features may allow paths of dispersal that facilitate gene flow, and these shifts may explain some of the patterns observed in Johnny darter. Ray et al. (2006) examined mtDNA variation among rainbow darter (*Etheostoma caeruleum*) populations (an ecologically similar species to Johnny darter) and found that sampled Wabash River populations grouped with both Lakes Michigan and Erie. Stream capture of parts of the Wabash River (a tributary of the Ohio River) by the Maumee River (Figure 1) may have facilitated gene flow between the two basins, as evidenced by shared haplotypes with Stony Creek (Lake Erie drainage) (Figure 6). Johnny darter populations appear to have followed a similar colonization pattern in which colonization of Lake Erie occurred via the captured portion of the Wabash river and propagated throughout the Lake Erie watershed, reducing structure based on east/west basins.

Construction of a neighbor-joining tree among Johnny darter population samples also revealed similarities among samples within major basins, especially Lake Huron; however, one Lake Michigan drainage (Muskegon River) grouped with Lake Huron drainages (Figure 8). Given the close proximity of the headwaters of the Muskegon River to the headwaters of several Lake Huron drainages (Figure 1) and the low, marshy topography of the region, it is possible that connections historically existed and allowed gene flow between basins. During high water events, connections may temporarily exist in the present day.

Additional unexpected patterns were noted in the maximum-likelihood tree of Johnny darter haplotypes and related taxa (Figure 6). The clustering of *E. podostemone* within Johnny darter haplotypes is inconsistent with current darter phylogenies. Darters are a highly speciose

group and their taxonomy is not fully resolved (Near et al. 2011, MacGuigan and Near 2018). The clade which includes Johnny darter (*E. nigrum*) has been studied, yielding two major lineages, one containing *E. nigrum, E. olmstedi*, and *E. perlongum*, and the other *E. podostemone* and *E. longimanum* (MacGuigan and Near 2018). Results from sequencing ND2 are not generally consistent with this result (Figure 6), possibly reflecting differences in patterns on inheritance and introgression among species. Understanding reasons behind the discordance between mtDNA and nuclear genes requires further study.

The position of *E. podostemone* is especially interesting as it is found within a group of Johnny darter haplotypes. *Etheostoma podostemone* is localized to a small and isolated region in Virginia and North Carolina, distant from the localities sampled here. Heckmann et al. (2009), using nuclear genes and mitochondrial gene cytochrome B, found a similar pattern of nesting of *E. podostemone* among *E. nigrum* haplotypes from the Mississippi River, Mobile River, and the Great Lakes. They propose this pattern may reflect ancient mitochondrial introgression. Introgression has often been identified as an important mechanism in the evolution of *Etheostoma* species (Ray et al. 2008, Bossu and Near 2009, MacGuigan and Near 2018).

Conclusions

I predicted that, as a recently introduced species, round goby populations would show low diversity and no geographic structure. Results were consistent with this hypothesis. Round goby populations were dominated by a single haplotype and few statistically significant differences were detected between pairs of populations using F-statistic analysis. This result reflects the history of round goby in the Great Lakes, having been founded from a single source population and rapid dispersal through the region.

The native Johnny darter is more likely to exhibit geographic structure and high levels of genetic diversity, shaped by past geological and evolutionary history. This study identified differences among populations within basins as a significant source of variation with limited divergence among regions. Examination of the distribution of haplotypes did identify evidence for historic stream capture among regions as an important factor in shaping the present day structure of Johnny darter populations in the Great Lakes region. Given this pattern was consistent with that of an ecologically similar species (*E. caeruleum*) it may follow for other similar species as well.

The results of this study inform subsequent chapters exploring variation in patterns of gene expression in round goby and Johnny darter. Historic factors can contribute to patterns of gene expression wherein populations on divergent evolutionary trajectories are more likely to have distinct gene expression patterns (Whithead and Crawford 2006). The populations included in the following chapters are the Rouge and Clinton Rivers. For round goby, historical factors are not likely important in gene expression patterns as this analysis failed to identify differences among these populations, consistent with their recent origin from a single source population. In the case of Johnny darter, I found significant divergence among river populations, suggesting that local forces are likely to have more influence than historic processes. Pairwise F_{ST} for the Rouge and Clinton rivers for Johnny darter was significant (F_{ST} =0.627, $P < 0.0001$), resulting from the high prevalence of private alleles at both sites (R and L in the Clinton River sample and T in the Rouge River sample). Although this results in high F_{ST}, both alleles are a single bp change from the most common allele, A (Figure 5). Given that these alleles are not divergent, these results do not likely reflect deep historical differences, but instead are driven by more recent local forces. Therefore, differences in patterns of gene expression within and among these species likely reflects forces such as founder effect in round goby and drift and/or local adaptation for Johnny darter.

TABLES AND FIGURES

Table 1. Sample size (N) and haplotype counts for round goby samples

			Haplotype				
Major Basin	**Locality**	**N**	**A**	**B**	**C**	**D**	**E**
Lake Erie	Stony Creek	30	30				
	Lower Rouge River	30	30				
Lake St. Clair	Clinton River	19	19				
Lake Huron	Rifle River	4	4				
	Au Sable River	20	19	1			
	Oqueoc River	40	39		1		
Lake Michigan	Manistee Lake	5	5				
	Penwater Lake	25	24		1		
	Muskegon River	20	17				3
	Crockery Creek	20	19		1		
	Kalamazoo River	14	14				
	St. Joseph River	20	17		2	1	

Table 2. Estimates of round goby molecular diversity for each of the drainages sampled. N and N_h refer to the number of individuals and number of haplotypes per sample, respectively. Standard deviations are provided for each estimate.

Major Basin	Locality	N	N_h	Gene Diversity	Mean No. of Pairwise Differences
Lake Erie	Stony Creek	30	1	0.000 +/- 0.000	0.000 +/- 0.000
	Lower Rouge River	30	1	0.000 +/- 0.000	0.000 +/- 0.000
Lake St. Clair	Clinton River	19	1	0.000 +/- 0.000	0.000 +/- 0.000
Lake Huron	Rifle River	4	1	0.000 +/- 0.000	0.000 +/- 0.000
	Au Sable River	20	2	0.100 +/- 0.088	0.100 +/- 0.178
	Oqueoc River	40	2	0.050 +/- 0.047	0.150 +/- 0.216
Lake Michigan	Manistee Lake	5	1	0.000 +/- 0.000	0.000 +/- 0.000
	Pentwater Lake	25	2	0.080 +/- 0.072	0.240 +/- 0.285
	Muskegon River	20	2	0.268 +/- 0.113	0.268 +/- 0.306
	Crockery Creek	20	2	0.100 +/- 0.088	0.300 +/- 0.326
	Kalamazoo River	14	1	0.000 +/- 0.000	0.000 +/- 0.000
	St. Joseph River	20	3	0.279 +/- 0.124	0.647 +/- 0.524

Table 3. Pairwise estimates of F_{ST} for round goby populations. Significant values are identified by the asterisk.

	St. Joseph River	Crockery Creek	Kalamazoo River	Manistee Lake	Muskegon River	Pentwater Lake	Au Sable River	Clinton River	Rouge River	Oaqueoc River	Rifle River
Crockery Creek	-0.019										
Kalamazoo River	0.047	-0.019									
Manistee Lake	-0.048	-0.105	0.000								
Muskegon River	0.084*	0.053	0.074	-0.024							
Pentwater Lake	0.001	-0.046	-0.026	-0.107	0.059						
Au Sable River	0.066	0.000	-0.019	-0.105	0.079	-0.004					
Clinton River	0.071	-0.003	0.000	0.000	0.101	-0.011	-0.003				
Lower Rouge River	0.110	0.021	0.000	0.000	0.145	0.007	0.021	0.000			
Oqueoc River	-0.046	-0.029	-0.033	-0.110	0.084*	-0.030	-0.005	-0.021	-0.007		
Rifle River	0.080	-0.138	0.000	0.000	-0.056	-0.138	-0.138	0.000	0.000	-0.142	
Stony Creek	0.110	0.021	0.000	0.000	0.145*	0.007	0.021	0.000	0.000	-0.007	0.000

Table 4. Sample size 'N' and haplotype counts for Johnny darter samples.

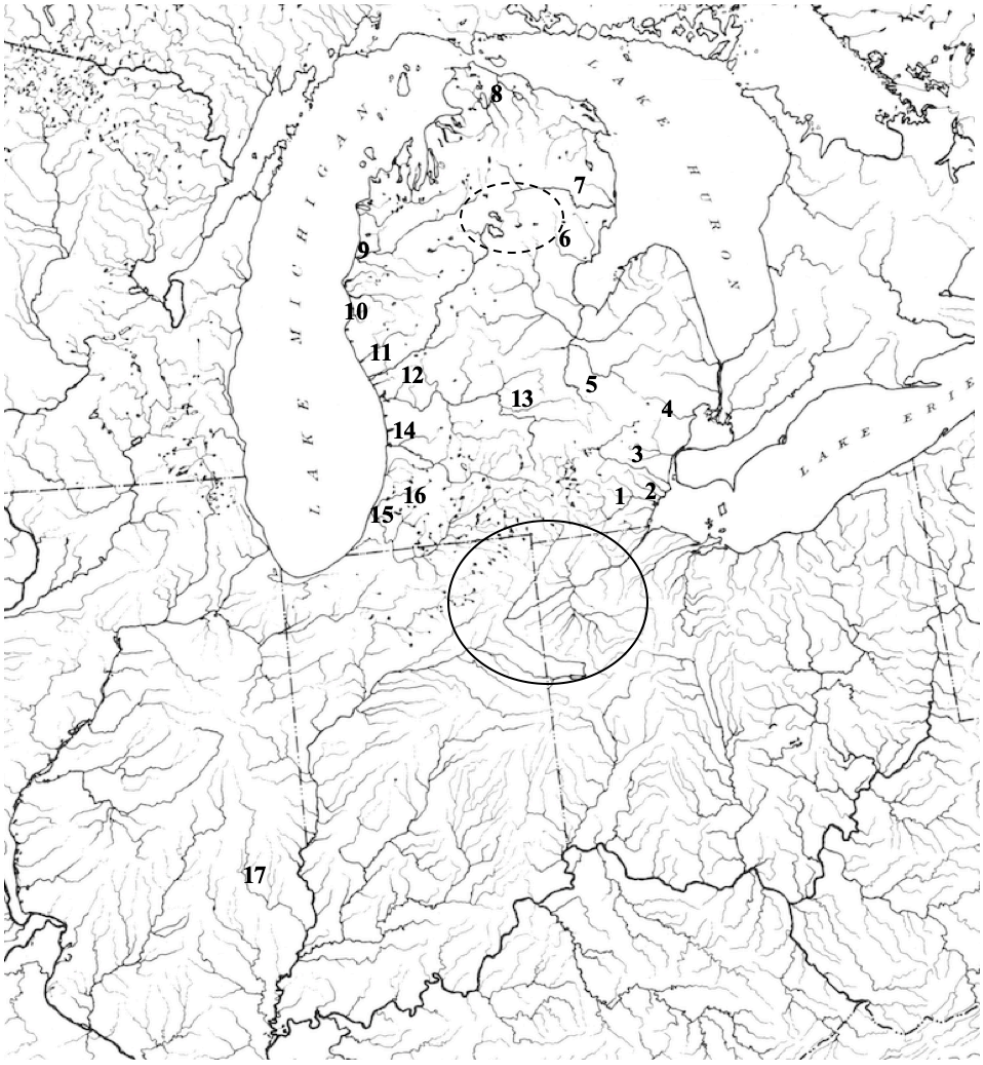
Major Basin	Locality	N	N_h	Gene Diversity	Mean No. of Pairwise Differences
Lake Erie	Stony Creek	30	1	0.000 +/- 0.000	0.000 +/- 0.000
	Lower Rouge River	30	1	0.000 +/- 0.000	0.000 +/- 0.000
Lake St. Clair	Clinton River	19	1	0.000 +/- 0.000	0.000 +/- 0.000
Lake Huron	Rifle River	4	1	0.000 +/- 0.000	0.000 +/- 0.000
	Au Sable River	20	2	0.100 +/- 0.088	0.100 +/- 0.178
	Oqueoc River	40	2	0.050 +/- 0.047	0.150 +/- 0.216
Lake Michigan	Manistee Lake	5	1	0.000 +/- 0.000	0.000 +/- 0.000
	Pentwater Lake	25	2	0.080 +/- 0.072	0.240 +/- 0.285
	Muskegon River	20	2	0.268 +/- 0.113	0.268 +/- 0.306
	Crockery Creek	20	2	0.100 +/- 0.088	0.300 +/- 0.326
	Kalamazoo River	14	1	0.000 +/- 0.000	0.000 +/- 0.000
	St. Joseph River	20	3	0.279 +/- 0.124	0.647 +/- 0.524

Table 5. Estimates of Johnny darter molecular diversity for each sample. N and N_h refer to the number of individuals and number of haplotypes per sample, respectively. Standard deviations are provided for each estimate.

Major Basin	Locality	N	N_h	Gene Diversity	Mean No. of Pairwise Differences
Lake Erie	Stoney Creek	17	5	0.764 +/- 0.065	9.662 +/- 4.660
	Raisin River	13	3	0.410 +/- 0.153	0.436 +/- 0.417
	Lower Rouge River	21	2	0.381 +/- 0.101	0.381 +/- 0.375
Lake St. Clair	Clinton River	25	3	0.477 +/- 0.086	0.873 +/- 0.634
Lake Huron	Shiawasee River	14	3	0.484+/- 0.143	0.527 +/- 0.467
	Rifle River	23	4	0.451 +/- 0.121	0.498 +/- 0.441
	Au Sable River	13	3	0.295 +/- 0.156	0.461 +/- 0.431
	Oqueoc River	11	2	0.182 +/- 0.144	1.272 +/- 0.864
Lake Michigan	Little Manistee River	11	5	0.873 +/- 0.059	5.055 +/- 2.658
	Muskegon River	12	6	0.758 +/- 0.122	1.152 +/- 0.799
	Red Cedar River	15	2	0.133 +/- 0.112	0.133 +/- 0.210
	Kalamazoo River	16	4	0.700 +/- 0.090	1.833+/- 1.111
	Dowagiac River	16	2	0.125 +/- 0.106	0.125 +/- 0.202

Table 6. Pairwise estimates of F_{ST} for Johnny darter populations. Significant values are identified by an asterisk.

	Stoney Creek	Raisin River	Lower Rouge	Clinton River	Shiawasee River	Rifle River	Au Sable River	Oqueoc River	Little Manistee River	Muskegon River	Red Cedar River	Kalamazoo River
Raisin River	0.016*											
Lower Rouge River	0.285*	0.593*										
Clinton River	0.288*	0.420*	0.627*									
Shiawasee River	0.167*	0.067	0.576*	0.418*								
Rifle River	0.223*	0.047	0.569*	0.435*	0.018							
Au Sable River	0.159*	0.028	0.582*	0.412*	0.042	0.023						
Oqueoc River	0.107*	0.023	0.467*	0.350*	0.034	0.040	0.007					
Little Manistee River	0.113*	0.414*	0.555*	0.531*	0.421*	0.493*	0.411*	0.258*				
Muskgon River	0.157*	0.120*	0.508*	0.392*	0.126*	0.140*	0.105*	0.070	0.381*			
Red Cedar River	0.277*	0.763*	0.837*	0.697*	0.734*	0.711*	0.753*	0.592*	0.550*	0.582*		
Kalamazoo River	0.101*	0.585*	0.695*	0.640*	0.584*	0.625*	0.581*	0.470*	0.290*	0.537*	0.717*	
Dowagiac River	0.286*	0.772*	0.842*	0.704*	0.742*	0.719*	0.761*	0.604*	0.562*	0.636*	0.931*	0.725*



1 Raisin*
2 Stony Creek
3 Lower Rouge
4 Clinton
5 Shiawasee*
6 Rifle
7 Au Sable
8 Oqueoc
9 Little Manistee/Manistee Lake
10 Pentwater**
11 Muskegon
12 Crockery Creek**
13 Red Cedar*
14 Kalamazoo
15 St. Joseph**
16 Dowagiac*
17 Embarras River***

*Johnny darter only
**round goby only
***sequence obtained from GenBank

Figure 1. Map of collection localities for round goby and Johnny darter. The dotted circle is the location of the headwaters of the Muskegon River and the Au Sable River. The solid circle shows the region where the Maumee River captured a portion of the Wabash River.

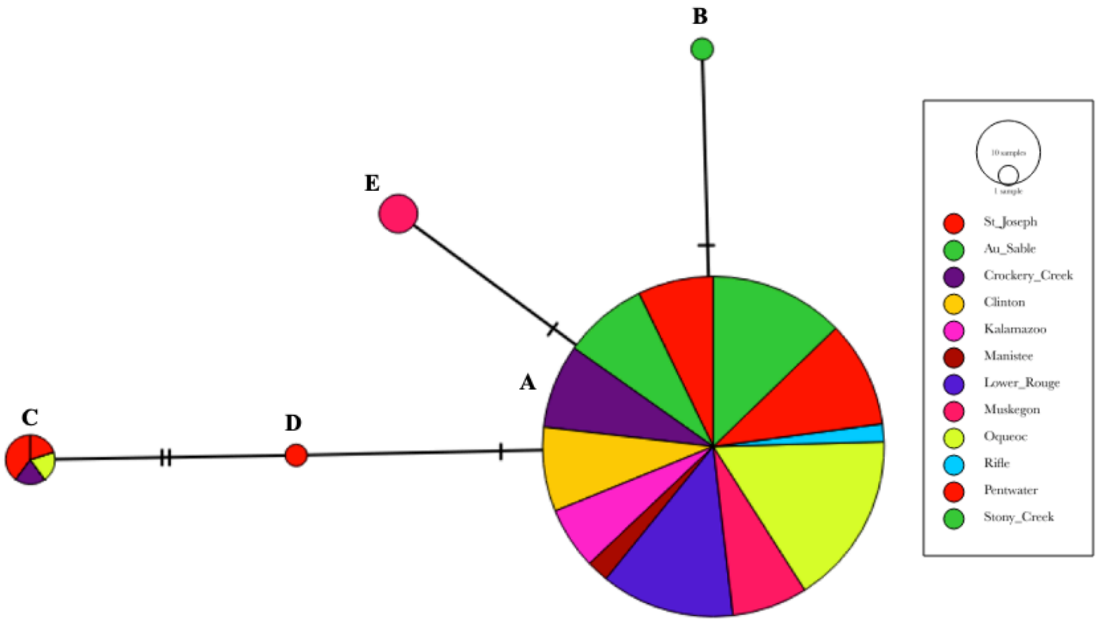


Figure 2. Median-joining network of round goby haplotypes. Pies show proportional representation of haplotypes by locality, with size of the circle reflecting the number of individuals. Mutations between haplotypes are indicated by vertical lines.

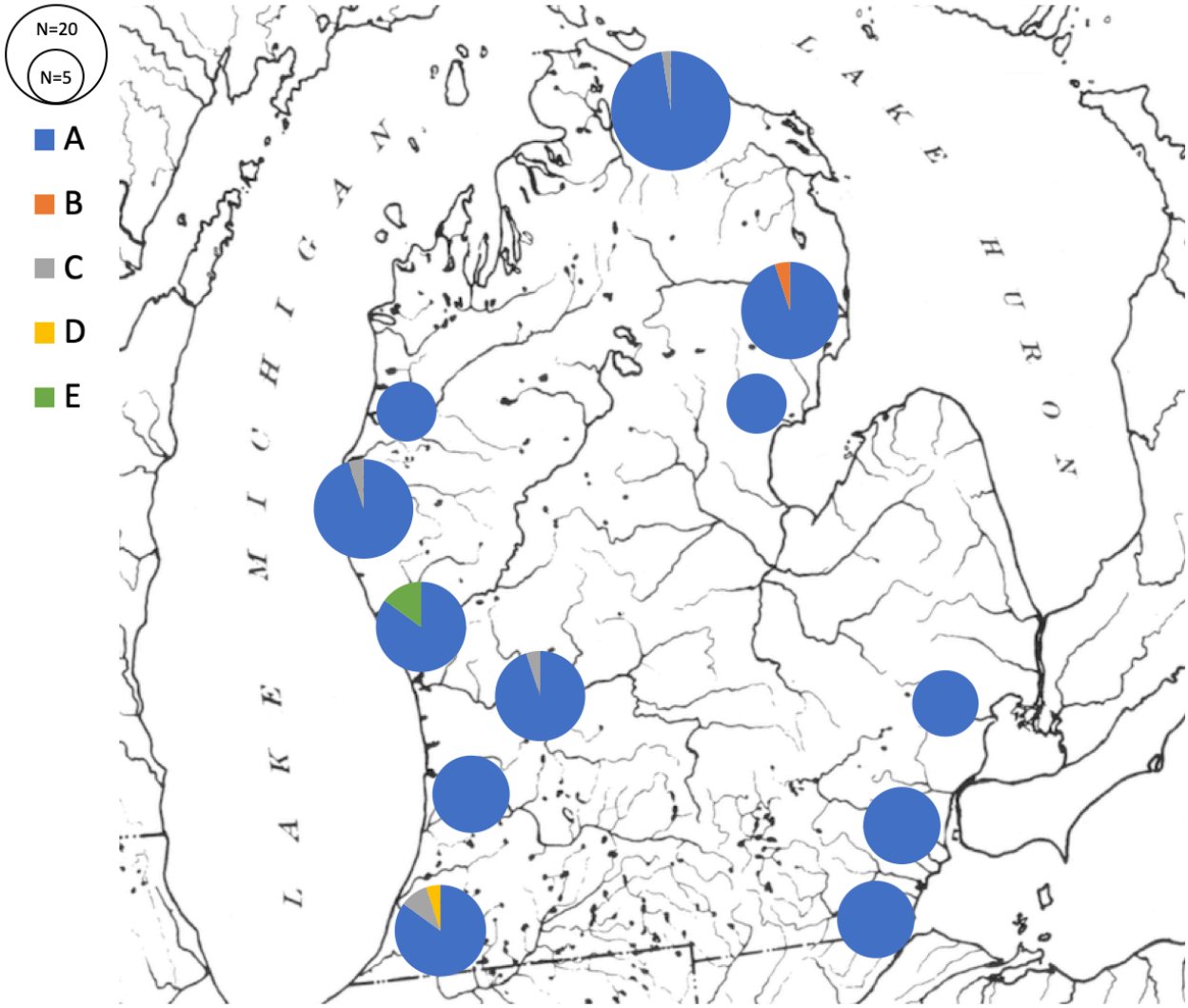


Figure 3. Distribution of haplotypes among sampling localities for round goby. The size of the circle reflects sample sizes and the haplotype frequency is reflected by size of slice.

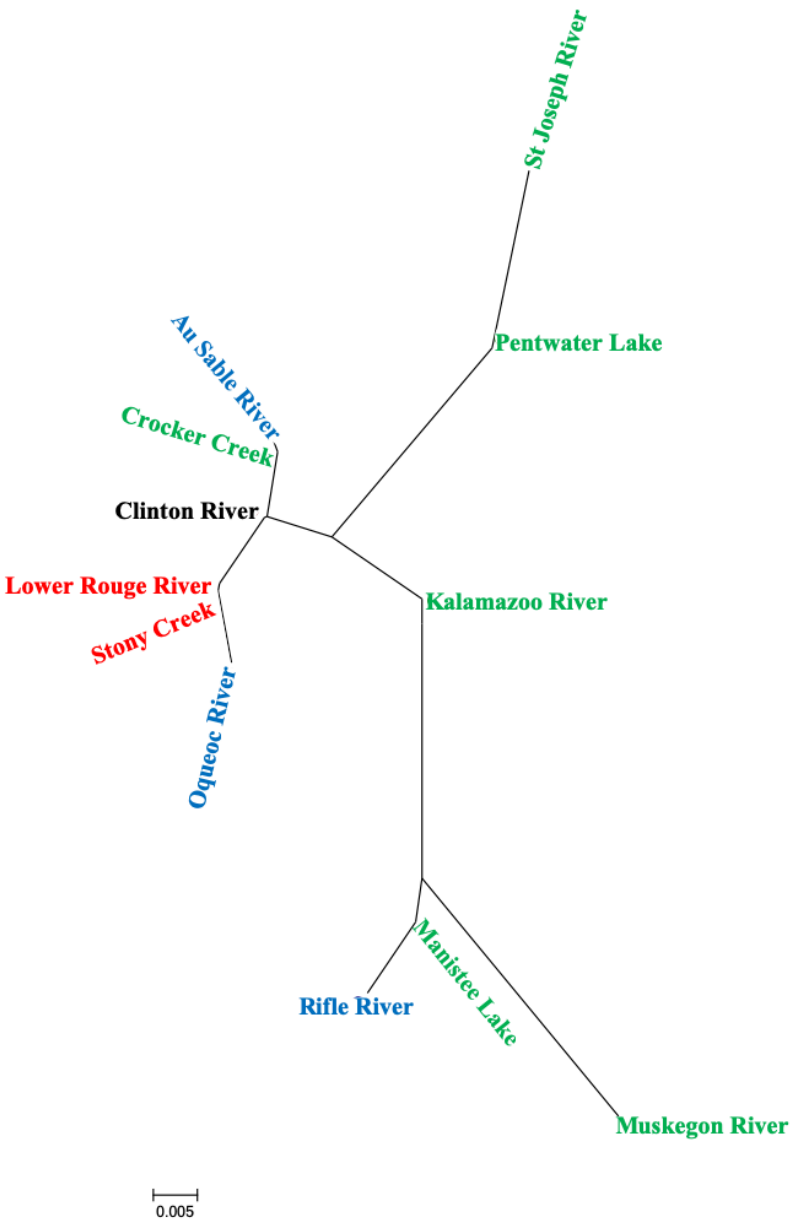


Figure 4. Neighbor-joining tree based of pairwise F_{ST} estimates for round goby. Color indicates major drainage: Black-Lake St. Clair, Green-Lake Michigan, Red-Lake Erie, Blue-Lake Huron. There was no overarching structure among round goby populations.

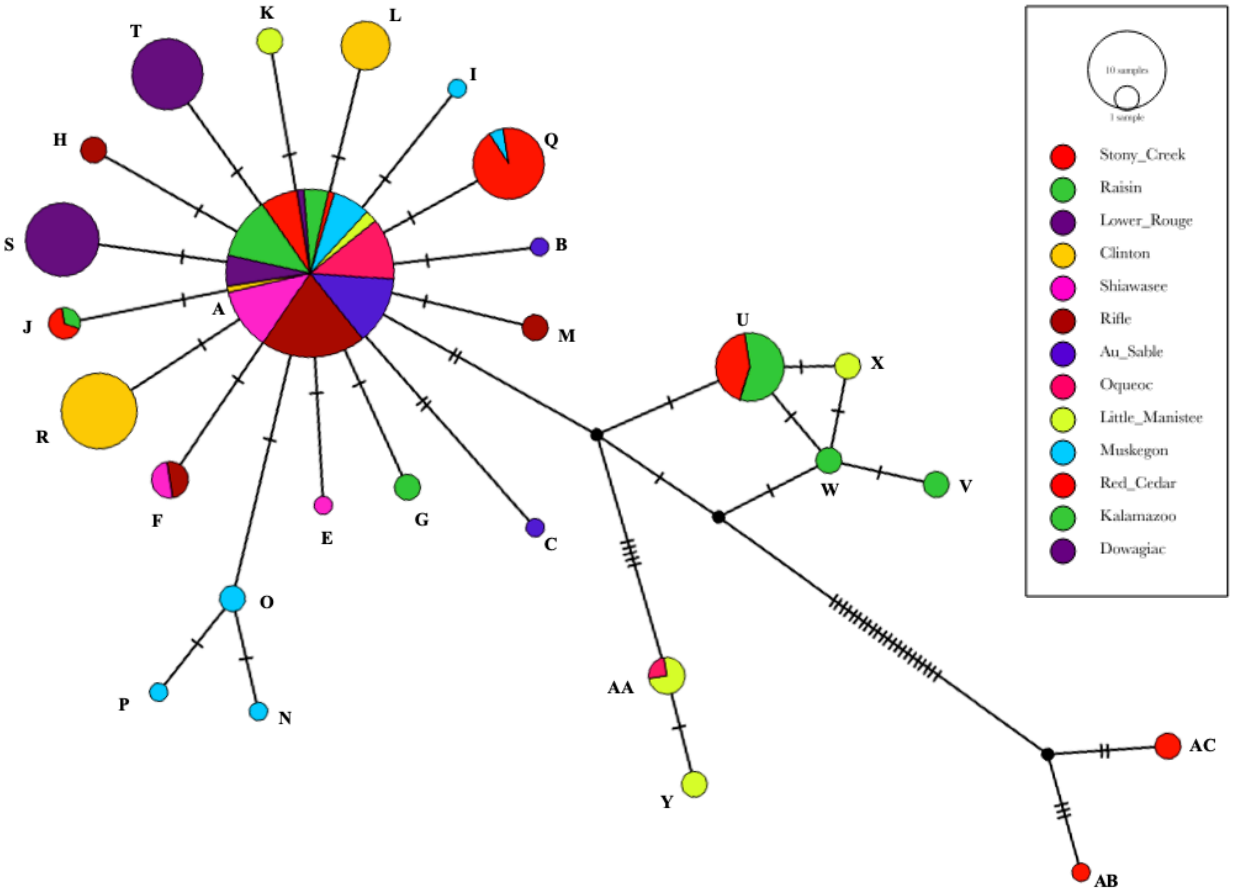


Figure 5. Median-joining network of Johnny darter haplotypes. Pies show proportional representation of haplotypes by locality, and size of the circle reflects number of individuals with that haplotype. Mutations between haplotypes are indicated by vertical lines.

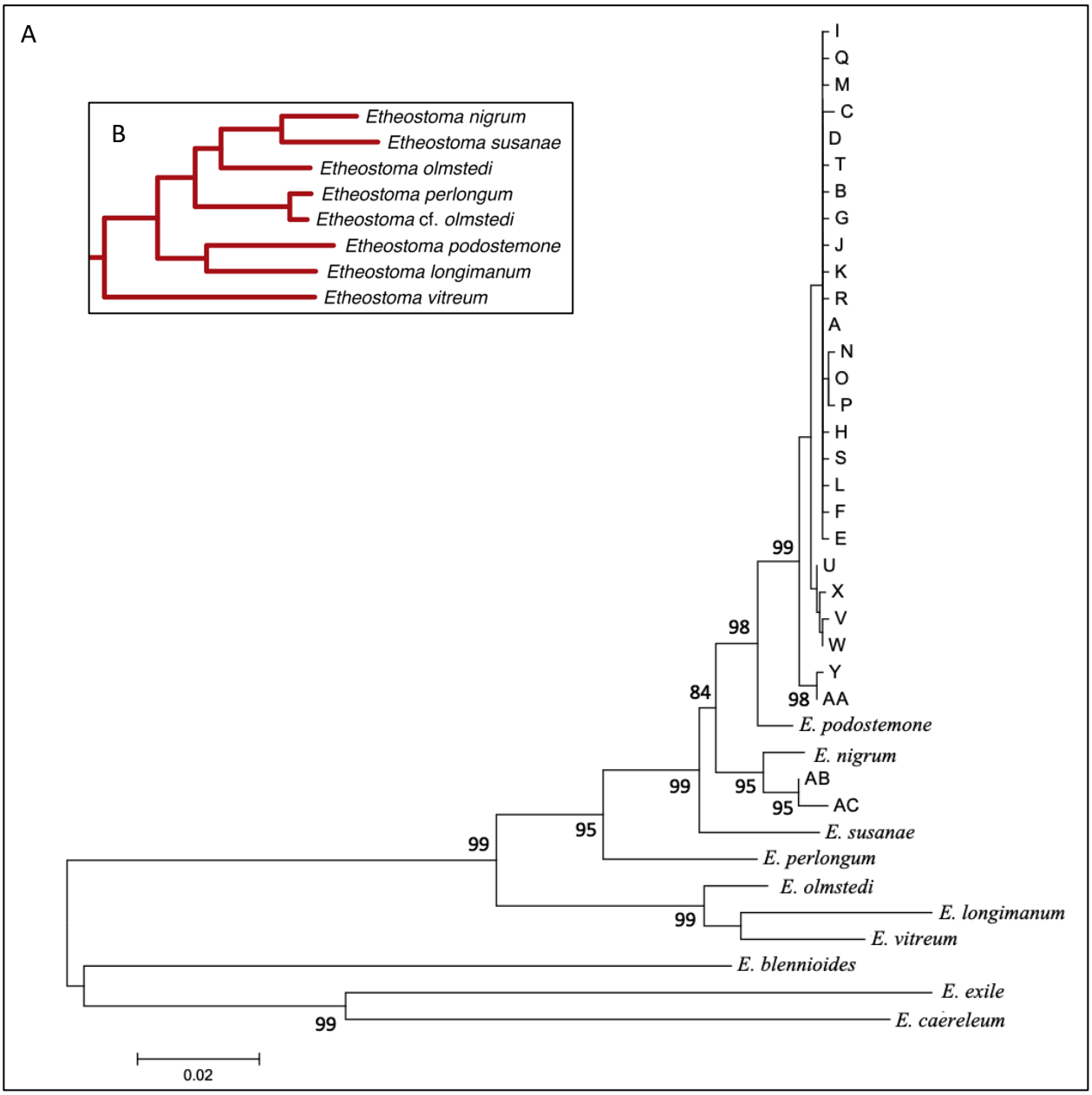


Figure 6. A) Neighbor-joining tree of Johnny darter haplotypes with bootstrap values indicated. The tree includes closely related species *E. olmstedi, E. podostemone*, and *E. perlongum*, *E. susanae, E. vitreum*, *E. longimanum* and more distantly related but co-occurring darter species *E. blennioides, E. exile*, and *E. caereleum*. The additional *E. nigrum* individual was sampled in the Wabash River drainage. B) Phylogeny based on single nucleotide polymorphisms by MacGuigan and Near (2018).

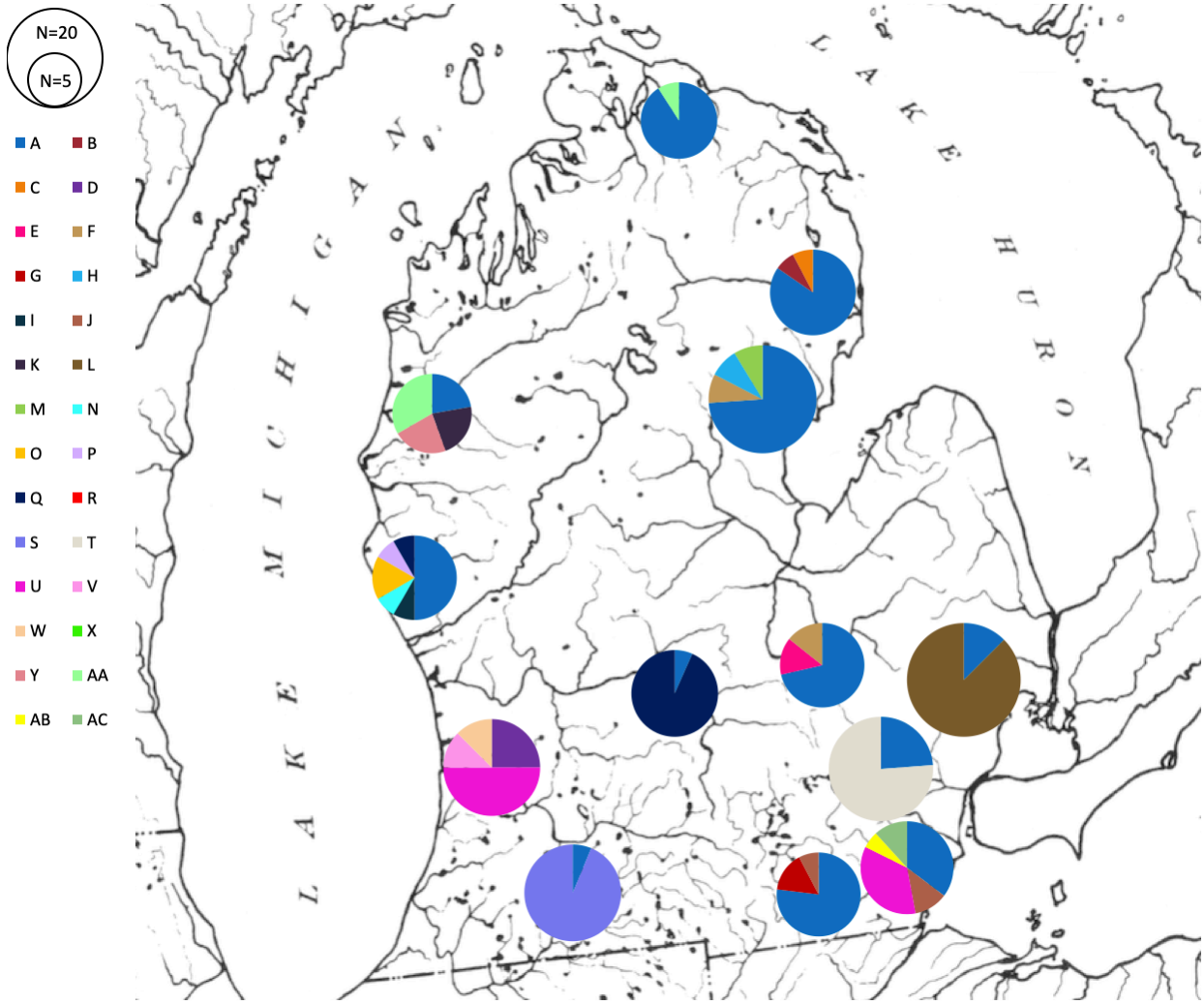


Figure 7. Distribution of haplotypes among sampling localities for Johnny darter. The size of the circle reflects sample size and the haplotype frequency is reflected by size of slice.

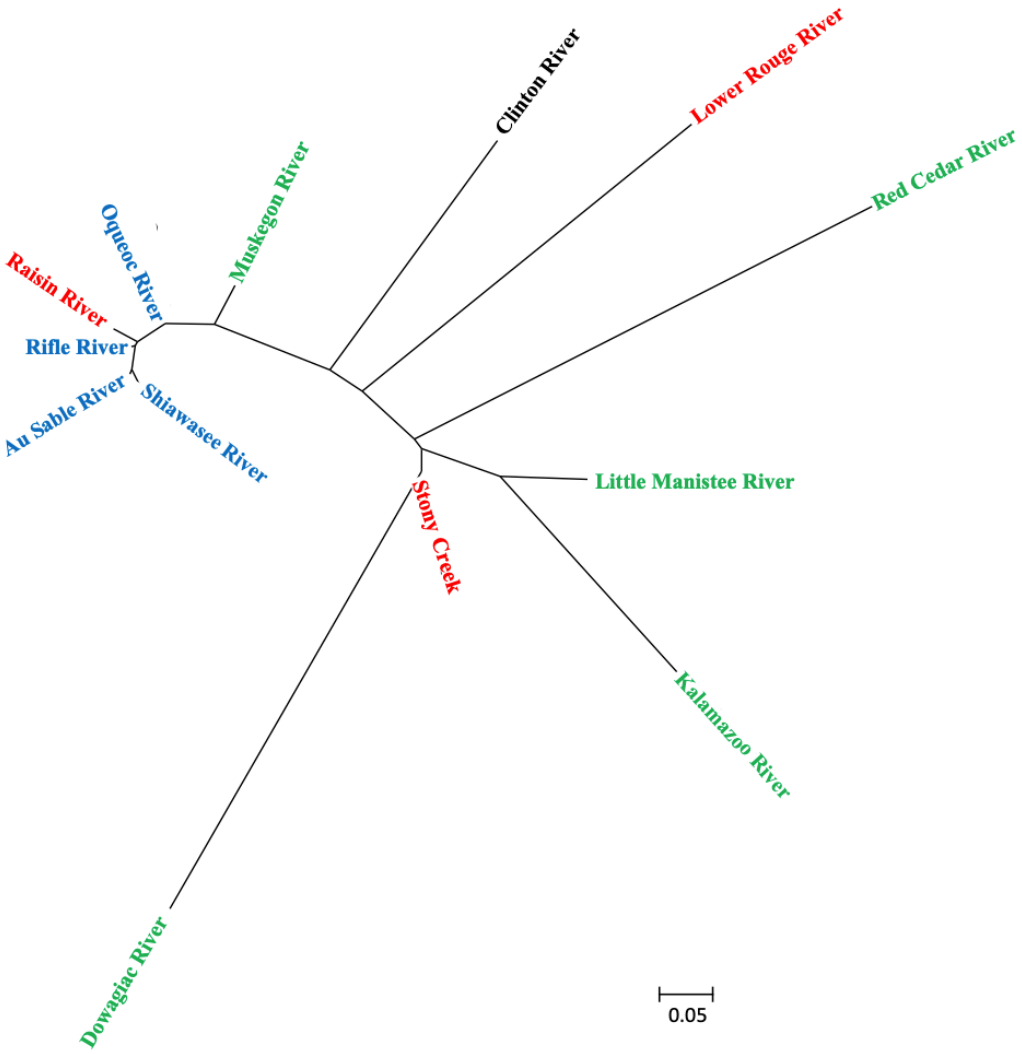


Figure 8. Neighbor-joining tree based of pairwise F_{ST} estimates for Johnny darter. Color indicates major basin: Black-Lake St. Clair, Green-Lake Michigan, Red-Lake Erie, Blue-Lake Huron. There is some clustering by major basin however the Muskegon River (Lake Michigan) and the Raisin (Lake Erie) cluster with the Lake Huron drainages while Stony Creek (Lake Erie) clusters among Lake Michigan drainages.

CHAPTER 2: TRANCRIPTOME VARIATION IN ROUND GOBY AND JOHNNY DARTER IN TWO SOUTHEAST MICHIGAN STREAMS

INTRODUCTION

Organisms vary in their response to abiotic and biotic factors in their environment and these differences will be reflected in an organism's gene expression profile. Habitats vary in resource availability, substrate, temperature, community composition, etc., with each potentially impacting organismal function. In wild populations, environmental heterogeneity is likely to maintain variation in different populations and patterns of gene expression should reflect variable environments (e.g. Buckley and Hofmann 2002; Culumber et al. 2014). In addition to environmental heterogeneity, local adaptation and genetic variation will also produce variation in organisms' physiological responses to the environment (Goetz and MacKenzie 2008). Differential expression analysis is used to identify genes that are significantly up or down regulated between groups of interest, providing an opportunity to quantify differences in how individuals, populations, and/or species respond to their environment.

Next generation sequencing technology has now made it possible to study ecological interactions through gene expression (Ekblom and Galindo 2011), providing a perspective of how physiological processes can change in varying environments of natural systems. RNA-seq is a recently developed tool for studying gene expression by sequencing all available transcripts from one or more tissues (e.g., the transcriptome). The RNA-seq workflow (Wang et al. 2009) involves isolating RNA from a specific tissue (i.e., liver), preparing a library of sequences by converting purified RNA to cDNA, adding an adapter for sequencer recognition of individual samples, and amplification. The library is sequenced using high-throughput sequencing technology, generating millions of short read (50-250 bp) sequences that are randomly distributed throughout the

transcriptome. Reads are mapped to a reference genome if available, or mapped to a reference transcriptome assembled from available transcripts. Here, differential expression analysis was applied to two common fishes of the Great Lakes Region, the native Johnny darter (*Etheostoma nigrum*) and invasive round goby (*Neogobius melanostomus*), in two Southeastern Michigan streams, the Rouge and Clinton rivers. This study utilizes RNA-seq technology to explore the patterns of gene expression in Johnny darter and round goby with the ultimate goal of understanding the mechanisms that drive the patterns observed at the population level.

The round goby is an invasive fish present in all five Great Lakes and is becoming increasingly common in their tributaries (Poos et al. 2010). Likely introduced via ballast water, round goby was first discovered in the St. Clair River in 1990 (Jude et al. 1991). Round goby has expanded rapidly throughout the Great Lakes system and its impact on fish communities in newly invaded streams is not well understood (Kornis et al. 2012). The relatively common Johnny darter (*Etheostoma nigrum*) is a native competitor with round goby. Given its niche overlap with round goby, Johnny darter is an ideal species for comparative study.

As a recent invader to the Great Lakes (Jude et al. 1991) and even more recently to inland streams, limited local adaptation is expected in round goby in the study streams (Bronnenhuber et al. 2011). Persistent propagule pressure is also likely to limit local adaptation and differentiation between source and riverine populations (Lockwood et al. 2005). Population genetic studies of round goby have found the Great Lakes population is derived from a single source population of the Black Sea (Brown and Stepien 2009). Further, the source populations for the Rouge and Clinton rivers, Lake Erie and Lake St. Clair, are not genetically distinct and also likely derived from a single source (Brown and Stepien 2009, Chapter 1). Therefore, with little genetic differentiation between these two rivers for round goby and persistent propagule pressure, local

adaptation should be limited and round goby populations of the two rivers are expected to have similar gene expression patterns.

Additionally, round goby is thought to be a generalist, exploiting many different habitats and resources. This research will explore how this strategy manifests at the gene expression level. A generalist species may show rapid acclimation to changing conditions (phenotypic plasticity) or have a static strategy that is successful across many environments. The expected gene expression patterns from these alternatives give rise to two hypotheses: 1) gene expression patterns will be significantly different in different environments, or 2) gene expression patterns will be similar across all environments. Phenotypic plasticity, although an intuitively appealing trait for an invader, is energetically costly and is not always associated with fitness advantages (Dewitt 1998, Richards et al. 2006). A common strategy that is successful in many environments could be advantageous for an invader; however, there is some evidence that phenotypic plasticity may be important to round goby's success (Wellband and Heath 2017).

Johnny darter is a native species; therefore, there is greater potential for this species to have adapted to local conditions in each stream. Although Johnny darter does have overlap in diet and habitat with round goby, its niche is narrower and gene expression patterns are expected to represent that of a more specialized species as compared to round goby (Paine et al. 1982, Pratt and Lauer 2013). Unique adaptation to local conditions in each stream may result in significant differences in gene expression between individuals from the two rivers. At sites within the same stream, factors like water chemistry, temperature, hydrology, biota, etc. are likely to be more similar than between streams. Thus gene expression patterns are predicted to be more similar within than between rivers.

In addition to comparing round goby and Johnny darter in their response to environmental variation, this study will compare populations of Johnny darter in the presence and absence of round goby. The presence of an exotic competitor presents a unique stressor for native species. Round goby and Johnny darter are likely to have niche overlap resulting in competitive interactions (Kornis et al. 2014), and there are examples where Johnny darter populations have declined in correlation with round goby invasion (Lauer et al. 2004). Laboratory experiments have shown that round goby is generally more aggressive than native species, thus round goby may present a novel competitive interaction for Johnny darter (Dubs and Corkum 1996, Balshine et al. 2005). The presence of a novel competitor may result in a shift in diet, habitat, or other physiological resources resulting in a shift in gene expression patterns. Energy directed toward competitive interactions may reduce an organism's efficiency in coping with other environmental stressors like pollutants and pathogens, potentially affecting their fitness (Smith et al. 2017, He et al. 2018). Given this difference in biotic environment, it is predicted that populations of Johnny darter in the presence of round goby will have gene expression patterns distinct from populations where round goby is absent.

METHODS

Sampling

Up to seven individuals of each species were collected from two sites on each of the Lower Rouge and Clinton rivers in May/April of 2015-2017 (Figure 9, Table 7). Darters were sampled at both sites on each river. The fish community at each site was sampled with a 3x1.25m nylon mesh seine (3.18mm mesh). Sampling was broad for characterization of community composition and was augmented with targeted sampling for the study species as necessary (Krabbenhoft 2019). Sites were chosen to sample Johnny darter populations in the presence and absence of round goby.

Gobies were only available at the downstream site on the Clinton, but had expanded to the upstream site of the Rouge River in 2017 and were collected accordingly. Sampling of the Lower Rouge River site and both Clinton River sites in 2017 did not yield enough individuals to be included in the study.

Fish were dissected streamside. Individuals were sacrificed by quickly severing the spinal column with a fresh razorblade and liver tissue was removed and placed in a tube with 0.5 mL of RNAlater. Samples were stored on ice until transfer to the lab freezer (-20°C) later the same day. Carcasses were preserved for further study (Krabbenhoft 2019). Care was taken to handle each fish in the same manner, with dissecting tools cleaned or changed between each fish and alternating between dissecting gobies and darters.

RNA Purification, Library Construction, and Sequencing

RNA was purified from liver tissue using the PureLink RNA mini kit (Thermo Fisher). Liver tissue was mechanically homogenized in TRIzol (Invitrogen) and purification was completed according to the manufacturer protocol with the addition of a DNase incubation to reduce contaminating genomic DNA. Quality and quantity were checked on an Agilent 2100 Bioanalyzer prior to sequencing. All samples had RNA Integrity Numbers (RINs) that were of sufficient quality for sequencing, with most > 9 (Table S1).

Samples were sent for library construction and sequencing by the RTSF Genomics Core at Michigan State University. Library prep was completed with Illumina TruSeq Stranded mRNA Library prep kit and a total of 85 individuals were sequenced in two pools over four lanes on Illumina HiSeq 4000 with 150 bp paired end reads.

Data Analysis

Genomes are not available from these species or closely related individuals to use as reference; therefore, reference transcriptomes were assembled *de novo.* Reads were trimmed with Trim Galore (Krueger 2015) to remove low quality reads and adapters. Reference transcriptomes were assembled with Trinity version 2.6.6 (Haas et al. 2013). Reads were mapped to the assembled reference with Bowtie2 (Langmead and Salzberg 2012) and assembled into a count matrix with RSEM (Li 2011).

Differential expression analysis was completed with the R package DESeq2 to identify differentially expressed genes between sampled groups (Love et al. 2014). Genes were filtered by including only genes with at least three individuals having minimum of 10 reads. Sex and library pool were included as covariates in the differential expression analysis, and Surrogate Variable Analysis (SVA) was included to identify and model any unrecognized batch effects (Leek et al. 2012), and it was limited to four surrogate variables. The developers of DESeq2 recommend this approach as opposed to removing batch affects prior to analysis.

Principal component analysis (PCA) was used to identify major sources of variation in gene expression. For visualization in principal component plots, the protocol outlined in Jaffe et al. (2015) using SVA was used to produce plots corrected for batch effects. NCBI's BLAST was used to annotate genes against the zebrafish (*Danio rerio*) genome. Q-values (Dabney et al. 2015) were used for control of false discovery rate with < 0.05 considered significant. For significant genes, those with available zebrafish annotations were uploaded to DAVID 6.8 (Huang et al. 2009) for gene ontology annotation and enrichment analysis. Gene ontology plots were created with GOplot (Walter et al. 2015).

RESULTS

Round Goby

Thirty-three round gobies were sequenced on Illumina Hi-seq 4000, generating a total of 568,826,171 reads (Table S1). A targeted minimum sequencing depth of 10,000,000 reads per individual was achieved. Variation in sequencing depth and RIN number was examined as source of expression variation and an effect was not observed. Trimmed reads were assembled *de novo* to a reference transcriptome resulting in 74,891 contigs with an N50 of 1260 bp (i.e. 50% of contigs are of this length or longer). Contigs were blasted against the zebrafish genome resulting in 34,313 annotated transcripts.

After filtering for low read counts transcripts, 31,001 transcripts were included in the analysis. DESeq2 uses Cook's distance to calculate each sample's influence on the fitted coefficients for genes. These distances were averaged and plotted in Figure S2. One outlier individual (RG17_006) was identified and removed from the analysis.

PCA was used to identify patterns of variation. PC1 accounts for 43.2% of the variance and correlates with differences between males and females (Figure 10A). Genes loading on PC1 are consistent with sex differences expected during spawning (e.g., vitellogenin, zona pellucida). The samples cluster distinctly by drainage along PC2, explaining 5.7% of the variance (Figure 10A). Clustering by site and year is observed by plotting PC2 and PC3 (Figure 10B). PC3 explains 3.8% of the variance. Within the Clinton River, the 2015 and 2016 samples form distinct groups. In the Rouge River, there is overlap between the Upper Rouge River sample from 2017 and the Lower Rouge River sample from 2015. There does not appear to be an overall pattern of site specific or annual clustering.

Populations were analyzed for differential expression in DESeq2. Sample sites and years within rivers were pooled to compare Rouge versus Clinton river samples, identifying 2660

differentially expressed genes. Of these differentially expressed genes, 1776 were found in the Zebrafish genome, and these were input to DAVID for ontology annotations and enrichment analysis. There were 16 gene ontology terms that were significantly enriched (Figure 11, Table 8). The most significantly enriched gene ontology terms were primarily related to ribosomal protein production and oxidation-reduction. For the cellular component category, overrepresented terms include intracellular ribonucleoprotein complex (GO:0030529, $p=1.79x10^{-5}$), endoplasmic reticulum (GO:0005783, $p=1.98x10^{-4}$), and U1sRNP (GO:0005685, $p=2.04x10^{-4}$). The only significantly enriched molecular function term was oxidoreductase activity (GO:0016491, $p=3.68x10^{-4}$). The most significantly enriched biological process terms were oxidation-reduction process (GO:0055114, $3.04x10^{-3}$), Spliceosomal snRNP assembly (GO:0000387, p=0.003), and protein folding (GO:0006457, p=0.006).

Pairwise comparisons of the number of differentially expressed genes between all groups were calculated (Figure 12A). In general, comparisons within the same river had more differentially expressed genes than between the two rivers. Most similar were Rouge River Lower 2016 and Rouge River Lower 2015 (480 genes) and the Clinton River Lower 2015 and 2016 (625 genes). An exception is Upper Rouge River 2017 which accounts for the two highest gene counts in comparisons with the Rouge River Lower 2016 and Clinton River Lower 2016 (2303 and 1903 differentially expressed genes, respectively). Rouge River Upper 2017 was more similar to Rouge River Lower 2015 (682 genes) and Clinton River Lower 2015 (809 genes).

Johnny Darter

Fifty-two Johnny darter samples were sequenced on Illumina Hi-seq 4000 generating a total of 860,548,239 reads, averaging approximately 16.5 million reads per individual. Assembly of the reference transcriptome resulted in 69,694 contigs with an N50 of 1787 bp. Contigs were

annotated by BLAST against zebrafish genome resulting in hits for 33,774 transcripts, respectively. After filtering out transcripts with less than 10 reads in more than three individuals, 33,559 transcripts were included in the analysis.

PCA was used to identify major sources of variation in the data set (Figure 13). PC1 explains 20.9% of the variation and clusters specimens by drainage (Figure 13A). PC2 explains 14.6% of the variation and corresponds to differences between males and females (Figure 13A). As with round goby, genes loading on this component are consistent with genes expected to be upregulated during spawning (e.g., vitellogenin, zona pellucida). Clustering by site can be visualized by plotting PC2 and PC3, with PC3 explaining 5.5% (Figure 13B). Rouge River individuals cluster by Upper vs Lower sites, correlated with PC3, with the exception of individuals from Rouge River Upper 2017 which form their own cluster. Clinton River samples also roughly cluster by Upper vs Lower sites along PC3 though the separation is not as distinct as the Rouge River. The 95% confidence interval ellipse for Clinton Lower 2015 overlaps both Clinton River Lower 2016 and Clinton River Upper 2016. However, there is not an overall pattern of clustering by upstream and downstream sites across the two rivers; Rouge River Lower sites cluster in the upper left quadrant and the Upper Rouge River sites cluster in the lower left quadrant while the Clinton River sites show the opposite pattern.

The groups were analyzed for differential expression between all sampling groups and between rivers with all groups (e.g., upper and lower samples) pooled within their respective rivers. There were 1226 genes differentially expressed between the Rouge and Clinton rivers. Of these, 840 had blast hits to the Zebrafish genome and were input to DAVID for gene ontology annotation and enrichment analysis. Compared to 16 enriched terms for round goby, only five GO terms were significantly enriched in Johnny darter (Figure 14, Table 9). Of the enriched terms, two

were related to oxidation reduction: oxidation reduction processes (GO:0055114, $p=1.96x10^{-3}$) in the biological processes category, and oxidoreductase activity (GO:0016491, $p=2.14x10^{-2}$) in the molecular function category. Enriched terms in the cellular processes category were cytosol (GO:0005829, $p=3.33x10^{-6}$), nucleolus (GO:0005730, $p=3.57x10^{-2}$), and endoplasmic reticulum (GO:0005783, $p=3.87x10^{-2}$).

Pairwise comparisons of differentially expressed genes between all groups are reported in Figure 12B. As with round goby, there were generally more differentially expressed genes in comparisons of groups between rivers than comparisons within the same river. Groups within the Rouge River tended to be more similar (257-613 genes) than groups within the Clinton River (658-937 genes). Both Clinton River Upper 2015 and 2016 appear to be the most distinct from almost all groups and account for the highest numbers of differentially expressed genes. Full gene and gene ontology lists are available upon request.

There were 910 genes that were differentially expressed between Johnny darter populations in the presence and absence of round goby. Clustering by goby presence did not appear in any of the principal components examined. Of these 910 genes, 668 had blast hits to the Zebrafish genome and were input to DAVID for gene ontology annotation and enrichment analysis. Unlike analyses by location, there were no significantly enriched categories (Figure 14).

DISCUSSION

In this study I used gene expression to compare how Johnny darter and round goby respond to variation in the environment. Principal component analysis of overall gene expression patterns identified locality specific gene expression patterns in Johnny darter. Johnny darter samples cluster tightly by river (Figure 13A), and also by sites within rivers (Figure 13B), while there is no

apparent clustering by sampling year, suggesting that expression profile is more locality specific. This clustering is less apparent in the Clinton River than the Rouge River. Although round goby also exhibit drainage specific patterns of expression (Figure 10A), they do not show the site specific clustering observed in Johnny darter (Figure 10B), suggesting that gene expression in round goby may be more dependent on transient environmental conditions.

The Johnny darter Rouge River Upper 2017 group stands out in the PCA (Figure 13B). In 2017, round goby increased dramatically in abundance at this site. Only one individual was sampled there in 2015 and 2016, representing 5% of all individuals sampled for both years; however, 58 gobies were collected in 2017, representing 42% of all individuals sampled at that locality in that year (Krabbenhoft 2019). The increase in round goby abundance could have impacted gene expression in Johnny darter, causing the separation of the site in the PCA. This factor, as well as the impact of other environmental variables in gene expression patterns, will be further explored in Chapter 3. The dramatic change in round goby abundance may also explain why Rouge River Upper 2017 is also distinct for round goby in terms of differentially expressed genes (Figure 11). The differences in expression could represent founder effect in these pioneer individuals.

Compared to the native species, round goby showed greater changes in levels of gene expression than the native Johnny darter. This analysis identified a similar number of expressed genes in each species, with 34,313 and 33,774 annotated transcripts (when compared to zebrafish) for round goby and Johnny darter, respectively. Despite this similarity, round goby had greater numbers of differentially expressed genes in 9 of the 10 pairwise comparisons (Table 8) as well as between the two rivers when all individuals were pooled. I predicted that since round goby invaded these streams recently, local adaption is likely to be minimal; therefore, similar gene expression

profiles across the two rivers was expected. However, I found that gene expression was more variable in round goby compared to Johnny darter but was not locality specific as in Johnny darter. This result suggests that round goby may not be genetically adapted to these localities but is able to respond to local environmental conditions with a plastic gene expression profile.

Although round goby had consistently more differentially expressed genes and more significantly enriched ontology terms, there was overlap in the enriched categories of the two species. In the pooled Rouge River versus Clinton River comparison, oxidation-reduction processes (GO:0055114), endoplasmic reticulum (GO: 0005685), and oxidoreductase acidity (GO:0016491) were significantly enriched and upregulated for both species. For Johnny darter, the number of differentially expressed genes for each of these categories were 44, 26, and 37, respectively, while for round goby there were 66, 47, and 60 differentially expressed genes, respectively. Although the two species seem to be responding similarly to environmental heterogeneity, response of round goby is more variable.

Though these results show greater gene expression plasticity in round goby than Johnny darter, the degree to which this expression plasticity translates to physiological plasticity and fitness advantage is unknown. Wellband and Heath (2017) compared gene expression patterns between round goby and a less successful invader, tubenose goby (*Proterorhinus semilunaris*). Tubenose goby is another introduced species to the region, also likely via ballast water, that comes from the same geographic region as round goby. Unlike round goby, tubenose goby has not proliferated rapidly and widely like round goby (GLANSIS). Wellband and Heath (2017) used differential expression analysis to assess the response of round and tubenose gobies to thermal stress. Round goby showed greater transcriptome plasticity as demonstrated with greater changes in gene expression levels in response to thermal stress than tubenose goby. This is consistent with

experimental findings of a larger thermal tolerance range in round goby compared to tubenose goby (Drulliard et al. 2018).

Additional studies in other organisms have documented the translation of gene expression plasticity to physiological plasticity. Lohmen et al. (2017) found that for stickleback transplanted between stream and lake environments, gene expression profiles of migrants shifted closer to that of natives, suggesting gene expression changes may be important in adaptation to the new environments. Cheviron et al. (2008) documented transcriptome plasticity in reciprocal transplants of sparrow populations at high elevation as compared to low altitude populations wherein the populations adapted to high elevations showed greater variation in patterns of gene expression. Although more work is required to understand the implications of gene expression plasticity in round goby, these examples support variability in gene expression translating to variability in physiological responses to different environments.

Plasticity is commonly hypothesized as an important characteristic of successful invasive species. High phenotypic plasticity has been documented in many invasive species, including other examples in round goby. Gutowsky et al. (2012) found plasticity in life history traits (age at maturity, growth rate, reproductive investment) associated with range expansion in round goby. Pettitt-Wade (2015) and Krabbenhoft (2019) both documented diet variability in round goby using stable isotope ratios. Gene expression plasticity observed in this study is consistent with these documented examples of physiological plasticity in round goby.

I also tested the related question, what does a generalist strategy look like at the gene expression level? The variation in gene expression profile is consistent with the hypothesis that gene expression (and presumably physiology) changes in round goby in response to environmental variation, as opposed to the strategy of maintaining a single, common expression profile across

varying environments. In this case, the ability of round goby to be successful across variable regions suggests that this species is reaping fitness benefits of plasticity despite energetic requirements. The link between plasticity and fitness is not straight forward, however, and the impact of plasticity on fitness varies (Richards et al. 2006, Davidson et al. 2011). While the literature on animals is limited, the fitness advantage of plasticity in invasive species appears to emerge in harsh environments (Richards et al. 2012, Xu et al. 2014). Invasion sites for aquatic organisms are often harsh environments, such as urban streams and major shipping routes, with multiple stressors present (Walsh et al. 2005). Such regions may provide an opportunity for highly plastic invaders to thrive and outcompete native species that are struggling to persist in these highly modified environments.

The impact of round goby on populations of Johnny darter was also examined. It was predicted that populations of the native species would differ in overall gene expression patterns in the presence of an ecologically similar exotic species. However, clustering by presence/absence of round goby was not identified in the PCA of overall patterns of gene expression. While no effect was detected, a shift resulting from the presence of an introduced competitor may be quite subtle and overshadowed by other sources of variation among sampling groups such as historic effects and local environmental conditions. There were 910 differentially expressed transcripts between the goby presence/absence groups, however gene ontology enrichment analysis did not reveal enriched pathways. Only one other vertebrate study could be identified that measured the gene expression response to competition. He et al. (2017) examined the transcriptome response of Atlantic salmon to competition by rearing Atlantic salmon with ecologically similar nonnatives. They also found that effects of source population were more significant than effects of competition, with only a small number of differentially expressed genes shared between populations. It may be

that the response of Johnny darter to competition with round goby is similarly population specific and dependent on local conditions. The response of Johnny darter populations to round goby invasion has been variable, with some populations declining with round goby invasion (Lauer et al. 2004, FOTR 2018), while some populations have persisted with no apparent negative impacts (Riley et al. 2008; Kornis et al. 2013). Local environmental conditions are likely an important determinant of the impacts of round goby invasion (Krabbenhoft 2019).

A major goal in conservation ecology is to understand what makes invasive species successful, with a thorough understanding of invasive species dynamics allowing for improved prevention and informed allocation of management resources. This research provides a comparison of the variation in gene expression within and between populations of a native species and an invasive competitor. I predicted round goby would show limited variation in gene expression given its relatively short residency in the Great Lakes region and in particular to the streams sampled for this study. However, round goby showed greater variation in gene expression as compared to the native Johnny darter, indicating rapid acclimation to new environments. The ability to respond rapidly to novel environments likely plays an important role in the rapid range expansion of this species; therefore, consideration of flexibility of gene expression as well as variation among habitats may be important factors when evaluating the probability of establishment of invasive species.

TABLES AND FIGURES

Table 7. Sampling locations for round goby and Johnny darter collections and number of individuals sequenced. Two sites were selected on each river such that Johnny darter could be sampled in the presence and absence of round goby. Up to seven individuals were included from each site. Only the Rouge River upstream sample was included for 2017 as sampling at other sites in 2017 did not yield a sufficient number of individuals to be included in the study.

		E. nigrum			*N. melanostomus*		
Site	**Coordinates**	**2015**	**2016**	**2017**	**2015**	**2016**	**2017**
Rouge Lower	42.285310, -83.388076	5	5	n/a	5	7	n/a
Rouge Upper	42.281972, -83.466054	7	6	7	n/a	n/a	7
Clinton Lower	42.671619, -83.095602	5	5	n/a	7	7	n/a
Clinton Upper	42.665730, -83.156182	5	7	n/a	n/a	n/a	n/a

Table 8. Significantly enriched gene ontology for differentially expressed genes in round goby between the Rouge and Clinton Rivers. Letter and color correspond to labels in Figure 11.

	ID	Description	No. of Genes	BH adjusted p-value
A	GO:0055114	oxidation-reduction process	66	2.60E-03
B	GO:0006457	protein folding	21	8.32E-03
C	GO:0000387	spliceosomal snRNP assembly	8	6.09E-03
D	GO:0071391	cellular response to estrogen stimulus	9	2.22E-02
E	GO:0006412	translation	37	1.82E-02
F	GO:0015031	protein transport	26	2.97E-02
G	GO:0030529	intracellular ribonucleoprotein complex	33	1.79E-05
H	GO:0005685	U1 snRNP	9	2.04E-04
I	GO:0005783	endoplasmic reticulum	47	1.98E-04
J	GO:0005739	mitochondrion	54	7.61E-04
K	GO:0000243	commitment complex	6	3.25E-03
L	GO:0005737	cytoplasm	172	2.81E-03
M	GO:0005840	ribosome	22	1.08E-02
N	GO:0034715	pICln-Sm protein complex	4	2.41E-02
O	GO:0005777	peroxisome	10	2.72E-02
P	GO:0016491	oxidoreductase activity	60	3.68E-04

Table 9. Significantly enriched gene ontology for differentially expressed genes in Johnny darter between the Rouge and Clinton Rivers. Letter and color correspond to labels in Figure 14.

	ID	Description	No. of Genes	BH adjusted p-value
A	GO:0055114	oxidation-reduction process	44	1.96E-03
B	GO:0005829	cytosol	39	3.33E-06
C	GO:0005730	nucleolus	14	3.57E-02
D	GO:0005783	endoplasmic reticulum	26	3.87E-02
E	GO:0016491	oxidoreductase activity	37	2.14E-02

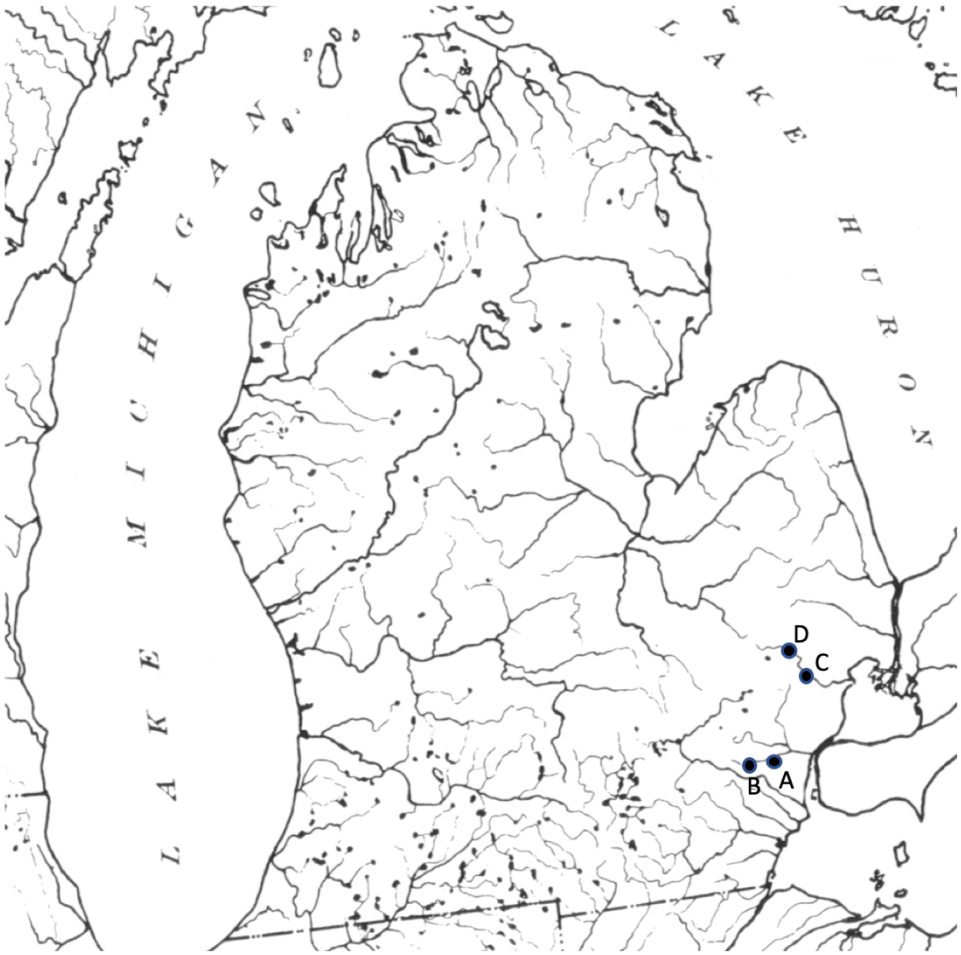


Figure 9. Map of sampling localities. A. Rouge Lower B. Rouge Upper C. Clinton Lower D. Clinton Upper

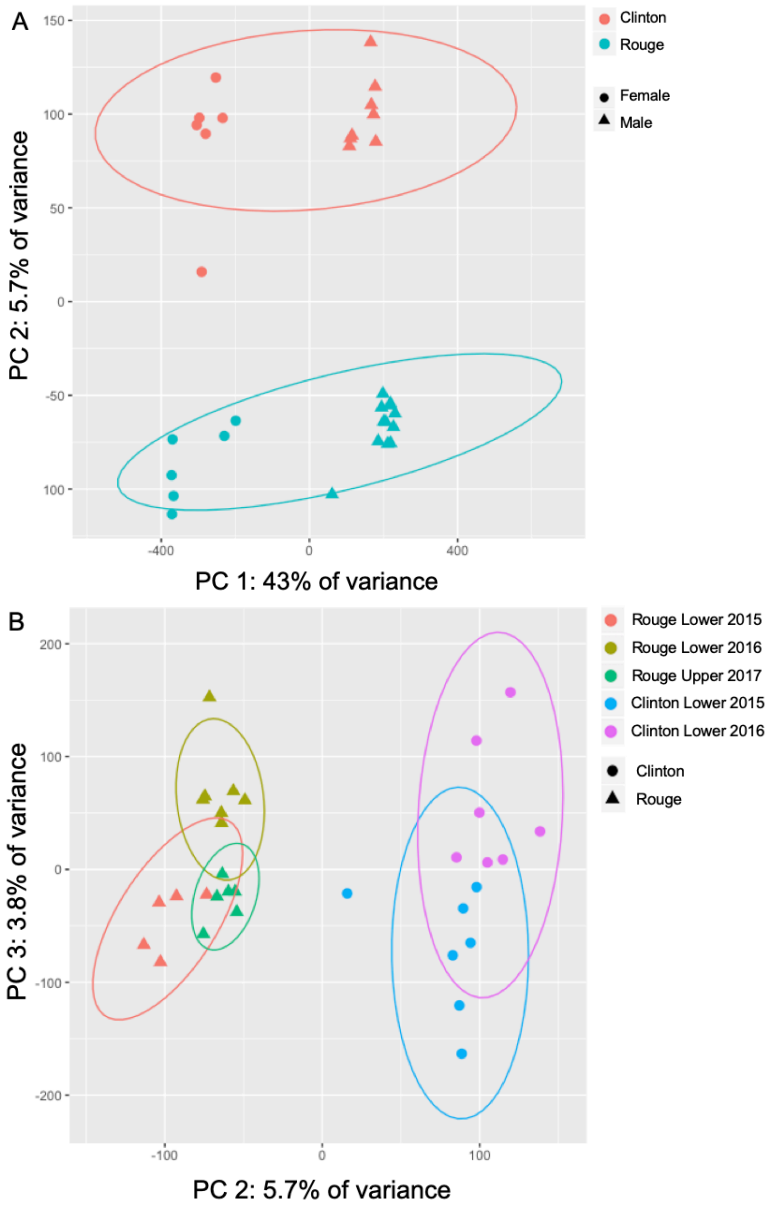


Figure 10. Results from principal component analysis of round goby. Legends identifying location and year are provided to the right of each plot and ellipses represent 95% confidence intervals. A) Plot of PC1 and PC2. Gene expression patterns cluster distinctly by sex, correlating with principal component 1, which accounts for 43.2% of the variance. Individuals cluster by drainage along PC2, explaining 5.7% of the variance. B) Plot of PC2 and PC3 (3.8% of variance) with clustering by year and site.

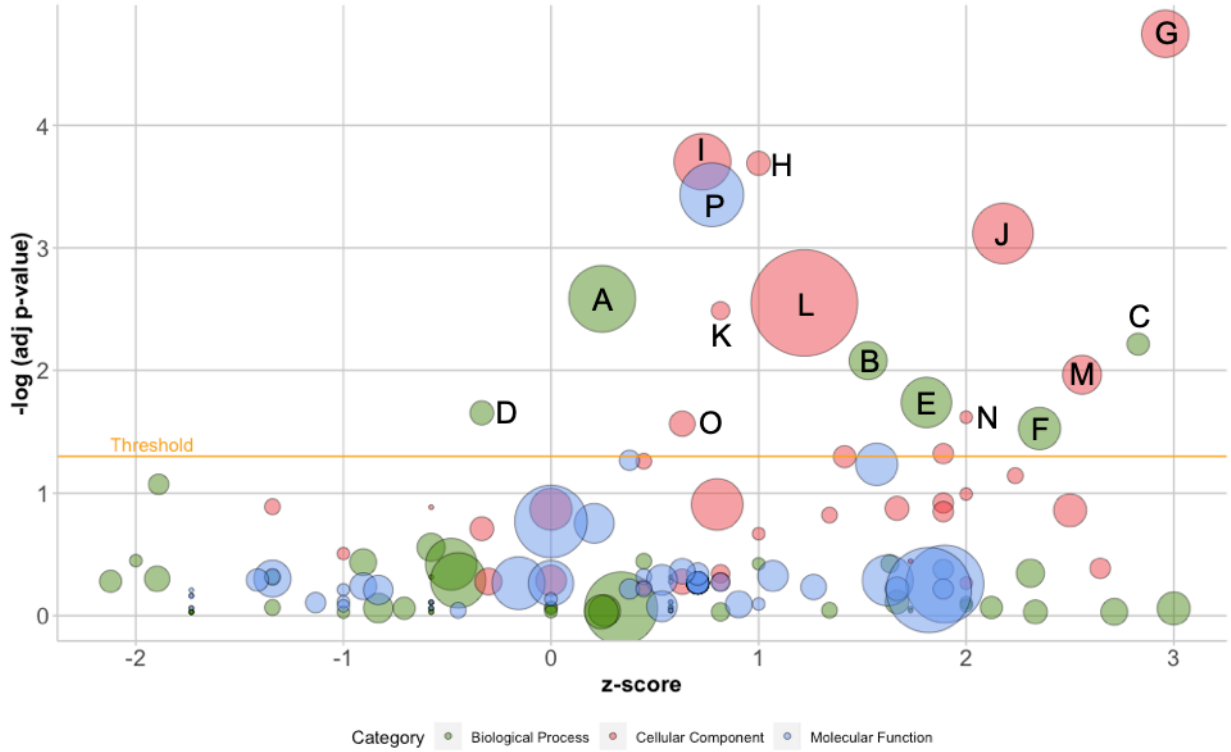


Figure 11. Gene ontologies for significantly differentially expressed genes in round goby between the Rouge and Clinton rivers. The x-axis indicates a pattern of up regulation for positive z-scores or down regulation for negative z-scores. Significantly enriched terms are labeled on the plot with letters which correspond to Table 8. The size of the circle corresponds the gene count for the term.

A

	Rouge River Lower 2015	Rouge River Lower 2016	Rouge River Upper 2017	Clinton River Lower 2015
Rouge River Lower 2016	480			
Rouge River Upper 2017	682	2303		
Clinton River Lower 2015	1187	1215	889	
Clinton River Lower 2016	1301	1294	1903	625

0-400
401-800
801-1200
1201-1700
1701+

B

	Rouge River Lower 2015	Rouge River Lower 2016	Rouge River Upper 2015	Rouge River Upper 2016	Rouge River Upper 2017	Clinton River Lower 2015	Clinton River Lower 2016	Clinton River Upper 2015
Rouge River Lower 2016	422							
Rouge River Upper 2015	328	527						
Rouge River Upper 2016	497	613	482					
Rouge River Upper 2017	257	358	306	473				
Clinton River Lower 2015	727	814	828	1023	480			
Clinton River Lower 2016	830	1043	902	981	568	658		
Clinton River Upper 2015	1112	1077	1093	1189	865	773	818	
Clinton River Upper 2016	1027	1093	1331	1150	537	857	905	937

Figure 12. Matrix of the number of differentially expressed genes in all pairwise comparisons for A) Johnny darter and B) round goby.

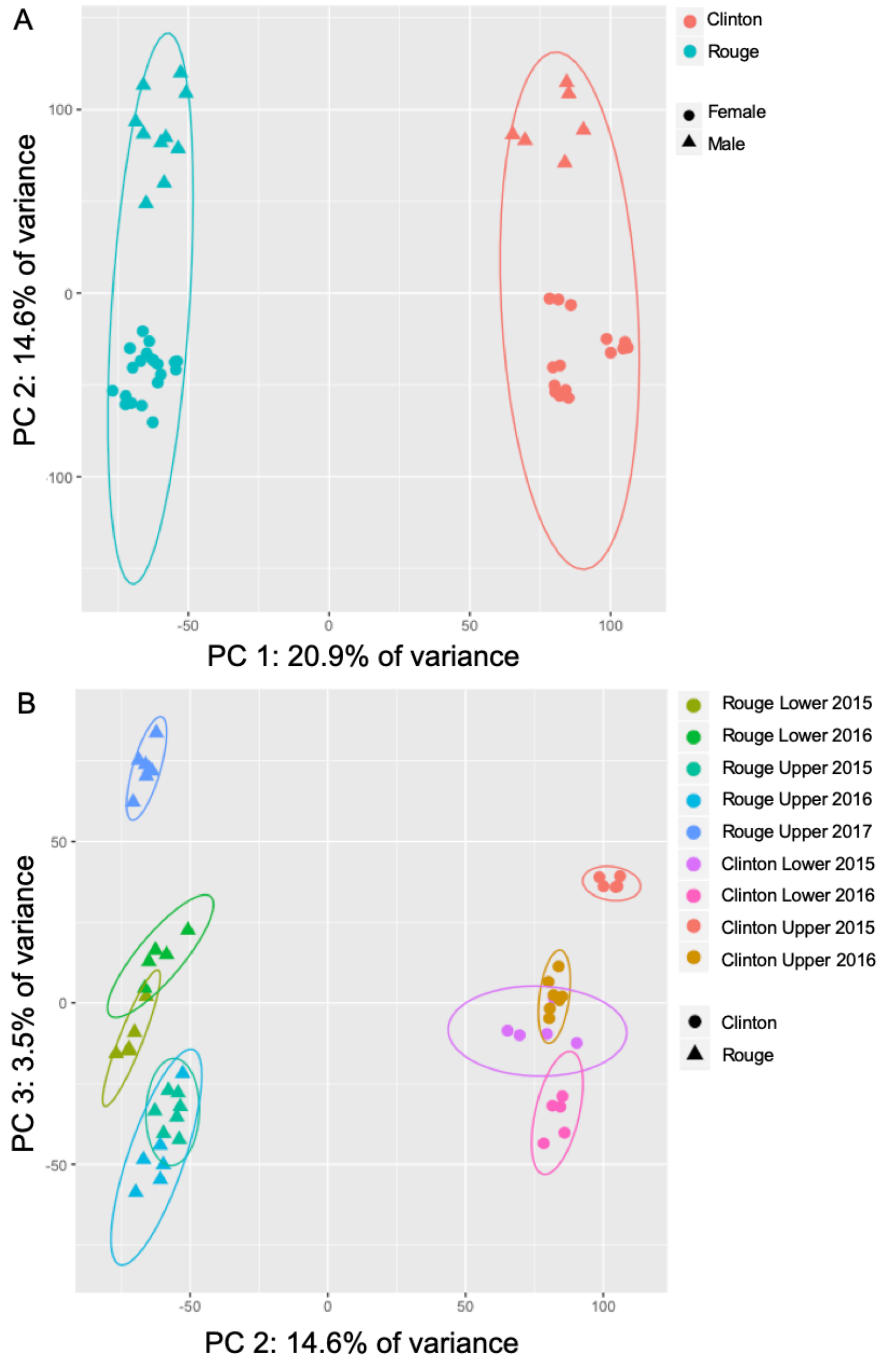


Figure 13. Results from principal component analysis of Johnny darter. Legends identifying location and year are provided to the right of each plot and ellipses represent 95% confidence intervals. A) Plot of PC1 and PC2. Gene expression patterns cluster distinctly by rivers (PC1, 20.9% of variance) and sex (PC2, 14.6% of the variance). B) Plot of principal components 2 and 3 shows clustering by site and sampling year with PC2 explaining 14.6% of the variance and PC3 explaining 3.5% of the variance.

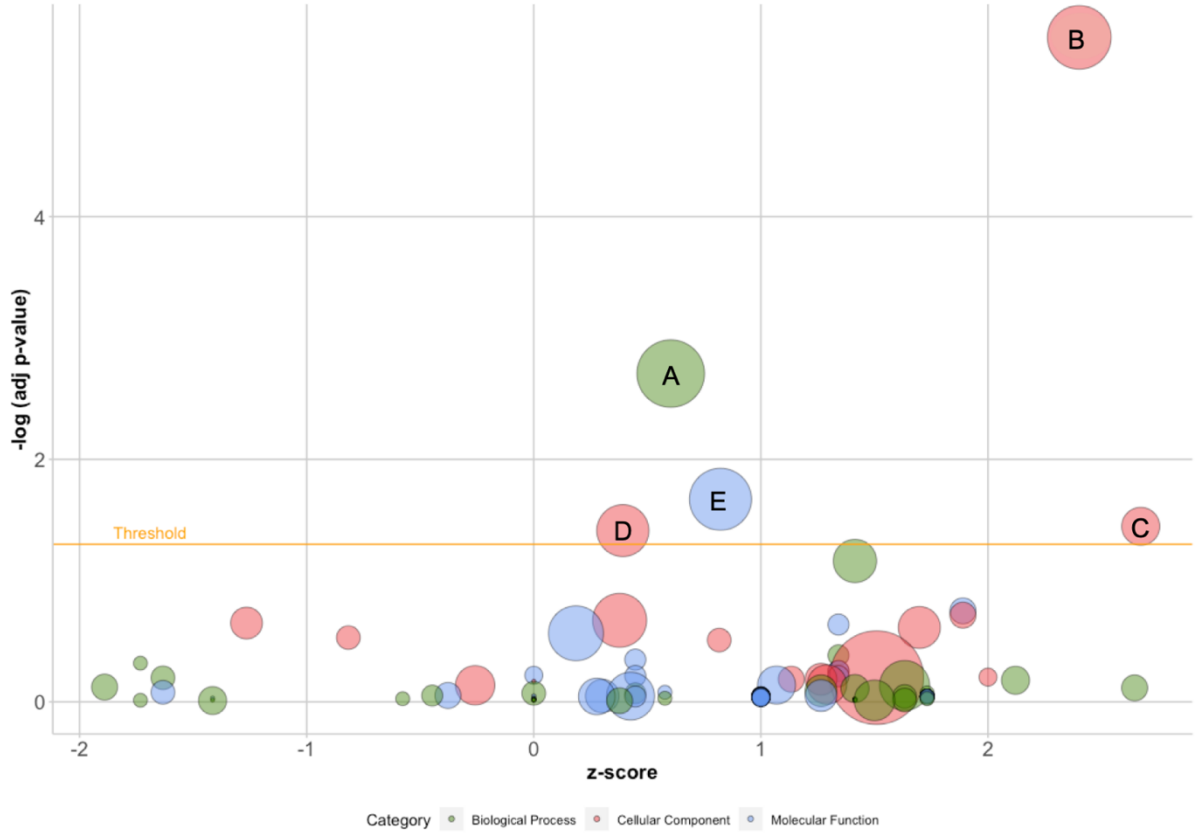


Figure 14. Gene ontologies for significantly differentially expressed genes in Johnny darter between the Rouge and Clinton rivers. The x-axis indicates a pattern of up regulation for positive z-scores or down regulation for negative z-scores. Significantly enriched terms are labeled on the plot with letters which correspond to Table 9. The size of the circle corresponds the gene count for the term.

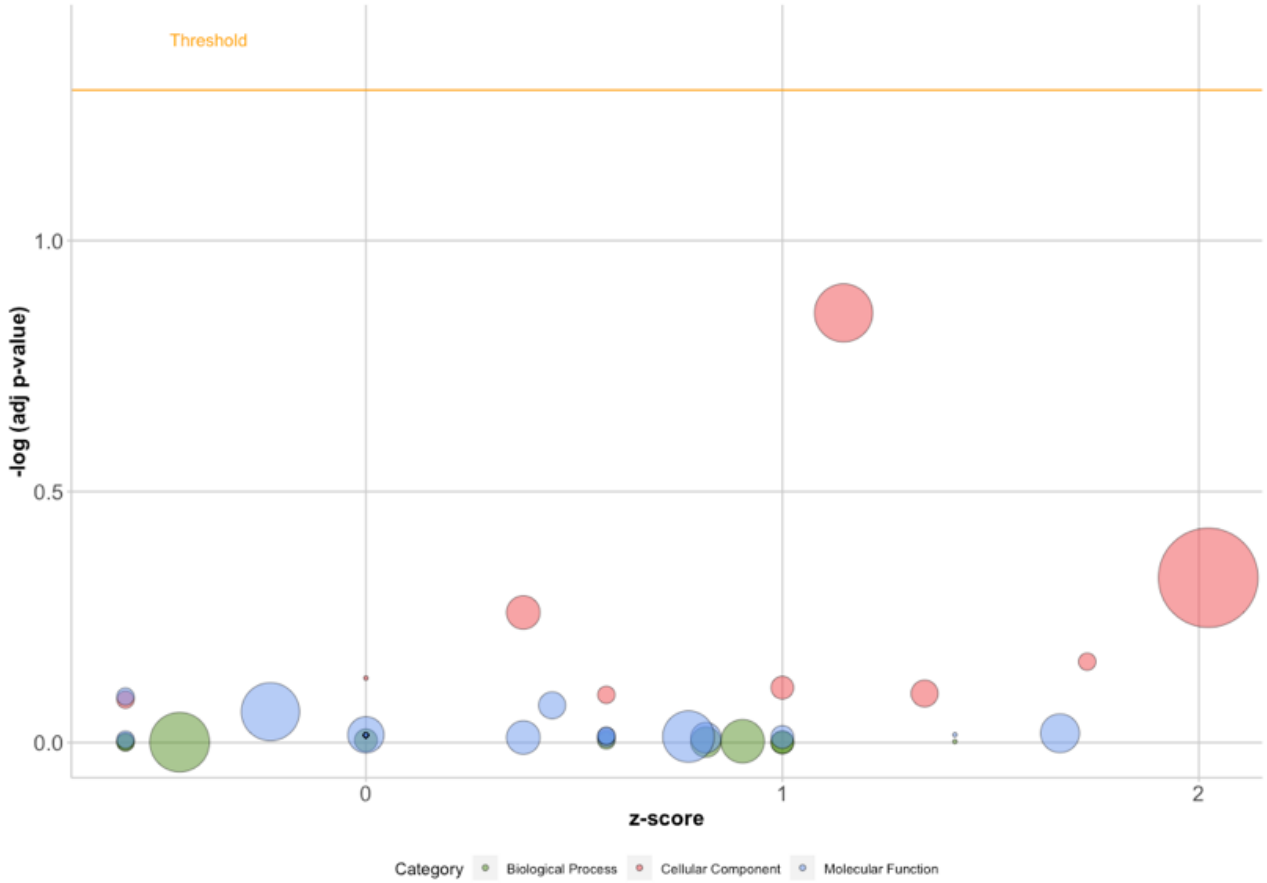


Figure 15. Gene ontology bubble plot for significantly differentially expressed genes in Johnny darter between goby present and goby absent sites. The x-axis indicates a pattern of up regulation for positive z-scores or down regulation for negative z-scores. There were no significantly enriched terms.

CHAPTER 3: GENE EXPRESSION RESPONSE OF ROUND GOBY AND JOHNNY DARTER TO ENVIRONMENTAL VARIABLES

INTRODUCTION

Tolerance to environmental stressors varies among organisms, and anthropogenic modification, contaminants, and other stressors may contribute to success of invasive species (Crooks et al. 2011, Airoldi 2015, Krabbenhoft 2019). Successful invaders are often species that are more tolerant of poor environmental quality, and differential response to environmental stressors may affect an organism's ability to persist when challenged by a more robust invader (Karatayev et al. 2009, Lenz et al. 2011). Further, urban systems, which are frequent sites of early invasion, are often characterized by high prevalence of environmental stressors, thus the localities where invasive species are introduced are also the localities most susceptible to invasion (Walsh et al. 2005). An understanding of how native and invasive taxa compare in the way they respond to environmental stressors will inform management efforts to reduce the spread of invasive species.

The invasive round goby has had variable impacts on native Johnny darter populations (Riley et al. 2008, Kornis et al. 2013). Environmental factors likely play a role in how these species interact with each other and the eventual outcome of this interaction. In Chapter 2, genes expression patterns were consistent with greater plasticity in populations of round goby as compared to Johnny darter. Here I will examine gene expression associations in these species with variables of their biotic and abiotic environment.

Round goby and Johnny darter are both generally thought to be tolerant species. The effects of pollutants have been studied in both species though no direct comparisons are available. Marentette et al. (2013) examined the response of gobies to pollutants from contaminated sites

versus non-contaminated sites, failing to detect differences in cortisol levels. Gobies from contaminated sites had smaller body size as compared to fish from reference sites, but prolonged exposure to pollutants did not suppress the stress response (Marentette et al. 2010). Exposure to wastewater effluent also had limited impact on round goby, increasing mortality at the highest exposures, but with no effect on mortality, behavior, or physiology at environmental levels of exposure (McCallum et al. 2017).

Fewer studies are available on the impact of stressors on Johnny darter. Krause et al. (2010) found pollution status of sampling site did not affect antipredator behavior in laboratory experiments; however, Reach and Berra (1987) found Johnny darters abundant at reference sites and absent at sites contaminated with sewage effluent. Lydy and Wissing (1988) found that exposure to copper lowered thermal tolerance in Johnny darter. This evidence suggests that both species are likely affected by pollutants, and I predict that both species will show changes in gene expression associated with chemical stressors; however, the available literature (Marentette et al. 2013, MarMcCallum et al. 2017) suggests that impacts at environmental levels of exposure may be less important in round goby.

In addition to the impact of chemical stressors, I will also examine the impact of biotic interactions on gene expression patterns. Due to overlap in spawning and foraging habitat, it is likely that direct competitive interactions occur between round goby and Johnny darter (Malone 2016). Direct interactions with other native species have been studied, finding round goby to be more aggressive than native species and to exclude native species from resources (Dubs and Corkum 1996, Janssen and Jude 2001). Coexistence with round goby appears to negatively affect growth in Johnny darter, suggesting these competitive interactions favor round goby (Kornis et al. 2014). Reallocation of resources to competition with exotic species presumably causes Johnny

darter to experience a greater degree of stress, therefore, the presence of round goby, especially at high densities, is expected to elicit a change in levels and patterns of gene expression in Johnny darter. Round goby also appears to be negatively impacted by intraspecific competition and are equally aggressive toward each other as they are towards other species, thus I also expect to find changes in levels and patterns of gene expression to be correlated with high conspecific density in round goby (Balshine et al. 2005, Kornis et al. 2014).

I will use whole transcriptome sequencing from samples of round goby and Johnny darter to examine the interactions discussed above. Whole transcriptome sequencing has the advantage of allowing for characterization of overall gene expression patterns to better interpret physiological changes associated with environmental stressors. These questions will be addressed by comparing changes in patterns of gene expression relative to various abiotic and biotic variables measured by Krabbenhoft (2019), using Weighted Gene Co-Expression Network Analysis (WGCNA). This method identifies groups of genes (modules) that are expressed in correlation with each other, and variation in levels of expression of these modules can be correlated with abiotic and biotic variation for round goby and Johnny darter. In Chapter 2, I found that round goby shows a more plastic gene expression pattern across localities than Johnny darter. Given this information, I expect the results of this analysis to show similar patterns, with round goby predicted to show more pronounced correlations to variables than Johnny darter.

First, the entire transcriptome from livers of round goby and Johnny darter will be analyzed individually with WGCNA, identifying groups of co-expressed genes and associated environmental factors. Second, to allow for direct comparison between Johnny darter and round goby, a subset of genes that are only present in the transcriptomes of both species, representing the conserved, ancestral portion of the genome, will be analyzed for co-expression. This method will

also provide a means to directly compare the two species that does not require annotation with a model organism making it possible to use genes without annotations as a means of comparison between species. Given that these are expected to be highly conserved genes, I predict the networks constructed to show high preservation.

METHODS

Sampling and Collections

The focus of this study was two drainages in southeastern Michigan, the Rouge and Clinton rivers. These rivers were selected because of 1) the distribution and abundance of the target species in these rivers, and 2) their drainage of urban and rural areas where a variety of stressors are present.

Methods for fish collection were as described in Chapter 2 and individuals included are detailed in Table S1. Fishes were collected from two sites on each of the Rouge and Clinton rivers with sites selected such that Johnny darter could be sampled in the presence and absence of round goby. This study was done in conjunction with that of Krabbenhoft (2019), who examined several biotic and abiotic variables in a number of Michigan river drainages, including these localities and sample periods. She characterized a suite of water quality data (e.g., temperature, pH, dissolved oxygen, conductivity [a proxy for salinity], copper concentration, fish community diversity, and density of round goby). Detailed methods for collection of water quality metrics and fish community sampling are available in Krabbenhoft (2019).

Data Analysis

Methods for construction of the gene count matrix are as described in Chapter 2. The counts were normalized with the R function *varianceStabilizingTransformation* in DESeq2 (version 1.24.0, Love et al. 2014). Some sampling groups are confounded by sex bias, so the *ComBat*

function in the SVA package for R (version 3.32.1, Leek et al. 2012) was used to remove the effect of sex from the expression data. Additionally, for darters, samples were sequenced in two separate Illumina lanes, and *ComBat* was used to remove potential batch effects. The resulting matrix was analyzed using the R package WGCNA (version 1.68, Langfelder et al. 2008). Samples were first clustered by Euclidian distance with the *hclust* function in the flashcluster package (version1.1.25, Müllner 2013) to check for and remove any outliers.

Using the *pickSoftThreshold* function, a soft thresholding power (beta) was chosen to fit a scale-free network, as recommended by Langfelder and Horvath (2014). The use of a soft threshold favors stronger correlations in network construction. Network construction is used to define co-expression modules and the pipeline for step-by-step signed network construction and module detection described in Langfelder and Horvath (2014) was used. A signed network separates positively and negatively correlated genes into different modules; an unsigned network can assign positively and negatively correlated genes in the same module. As the goal of this analysis is to infer biological processes from the correlation network, a signed network is more appropriate.

Expression patterns were analyzed to define gene "modules" (groups of genes that exhibited correlated changes in expression) using Pearson correlation. Minimum module size was set to 30 genes as recommended by the program authors and routinely used by researchers. Modules were analyzed for correlation with water quality and fish community metrics. Standard length of individuals was also included as a surrogate variable for life stage (which is generally related to size) as this feature is likely an important factor in gene expression, especially relative to reproduction. In WGCNA, variables tested are generally referred to as 'traits' and will be referred to as such going forward. Correlations were corrected for multiple testing with the Holm procedure (Holm 1979). Gene lists were input to DAVID 6.8 (Huang et al. 2009) for gene ontology

annotation and enrichment analysis. Gene ontology analysis is a process that classifies genes by their function or products to universally recognized 'terms.' Each term has a unique descriptor and GO ID number. Enrichment analysis identifies terms that are statistically overrepresented in a list of genes, including those in GO groups. Descriptions of gene functions, unless otherwise noted, were obtained from OMIM (Online Mendelian Inheritance in Man, https://omim.org).

Reciprocal BLAST and module preservation

To identify the set of genes present in the transcriptomes of both species, a BLAST search (Altschul et al. 1990) was conducted for each species against the other, using an e-value of 0.0001 and word size of 3. Results were filtered to include only genes that were each other's best hit using the programs and methods described in Warren et al. (2014). The resulting matrix was analyzed in WGCNA with the same parameters outlined above. The module Preservation function in WGCNA was used with 200 permutations to test for conservation of modules between the two species. This analysis was run using round goby as reference and Johnny darter as comparison, and then the reciprocal. This program generates a Z score, which compares the module against the expectation for a random set of genes. A module $Z_{summary}$ score >2 was used as the threshold for module preservation between species, as recommended by the developers of WGCNA.

RESULTS AND DISCUSSION

ROUND GOBY

Network Construction

Results for sequencing output, reference assembly, mapping, and annotation are detailed in Chapter 2. After filtering and correcting for sex, the resulting read count matrix of 31,001 transcripts was input to WGCNA for analysis. Samples were first clustered by overall expression

patterns in a distance tree to check for outliers (Figure S3). Two individuals met the threshold for outliers (RG15_003 and RG15_009) and were removed from later analyses.

A soft threshold power beta of 15 for a signed network was chosen (Figure S4). Network construction and module detection resulted in 47 modules. Highly correlated modules ($r > 0.75$) were merged, reducing the number of modules to 31, ranging in size from 35 genes to 9673 genes. Modules were then analyzed using DAVID for enriched gene ontology terms. Modules were assigned a color by WGCNA and, for ease of discussion, modules were renamed by the most significantly enriched gene ontology term after Benjamini-Hochberg correction. For modules that had no enriched terms, the module was named for term with the lowest p-value. Modules are summarized in Table 10, indicating the most significantly enriched gene ontology terms.

Module-Trait Correlations

Modules were analyzed for correlations with traits, with p-values corrected for multiple testing, and are presented in Figure 16. Significant module-trait correlations were identified for temperature, standard length, pH, species diversity, and round goby density (Table 10). There were no significant correlations with dissolved oxygen or copper.

Temperature

Four modules were significantly correlated with temperature. Of these, three were negatively correlated. The negative correlation indicates that expression levels of genes in this module decrease with increasing temperature. The significant mitochondrion1 module ($r=-0.61$, $p=8.49 \times 10^{-3}$) had 17 enriched terms. Along with mitochondrion (GO:0005739), the module had enriched terms related to hydrolase activity (GO:0016787), oxidation reduction (GO:0055114 and GO:0016491), and metabolism (GO:0008152). The mitochondrion2 module was also negatively correlated with temperature ($r=-0.65$, $p=2.60 \times 10^{-3}$) and had 12 significantly enriched terms. These terms were similar to those in the mitochondrion1 module, with additional metabolic terms

enriched (lipid metabolic process GO:0006629, steroid metabolic process GO:0008202). The oxidoreductase activity2 module was also negatively correlated ($r=-0.58$, $p=1.88x10^{-2}$) (GO:0016491) and had five enriched terms, most of which were involved in oxidation-reduction (GO:0055114, GO:0016705). Mitochondria function in energy production and respiration, thus negative correlation of mitochondrion terms likely indicates a downregulation of these processes with increasing temperature. The downregulation of oxidation-reduction processes in the same modules likely follows the downregulation in respiration and energy production as these processes are major sources of reactive oxygen species (Lushchek et al. 2006). This result is somewhat surprising as respiration and metabolism would be expected to slow with decreasing temperature (Clark and Johnston 1999). One collection event for round gobies was during unseasonably warm temperatures and could account for this anomaly. Under thermal stress, fish have reduced efficiency to supply oxygen for respiration (Farrell 2002). Although the temperature experienced is not outside of the tolerance range for round goby, the unseasonal and abrupt change could have caused thermal stress (Cross et al. 2009). WGCNA calculates a gene significance value (GS) which is a measure of the correlation of an individual gene to each trait. For genes in these modules with strong negative GS values for temperature, removing these individuals appears to confirm this interpretation. For example, for the gene chitinase, acidic.6 (chia6), a metabolic gene in the mitochondrion1 module, correlation with temperature is -0.83. When the two individuals from the high temperature event are removed, the correlation is -0.37. Similarly, for a gene in the mitochondrion2 module, mitochondrion intermediate peptidase (mipep), the correlation changes from -0.67 to 0.04 when those individuals are removed.

The antigen processing and presentation module was positively correlated with temperature ($r=0.61$, $p=7.98x10^{-3}$) and had no significantly enriched terms. Although not

significantly enriched, the module was composed of many terms related to immune response (GO:0042613, GO:0097169, GO:0072557). Upregulation of immune response with increasing temperature is common in fish and thus positive correlation of this module fits the expected pattern (Boltaña et al. 2013, Mohammed et al. 2016).

Standard Length

The catalytic activity module (GO:0003824) was the only module correlated with standard length ($r=0.56$, $p=6.7x10^{-4}$). This module contained 7552 genes and is the second largest module. Gene ontology enrichment analysis found 42 significantly enriched terms. In addition to catalytic activity, other enriched terms were oxidoreductase activity (GO:0016491), and cytosol (GO:0005829). Given the range of standard length (33mm to 80mm) for round gobies in the sample, it is surprising that there were not associations with more modules. Enriched terms in the catalytic module do not have a clear connection to growth or life stage and could result from a correlation between standard length and conductivity ($r=-0.41$, $p=2.1x10^{-2}$) and/or pH ($r=-0.44$, $p=1.3x10^{-2}$), which were both negatively correlated with the catalytic module.

Several genes that had high GS for standard length were related to reproductive processes such as vitellogenin and apolipoprotein, suggesting a relationship between size and reproductive maturity. A correlation with standard length and fecundity in round gobies has been previously established (MacInnis and Corkum 2000). Spawning in females occurs at >40 mm, thus the range sampled here likely includes individuals that are not yet reproductively mature (L'avrinčíková and Kováč 2007). Additionally, round goby are noted to spawn throughout the season with a variable proportion of the population reproductive at any one time (MacInnish and Corkum 2000). If reproductive activity is driving the association with the module and only a portion of the

individuals are reproductively active, this may explain why only one module was associated with this factor.

Conductivity

Three modules were significantly correlated with conductivity. The nucleolus2 ($r=-0.59$ $p=1.33\times10^{-2}$) and catalytic activity ($r=-0.61$, $p=7.68\times 10^{-3}$) modules were negatively correlated with conductivity while the metabolic process module was positively correlated with conductivity ($r=-0.67$, $p=1.28\times10^{-3}$). Although these modules had large numbers of significantly enriched genes, the annotated genes provided limited insight into understanding response of round goby to salinity as genes known to be involved in osmoregulation in fish did not have annotations in the round goby transcriptome. However, one relevant gene was identified in the metabolic process module, the sodium channel protein type 4 subunit alpha B (scn4ab). This gene functions in sodium ion transport and had high GS for conductivity.

pH

Three modules had significant correlations with pH. The DNA replication and catalytic activity modules were negatively correlated with pH ($r=-0.59$, $p=4.4\times10^{-4}$ and $r=-0.56$, $p=2.99\times10^{-2}$, respectively) while the response to cAMP module was positively correlated with that variable ($r=0.65$, $p=6.5\times10^{-5}$). Cyclin adenosine 3’5’ monophosphate (cAMP) is a common cell signaling pathway for many biological processes, thus this term indicates a response to this signal. Also enriched was activation of MAPKKK (mitogen activated protein kinase) activity (GO:0000185). Zhao et al. (2015) studied gene expression response in tilapia to alkalinity challenges, and there was notable overlap with the terms and genes associated with pH in round goby found here with their results. They identified upregulation of the MAPK pathway in fish exposed to an alkaline environment. They also noted differential expression of eukaryotic translation initiation factor 4E-

binding protein 3 (eif4ebp3), which also functions in MAPK signaling, and a similar gene was strongly associated with pH for round goby (eif4ebp2). Measured pH at sampling sites was generally within the normal tolerance range for most fishes (7.25-8.92), though the highest pH may be approaching the upper limit for pH tolerance. Therefore, the genes and terms associated with pH in round goby are consistent with the expected physiological response to alkalinity.

Shannon Diversity

Four modules were significantly correlated with Shannon diversity and also represented the two strongest module-trait correlations, the antigen processing and presentation module ($r=-0.88$ $p=2.15\times10^{-9}$) and the response to cAMP module ($r=-0.57$ $p=2.16\times10^{-2}$). The MHC class II protein complex module ($r=0.56$ $p=3.01\times10^{-02}$) along with the oxidoreductaseactivity2 ($r=0.88$ $p=2.36\times10^{-9}$) were positively correlated with diversity. Most likely, these are not a direct association with diversity, but an association with factors that are correlated with differences in species diversity.

Round Goby Density

Three modules had significant correlations with goby density. The response to cAMP and antigen processing and presentation modules were negatively correlated with density ($r=0.61$, $p=6.43\times10^{-4}$ and $r=0.68$, $p=8.91\times10^{-3}$, respectively) while the oxidoreductase activity2 module was positively correlated ($r=-0.67$ $p=1.07\times10^{-3}$). These modules also had correlations with Shannon diversity, though with opposite signs. Pearson correlation was calculated for round goby density and Shannon diversity for sites where round goby was sampled, identifying a strong negative correlation between the two traits ($r=-0.71$, $p=8.1\times10^{-6}$). This relationship likely explains the association of the same modules with both factors.

JOHNNY DARTER

Network Construction

Results for sequencing output, reference assembly, mapping, and annotation are detailed in Chapter 2. After filtering and batch corrections, the resulting read count matrix of 33,559 transcripts was input to WGCNA for analysis. Samples are first clustered by overall expression patterns in a distance tree to check for outliers (Figure S5). One outlier was identified (JD16_20) and removed.

A soft threshold power beta of 12 for a signed network was chosen (Figure S6). Network construction and module detection identified 70 modules. Modules that were highly correlated ($r>0.75$) were merged, reducing the number of modules to 52, ranging in size from 44 genes to 4488 genes. Each module was analyzed for gene ontology enrichment using DAVID and named for the most significant ontology term as with round goby. Composition of modules and results of gene enrichment are detailed in Table 11.

Module-Trait Correlations

As with round goby, modules were analyzed for correlations with traits and are presented in Figure 17. Significant module-trait correlations (Table 11) were identified for temperature, standard length, conductivity, dissolved oxygen, copper, Shannon diversity, and round goby density. Notably, there were no significant correlations with pH.

Temperature

Two modules had significant correlations with temperature. The strongest correlation was with the oxidoreductase activity6 module ($r=0.49$, $p=1.23\text{x}10^{-2}$). In addition to oxidoreductase activity (GO:0016491), several other processes in this module were also enriched (mitochondrion, GO:0005739; catalytic activity, GO:0003824; oxidation-reduction processes, GO:0055114; metabolic processes, GO:0008152; and hydrolase activity, GO:001687). These processes indicate

an upregulation of respiration and metabolism with increasing temperature and is the expected pattern in fishes (Clark and Johnston 1999). The pre-autophagosomal structure module was negatively correlated with temperature ($r=-0.46$, $p=3.06\times10^{-3}$). Significantly enriched gene ontology terms in this module were immune response related, including pre-autophagosomal structure (GO:0000407) and autophagosome (GO:0005776). This pattern suggests a downregulation of respiration (mitochondrion, metabolic processes) with increasing temperature, which is the opposite pattern expected (Boltaña et al. 2013, Mohammed et al. 2016).

Standard Length

Four modules had significant correlations with standard length. The ribosome module was positivity correlated ($r=0.56$, $p=7.19\times10^{-4}$) with most terms related to ribosome function including ribosome (GO:0005840), intracellular ribonucleoprotein complex (GO:0030529), and structural constituent of ribosome (GO:0003735). The catalytic activity module ($r=-0.54$, $p=1.72\times10^{-3}$), oxidoreductase activity5 module ($r=-0.50$, $p=1.02\times10^{-2}$), and oxidoreductase activity4 module ($r=-0.47$, $p=2.57\times10^{-2}$) were all negatively correlated with standard length. Many enriched terms in these three modules were related to oxidation reduction processes (GO:0016491, GO:0055114) and catalytic activity (GO:0045261, GO:0003824). The oxidoredutase4 module had no significantly enriched ontology terms. The gene with the strongest negative GS for standard length was insulin-like growth factor 2 binding protein 2a (IGF2bp2a). Insulin-like growth factor is a common growth regulation gene and the associated binding proteins can repress growth, often after development, thus this gene is often downregulated in the larger individuals (Wu et al. 2019). Additionally, the same reproductive genes identified in round goby (vitellogenin, apolipoprotein) had high positive GS for standard length, indicating that some darters were likely not sexually mature.

Conductivity

Only the ribosome module had a significant correlation with conductivity (r=-0.48, $p=1.67x10^{-2}$). Twenty gene ontology terms were enriched in this module and the most significant were ribosome (GO:005840), structural constituent of ribosome (GO:0003735), and intracellular ribonucleoprotein complex (GO:0030529). As with round goby, genes that are well known to function in osmoregulation in fish did not have strong blast matches for Johnny darter; therefore, this result for both species may reflect the lack of genomic resources. Enriched ontology terms and genes with high GS could not be directly associated with osmoregulation based on available literature.

Dissolved Oxygen

The antigen processing and presentation module was negatively correlated with dissolved oxygen (r=0.91, $p=1.267x10^{-2}$). Enriched terms for the module were antigen processing and presentation (GO:0019882) and immune response (GO:0006955), indicating decreases in immune processes with increases in dissolved oxygen. This association is contrary to expectations, as exposure to hypoxia in fish typically reduces immune response (Welker et al. 2007, Kvamme et al. 2013).

Copper

Copper concentration was significantly correlated with the ribosome2 module and represented the strongest correlation for Johnny darter (r=0.91, $p=3.01x10^{-18}$). In fish, copper contamination in water induces oxidative stress and changes in blood chemistry (Singh et al. 2008, Eykmans et al. 2011, Kaviani et al. 2019). Several genes with high GS for copper concentration in Johnny darter were related to blood plasma components including fibrinogen beta and fibrinogen gamma, and prostaglandin synthase, consistent with the changes in blood chemistry that have been associated with copper exposure. Enrichment analysis of the ribosome2 module primarily

identified genes related to ribosome function; however, terms related to oxidation-reduction were also enriched. Genes with oxidative-reduction function, including members of the cytochrome p450 family, cytochrome c oxidase, and alcohol dehydrogenase, had high GS for copper. Glutathione peroxidase was also strongly associated with copper and functions in clearing reactive oxygen species. Although it is not possible to assess the Johnny darter's tolerance to copper from this information, there appears to be evidence of an impact (changes in blood chemistry) and a response (oxidation-reduction).

Shannon Diversity

Two modules had significant correlations with the Shannon diversity index. The triglyceride catabolic process module was positively correlated with Shannon diversity ($r=0.45$, $p=8.02\times10^{-4}$), and triglyceride catabolic process (GO:0019433) was the only term enriched in that module. The mitochondrion5 module was negatively correlated ($r=-0.68$, $p=2.8\times10^{-6}$). No terms were significantly enriched in the mitochondrian5 module.

Goby Density

Levels of goby density were significantly correlated with the ribosome2 module ($r=0.76$, $p=5.05\times10^{-4}$), and this module also had strong positive correlation with copper concentration. Goby density was negatively correlated with copper ($r=0.77$, $p=5.4\times10^{-}4$); therefore, this correlation is likely driving the correlation of both the traits with the ribosome module. As genes and ontologies in the module are consistent with what is known about how fish respond to copper, it is more likely that copper is truly associated with this module, and therefore, this is also driving the impact of goby density on Johnny darter.

Comparing round goby and Johnny darter expression patterns

Although more modules were identified in Johnny darter than in round goby, more module-trait correlations were observed in round goby. These also tended to be stronger than correlations found in Johnny darter (average correlation of 0.64 for round goby versus 0.51 for Johnny darter). Of the traits measured, temperature, standard length, conductivity, dissolved oxygen, pH, and copper had different impacts for Johnny darter and round goby. Additionally, both species showed significant correlations with species diversity and round goby density, though these may be confounded by other factors (such as copper concentration).

These results should be interpreted with caution and are not meant to imply causation. Streams systems are dynamic and many factors are interrelated and likely influenced by factors not measured or considered. Nevertheless, studies such as this provide baseline information for future studies of the impacts of significant environmental variables and their role in adaptation of native and invasive species to their environments.

Temperature

Temperature has been previously identified as an important factor in the ability of round goby to rapidly expand its invasive range (Wellband and Heath 2017, Wellband et al. 2018). Here, Johnny darter and round goby exhibited contrasting correlations with temperature. Johnny darter and round goby differ in their thermal tolerance ranges. CTMax (critical thermal maximum) for Johnny darter is 30.9ºC as compared to 33.4ºC for round goby (Ingersoll and Claussen 1984, Cross and Rawding 2009). For this study, measured temperatures ranged from 11.8 ºC to 19.2 ºC for round goby and 11.8ºC to 16.6ºC for Johnny darter. Although temperatures observed during specimen collection did not vary as much as in studies that specifically test for thermal stress, a correlation with variation in temperature was observed. There were similarities in enriched gene

ontology terms for the modules that had the strongest correlation for each species with temperature. Both had enrichment of several process related to oxidation-reduction and mitochondria; however, these were upregulated with increasing temperature in Johnny darter and downregulated in round goby. It is possible the unexpected associations in round goby may be the result of a slowing of these process with unseasonably high temperatures experienced for one collection event for round gobies (Johnny darter were not collected on the same occasion).

The enriched terms in the negatively correlated module for Johnny darter were related to immune response (pre-autophagosomal structure, autophagosome). The positively correlated module for round goby had no significantly enriched terms, but did have several terms related to immune function (IPAF inflammasome complex, MHC class II protein complex). Suboptimal temperatures are known to impact immune response in fish with warmer water temperatures tending to activate immune response (Boltaña et al. 2013, Mohammed et al. 2016), thus correlations in round goby were consistent with this expectation. The contrasting patterns observed may reflect differences in the way that Johnny darter and round goby respond to temperature changes. Fish species commonly vary in the rate of their response to changes in temperature (Haverinen and Vornanen 2009). Fishes were collected in April and May when temperatures are highly variable. Both Rouge River sites are also downstream of a wastewater treatment outflow that can cause daily temperature fluctuations, depending on discharge (B. Mueller, pers. comm.). It is possible that temperature fluctuations could have occurred prior to our sampling and that the pattern observed reflects an immune response in Johnny darter that has not yet adjusted to a recent change in conditions.

Standard Length

Both round goby and Johnny darter had significant correlations with standard length. With the exception of one gene identified in Johnny darter, insulin like growth factor binding protein 2a (IGF2bp2a), most genes and enriched processes were not clearly connected with growth. Rather, for both species, highly correlated genes were related to reproductive processes. This result is not surprising as larger individuals were more reproductively active than smaller ones that may not have been sexually mature; however, there were four significantly associated modules for Johnny darter and only one for round goby. The catalytic activity module was positively correlated with standard length for round goby and was also a very large module at 7552 genes, much larger than the modules associated with standard length for Johnny darter. In addition to variation in reproductive timing in round goby, this difference may explain why there were more modules associated with standard length for Johnny darter despite the greater range of standard length for round goby than Johnny darter (33-80mm and 35-64mm, respectively). For round goby, one large module may encompass these processes while for Johnny darter several smaller modules are involved.

Conductivity

As round goby is native to brackish waters, it is expected to be more tolerant to changes in salinity (Karsiotis et al. 2012). For round goby, there were correlations for conductivity with three modules while Johnny darter only had one module correlated with conductivity. Enrichment analysis provided limited insight for both species as many genes did not have annotations, and genes typically associated with osmoregulation were not evident in modules identified for either Johnny darter or round goby. One noteworthy exception was sodium channel protein type 4 subunit alpha B (scn4ab). This gene functions in sodium ion transport and was correlated with conductivity

for round goby. Similar correlations were not observed for Johnny darter. Although annotations were limited for genes associated with this trait, the difference in the number of modules and stronger correlations for round goby as compared to Johnny darter may demonstrate a difference in the way these species respond to variation in salinity.

Dissolved oxygen

Observationally, round goby and Johnny darter differ in their tolerance for dissolved oxygen, with round goby appearing to tolerate hypoxia better than Johnny darter (A. Wicks, pers. comm.). Round goby had no modules that were significantly associated with dissolved oxygen. Johnny darter had one module that was negatively correlated with that trait, and this module contained several terms related to immune function. This result is consistent with downregulation in immune response with increases in dissolved oxygen, a pattern opposite to the one expected (Welker et al. 2007, Kvamme et al. 2013). As with temperature, this may suggest a delay in response to changing conditions. A typical adaptive response to low dissolved oxygen would be a reduction in respiration and metabolic processes (Richards 2011). Since these changes were not noted in either species, it may be that dissolved oxygen conditions experienced during collection did not reach levels that would elicit this response.

pH

For round goby three modules were correlated with pH while there were no correlations with pH for Johnny darter. The genes that were strongly associated with pH for round goby were consistent with other studies that examined the gene expression response of fish to alkalinity. Johnny darter was also sampled at the same locations with higher pH values experience by round goby; however, the same gene expression pattern was not observed. Jakubčinová et al. (2018) characterized habitats occupied by round goby, and these tended to have slightly alkaline water.

Sea water has a higher pH than freshwater and as a native to the brackish Black Sea, round goby may be able to more readily respond to changes in pH than Johnny darter.

Copper

Johnny darter had one module significantly correlated with copper concentration, consistent with expected changes in blood chemistry and oxidative processes that are known to occur with exposure to copper in fishes. There were no such changes noted in round goby, possibly indicating that this concentration of copper is not inducing the same effect in round goby as in Johnny darter. Krabbenhoft (2019) also noted the importance of copper concentration and found it to be predictive of round goby invasion success. Although it is not likely that copper is uniquely causal to successful round goby invasion, it is likely that these sites with copper contamination are sites with a number of other stressors present and may offer advantages for round goby if they are more tolerant to such stressors.

Shannon species diversity and round goby density

For both species, Shannon diversity and round goby density represented some of the strongest correlations; however, interpretation of these results was challenging to assess due to correlations with other traits. For Johnny darter, the ribosome2 module correlated with levels of round goby density and was also correlated with copper concentration. Round goby had modules strongly correlated with Shannon diversity and levels of round goby density therefore the importance of these two variables could not be separated due to their correlation. Johnny darter had a module significantly correlated to species diversity that was not conflated by round goby density.

Ecosystems exhibit a web of interactions of biotic and abiotic factors, making it very difficult to easily interpret biological implications from gene expression patterns. Species diversity

can influence a number of ecological and physiochemical parameters (Duffy et al. 2007, Cardinale 2011); therefore, observed gene expression patterns could be a response to parameters associated with diversity, competition within or between species, or facultative interactions with other species (Cardinale et al. 2006, Azour et al. 2015, Wright et al. 2017). Despite the limitations for interpretation of specific physiological impacts, it does demonstrate significant associations in both round goby and Johnny darter to species diversity.

Reciprocal BLAST and module preservation

Reference transcriptomes of round goby and Johnny darter were BLASTed against each other, adjusted for batch effects, and filtered in the same manner as the full gene expression matrix. This approach was expected to identify genes that could be identified as "homologous," yielding 5421 shared genes. For Johnny darter, after removing low read count genes, the gene list of 4450 genes was analyzed with WGCNA. As with the full data set, JD126_20 was flagged as an outlier and removed. The module network was constructed with a soft threshold power beta of 5 (Figure S7), yielding a network that consisted of thirteen modules ranging in size from 64 to 1071 genes. None of these met the criteria for merging. This network was first used as the reference network to analyze module preservation with Round goby. Results are presented in Figure 18A. A Zsummary score above 2 is considered the threshold for preservation. No module met this threshold.

The same approach was used with round goby as the reference network. After batch correction and filtering for low read count genes, the resulting list of 4314 genes was analyzed with WGCNA. With a smaller list of genes, two individuals were identified as outliers and removed (RG15_003 and RG15_009). The module network was constructed with a soft threshold power beta of 12 (Figure S8), yielding a network that included 14 modules. These were ultimately

merged to 11 modules ranging in size from 63 to 1904 genes. This network was used as the reference and analyzed for module preservation with Johnny darter (Figure 18B). As with the comparison with Johnny darter as a reference, no modules met the threshold for preservation.

I predicted that the networks defined from these genes would be preserved between round goby and Johnny darter. Networks were constructed for both Johnny darter and for round goby and each was used as a reference and in both cases, no module met the threshold for preservation. Few other examples of application of this tool across species could be identified. Cheviron et al. (2017) found high perseveration for the majority of modules defined for two species of songbirds in response to seasonal changes. Filteau et al. (2019) tested module preservation between ecotypes of whitefish and sockeye salmon and found preservation for approximately half of their modules and those that were preserved across species were those with strong association with phenotypes.

Contrary to predictions, Johnny darter and round goby differ greatly in expression patterns even for highly conserved genes. The lack of module preservation suggests that Johnny darter and round goby are "using" this common set of genes in different ways. Alterations in expression are likely an important means of adaptation and evolution (Carroll 2008). The results of Chapter 2 indicated differences in gene expression patterns within and between populations of Johnny darter and round goby. Gene expression differences are a significant source of phenotypic variation both within species and between species (Oleksiak 2002, Wittkopp et al. 2004, Whitehead and Crawford 2006). Alterations in expression may be an efficient mechanism of evolution, allowing for adaption to new environments without novel genetic material.

Conclusions

Weighted gene co-expression network analysis was applied to two species of fish sampled in a variety of environmental conditions, identifying groups of genes correlated with several environmental factors. This approach also facilitated the comparison of gene expression patterns of these two species associated with variation in their environment. I observed that the module-trait correlations for Johnny darter and Round goby are markedly different. Although this study did not test how these changes in gene expression relate to the organism's physical response to environmental stressors, the marked difference in gene expression patterns in correlation with environmental variation in the invasive round goby and native Johnny darter suggest that the species differ in their tolerance of stressors.

Round goby showed no correlations with copper concentration while Johnny darter showed correlations consistent with expectations based on studies of fish exposed to copper. Negative effects of copper on Johnny darter have previously been established as Lydy and Wissing (1988) showed that copper exposure reduced tolerance to thermal stress in Johnny darter. For studies that measured response of round goby to pollutants, round goby did not show a response at environmental levels of exposure (Marentette et al. 2013, McCallum et al. 2017). Krabbenhoft (2019) identified copper contamination as predictive of round goby invasion. That Johnny darter and round goby differ in these correlations at the gene expression level and this factor is associated with round goby invasion may suggest that a differential response to stressors is important to invasion success. A more robust tolerance and response to stressors as compared to native species would likely give round goby an advantage in expanding its range in degraded habitats.

Round goby and Johnny darter also differed in their correlations with temperature, pH, and conductivity. The round goby population that invaded the Great Lakes is native to brackish waters

while Johnny darter is an exclusively freshwater species. As round goby is adapted to variable salinity and pH of brackish waters, this may provide an advantage in impacted areas (Karsiotis et al. 2012). The native range of round goby is characterized by short term variation in climate, and this species has broad thermal tolerance (Neuman 1991, Cross and Rawding 2009). Impacted streams like those sampled for this study receive inputs from agricultural runoff, stormwater runoff, and wastewater effluents that can results in fluctuations in water chemistry and temperature (Walsh et al. 2005, Cokeril et al. 2017). This is similar to the native habitat of round goby, characterized by high variability, and may have favored plasticity that is well suited to degraded streams that are commonly early sites of invasion. These features appear to contribute to the success of round goby as an invader of the Great Lakes and their tributaries.

Table 10. Composition of modules identified in construction of gene co-expression network for round goby.

Module Name	Number of Genes	Top Significant GO Terms	Negatively Correlated Traits	Positively Correlated Traits
poly(A) RNA binding	82	N/A		
endoplasmic reticulum	63	GO:0005783~endoplasmic reticulum		
nucleolus	501	GO:0005730~nucleolus, GO:0003723~RNA binding, GO:0008168~methyltransferase activity		
ATP-dependent helicase activity	282	GO:0008026~ATP-dependent helicase activity		
mitochondrion1	3448	GO:0005739~mitochondrion, GO:0005737~cytoplasm, GO:0016787~hydrolase activity	Temperature	
mitochondrion2	1185	GO:0005739~mitochondrion, GO:0055114~oxidation-reduction process, GO:0016491~oxidoreductase activity	Temperature	
nucleolus2	1491	GO:0005730~nucleolus, GO:0003723~RNA binding, GO:0005739~mitochondrion	Conductivity	
catalytic activity	7552	GO:0003824~catalytic activity, GO:0016491~oxidoreductase activity, GO:0005829~cytosol	Conductivity, pH	Standard Length
oxidoreductase activity	973	GO:0016491~oxidoreductase activity, GO:0055114~oxidation-reduction process, GO:0005739~mitochondrion		
mitochondrion3	498	GO:0005739~mitochondrion, GO:0005747~mitochondrial respiratory chain complex I, GO:0003735~structural constituent of ribosome		
ribosome	283	GO:0005840~ribosome, GO:0003735~structural constituent of ribosome, GO:0006412~translation		
mitochondrial inner membrane	384	GO:0005743~mitochondrial inner membrane, GO:0005739~mitochondrion		
cytoplasm	103	GO:0005737~cytoplasm, GO:0016491~oxidoreductase activity		
endopeptidase inhibitor activity	454	GO:0004866~endopeptidase inhibitor activity, GO:0005615~extracellular space, GO:0005576~extracellular region		
regulation of translational initiation	130	GO:0006446~regulation of translational initiation, GO:0016282~eukaryotic 43S preinitiation complex, GO:0033290~eukaryotic 48S preinitiation complex		

Table 10 Continued.

Module Name	Number of Genes	Top Significant GO Terms	Negatively Correlated Traits	Positively Correlated Traits
peptidase activity	333	GO:0008233~peptidase activity, GO:0008236~serine-type peptidase activity, GO:0016787~hydrolase activity		
DNA replication	50	GO:0006260~DNA replication, GO:0006270~DNA replication initiation, GO:0005634~nucleus	pH	
cell cycle	216	GO:0007049~cell cycle, GO:0007059~chromosome segregation, GO:0007067~mitotic nuclear division		
lipid transporter activity	207	GO:0005319~lipid transporter activity, GO:0006869~lipid transport, GO:0071391~cellular response to estrogen stimulus		
catalytic activity1	100	GO:0003824~catalytic activity, GO:0008152~metabolic process, GO:0055114~oxidation-reduction process		
response to cAMP	270	GO:0051591~response to cAMP, GO:0051726~regulation of cell cycle, GO:0045597~positive regulation of cell differentiation	Diversity	pH, Round Goby Density
antigen processing and presentation	462	N/A	Diversity	Temperature, Round Goby Density
respiratory chain complex IV	63	GO:0045277~respiratory chain complex IV, GO:0006869~lipid transport, GO:0042627~chylomicron		
hemoglobin complex	35	GO:0005833~hemoglobin complex, GO:0019825~oxygen binding, GO:0005344~oxygen transporter activity		
translation	70	N/A		
oxidoreductase activity2	551	GO:0016491~oxidoreductase activity, GO:0055114~oxidation-reduction process, GO:0016705~oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	Temperature, Round Goby Density	Diversity
MHC class II protein complex	63	N/A		Diversity
thiosulfate sulfurtransferase activity	36	N/A		
glutathione transferase activity	100	N/A		
mitochondrion4	1234	GO:0005739~mitochondrion, GO:0030529~intracellular ribonucleoprotein complex, GO:0016746~transferase activity, transferring acyl groups		
metabolic process	9673	GO:0008152~metabolic process, GO:0003824~catalytic activity, GO:0005739~mitochondrion		Conductivity

Table 11. Composition of modules identified in construction of gene co-expression network for Johnny darter.

Module Name	Number of Genes	Top Significant GO Terms	Negatively Correlated Traits	Positively Correlated Traits
triglyceride catabolic process	136	GO:0019433~triglyceride catabolic process		Diversity
cilium morphogenesis	103	N/A		
Hydrolase	737	N/A		
autophagosome assembly	276	GO:0000045~autophagosome assembly		
aromatic amino acid family metabolic process	80	GO:0009072~aromatic amino acid family metabolic process, GO:0008483~transaminase activity, GO:0003824~catalytic activity		
extracellular region	47	N/A		
phosphatidylinositol-3-phosphate binding	66	N/A		
pre-autophagosomal structure	185	GO:0000407~pre-autophagosomal structure, GO:0005776~autophagosome	Temperature	
oxidoreductase activity	2613	GO:0016491~oxidoreductase activity, GO:0003824~catalytic activity, GO:0008483~transaminase activity		
nucleus	44	GO:0005634~nucleus, GO:0005764~lysosome		
response to virus	63	GO:0009615~response to virus, GO:0010998~regulation of translational initiation by eIF2 alpha phosphorylation, GO:0045071~negative regulation of viral genome replication		
antigen processing and presentation	68	GO:0019882~antigen processing and presentation, GO:0006955~immune response	Dissolved Oxygen	
serine-type endopeptidase inhibitor activity	34	N/A		
endopeptidase inhibitor activity	351	GO:0004866~endopeptidase inhibitor activity, GO:0005615~extracellular space, GO:0005576~extracellular region		
endopeptidase inhibitor activity2	1544	GO:0004866~endopeptidase inhibitor activity, GO:0016787~hydrolase activity, GO:0003824~catalytic activity		
oxidation-reduction process	1351	GO:0055114~oxidation-reduction process, GO:0016491~oxidoreductase activity, GO:0008483~transaminase activity		
acute-phase response	181	N/A		
serine-type peptidase activity3	237	GO:0008236~serine-type peptidase activity, GO:0008233~peptidase activity, GO:0006508~proteolysis		

Table 11 Continued.

Module Name	Number of Genes	Top Significant GO Terms	Negatively Correlated Traits	Positively Correlated Traits
ribosome	115	GO:0005840~ribosome, GO:0030529~intracellular ribonucleoprotein complex, GO:0003735~structural constituent of ribosome		Standard Length
nucleolus	402	GO:0005730~nucleolus, GO:0006364~rRNA processing, GO:0032040~small-subunit processome		
translation	320	GO:0006412~translation ,GO:0003743~translation initiation factor activity , GO:0006413~translational initiation		
ribosome3	1992	GO:0005840~ribosome, GO:0003735~structural constituent of ribosome, GO:0030529~intracellular ribonucleoprotein complex		
cytoplasm	662	GO:0005737~cytoplasm, GO:0005739~mitochondrion, GO:0005634~nucleus		
endoplasmic reticulum	2431	GO:0005783~endoplasmic reticulum, GO:0005789~endoplasmic reticulum membrane, GO:0006888~ER to Golgi vesicle-mediated transport		
mitochondrion	289	GO:0005739~mitochondrion, GO:0004129~cytochrome-c oxidase activity, GO:0005747~mitochondrial respiratory chain complex I		
fibrinogen complex	62	GO:0005577~fibrinogen complex, GO:0072378~blood coagulation, fibrin clot formation, GO:0051258~protein polymerization		
mRNA binding	192	GO:0003729~mRNA binding, GO:0005681~spliceosomal complex, GO:0003723~RNA binding		
ATP-dependent helicase activity	694	GO:0008026~ATP-dependent helicase activity		
oxidoreductase activity2	3736	GO:0016491~oxidoreductase activity, GO:0005739~mitochondrion, GO:0055114~oxidation-reduction process		
ribosome2	303	GO:0005840~ribosome, GO:0003735~structural constituent of ribosome, GO:0030529~intracellular ribonucleoprotein complex	Conductivity	Copper, Goby Density
intracellular	50	N/A		
lipid transporter activity	1050	GO:0005319~lipid transporter activity, GO:0071391~cellular response to estrogen stimulus, GO:0003824~catalytic activity		
serine-type endopeptidase inhibitor activity2	161	N/A		
nucleic acid binding	89	N/A		

Table 11 Continued.

Module Name	Number of Genes	Top Significant GO Terms	Negatively Correlated Traits	Positively Correlated Traits
positive regulation of macromolecule metabolic process	41	N/A		
cytoplasm2	4488	GO:0005737~cytoplasm, GO:0000166~nucleotide binding, GO:0005829~cytosol		
translation3	81	N/A		
mRNA splicing, via spliceosome	164	GO:0000398~mRNA splicing, via spliceosome, GO:0006412~translation, GO:0005681~spliceosomal complex		
mitochondrion2	889	GO:0005739~mitochondrion, GO:0005829~cytosol, GO:0030529~intracellular ribonucleoprotein complex		
nucleotide binding	1178	GO:0000166~nucleotide binding		
catalytic activity	171	GO:0003824~catalytic activity, GO:0045261~proton-transporting ATP synthase complex, catalytic core F(1), GO:0046933~proton-transporting ATP synthase activity, rotational mechanism	Standard Length	
oxidoreductase activity3	273	GO:0016491~oxidoreductase activity		
oxidoreductase activity4	324	N/A	Standard Length	
oxidoreductase activity5	867	GO:0016491~oxidoreductase activity, GO:0055114~oxidation-reduction process, GO:0003824~catalytic activity	Standard Length	
oxidoreductase activity6	1749	GO:0016491~oxidoreductase activity, GO:0005739~mitochondrion, GO:0003824~catalytic activity		Temperature
cytosol	79	GO:0005829~cytosol		
mitochondrion3	1039	GO:0005739~mitochondrion, GO:0005743~mitochondrial inner membrane, GO:0016491~oxidoreductase activity		
metabolic process	1199	GO:0008152~metabolic process, GO:0005739~mitochondrion, GO:0016491~oxidoreductase activity		
mRNA splicing, via spliceosome2	37	GO:0000398~mRNA splicing, via spliceosome, GO:0006412~translation, GO:0005681~spliceosomal complex		
cytosol2	66	N/A		
mitochondrion4	171	N/A		
mitochondrion5	210	N/A	Diversity	

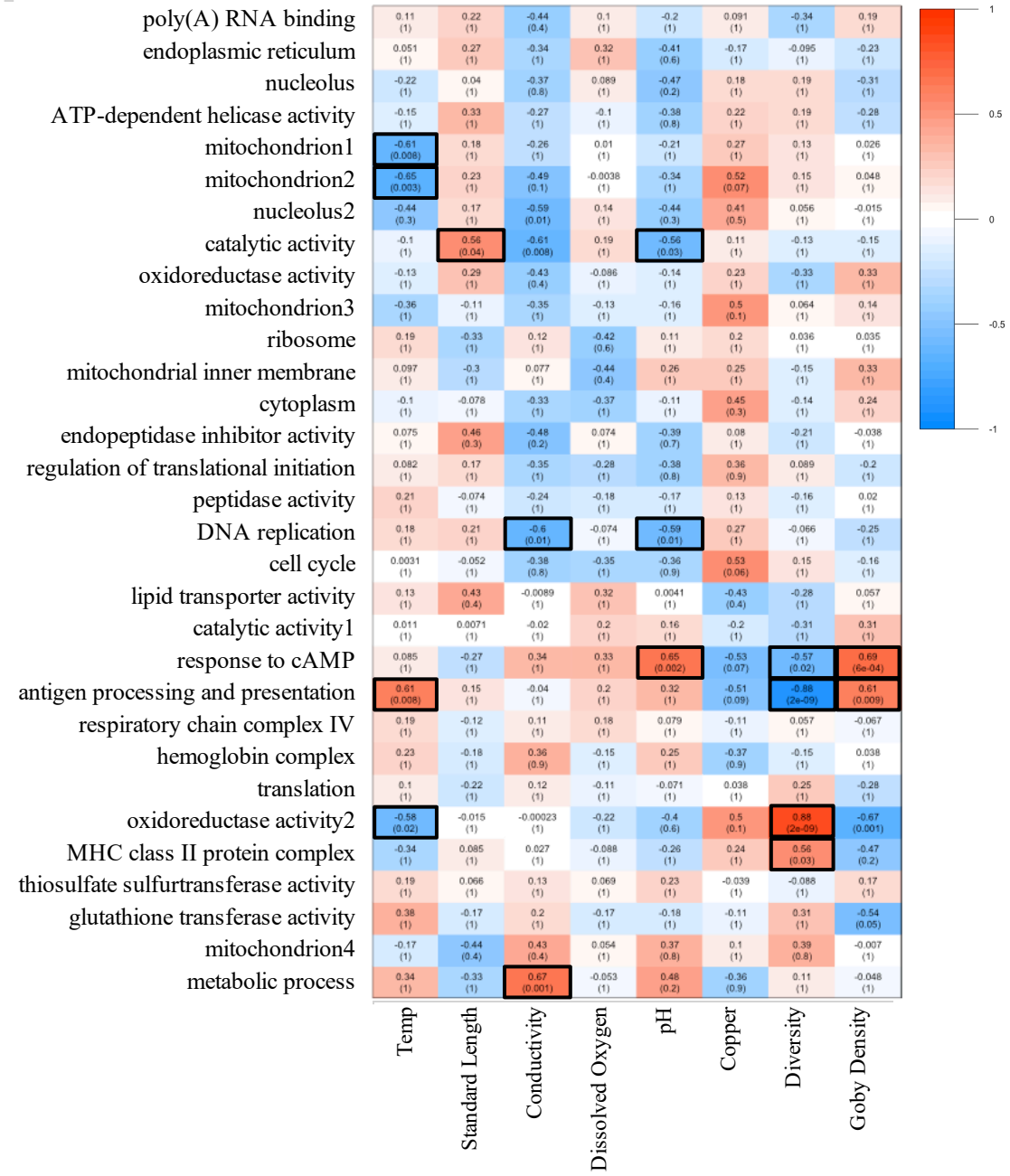


Figure 16. Heat map of module-trait relationships for round goby. Red represents positive correlations and blue represents negative correlations. The top number in each cell is the correlation value with p-values below in parentheses. Significant correlations are outlined in black.

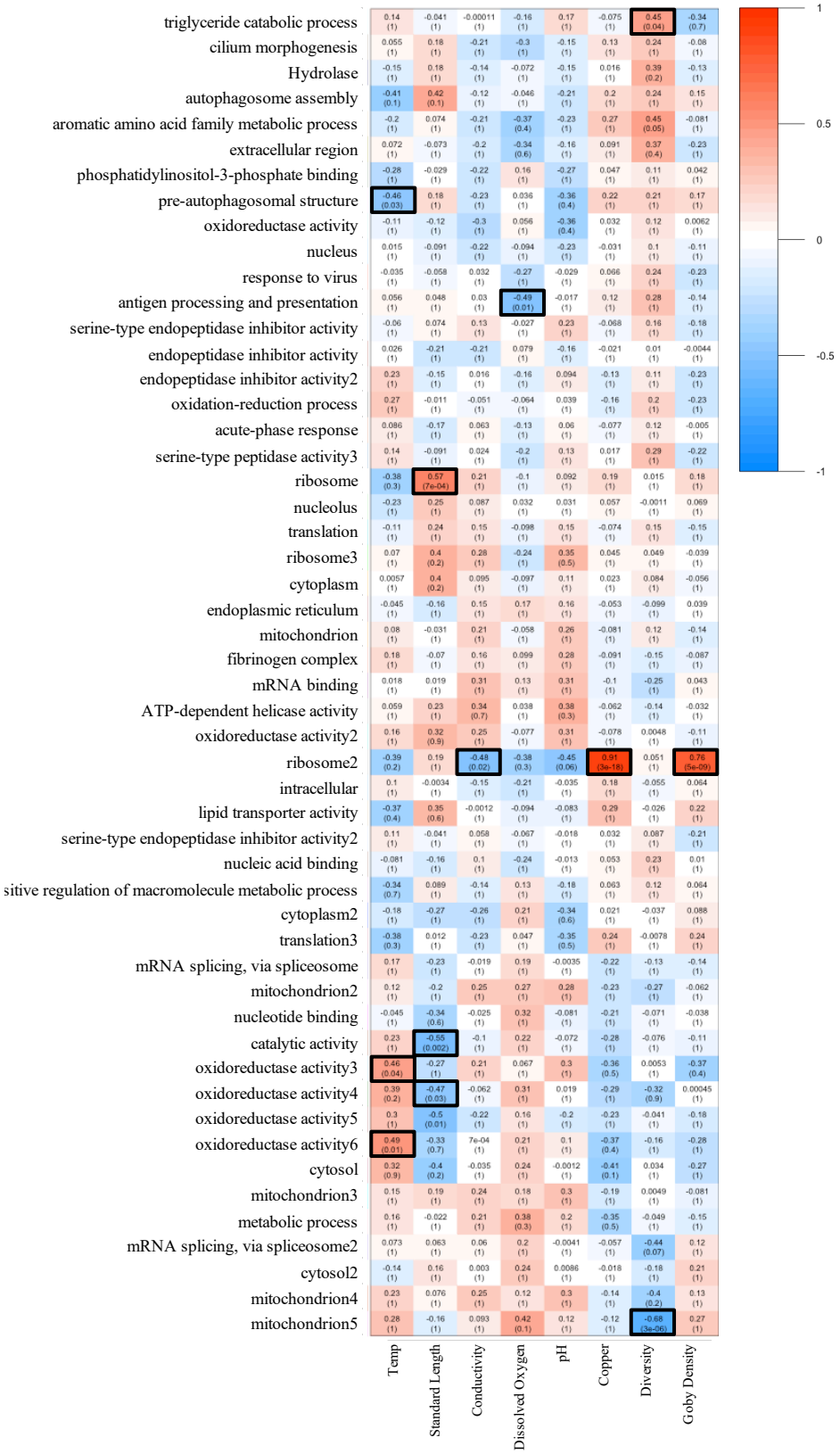


Figure 17. Heat map of module-trait relationships for Johnny darter. Red represents positive correlations and blue represents negative correlations. The top number in each cell is the correlation value with p-values below in parentheses. Significant correlations are outlined in black.

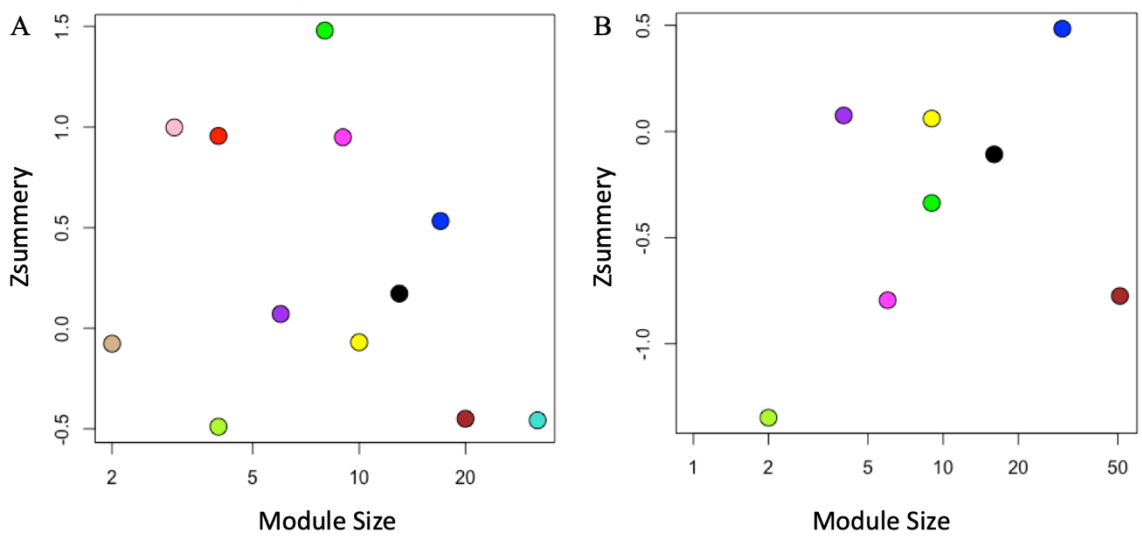


Figure 18. Module preservation Zsummary for A) Johnny darter network as reference and B) Round goby network as a reference. A Zsummery score of >2 is the threshold for preservation. No modules for either comparison met this threshold.

SIGNIFICANCE

Through the use of population genetic methods and genomic tools, this project expands the understanding of interactions between native and invasive species with their environment. This approach has rarely, if ever, been tried because of the many challenges presented when using uncontrolled wild populations; however, studies of non-model organisms in their natural environments are critically important as some effects may only be seen in the wild (Todd et al. 2016). In the wild, organisms are influenced by dynamic and synergistic factors, both biotic and abiotic which I attempted to capture in this research.

This is the first study to examine population genetics of Johnny darter in the Great Lakes region. The characterization the population genetic structure of Johnny darter in Lower Michigan offers insight into the historic processes that shaped the distribution of the species. I found high levels of genetic diversity and significant divergence among Johnny darter populations, demonstrating that local forces have been especially important in shaping the population structure in Johnny darter. Contrary to predictions, I did not detect significant differences between the east and west basins of Lower Michigan. However, multiple lineages were identified within Johnny darter populations. Future research could expand to a broader geographical range with markers that would allow for greater historical perspective.

Unlike the native species, I found low mitochondrial DNA diversity in round goby and no clear population structure, not surprising given the recent invasion of this species. The population structure of round goby in the Great Lakes and major waterways has been studied previously; this research contributes to this body of knowledge with the characterization of genetic diversity in round goby in recently invaded inland streams.

Understanding the evolutionary history of these species in Lower Michigan stream was an important precursor to the subsequent analysis of gene expression patterns. Evolutionary history

of a population may have a significant influence on patterns of gene expression, with patterns determined by genetic differences that have accumulated over time. Having completed this analysis, I concluded that there are limited levels of divergence among populations; however, local populations differ in the frequency of relatively similar haplotypes. Therefore, local evolutionary forces are more likely to play a significant role in gene expression patterns for round goby and Johnny darter than relatively deep evolutionary divergence.

Chapter 2 characterized patterns of gene expression and described hierarchical variation in study populations, exploring how these species respond to environmental heterogeneity. Here I found that gene expression patterns were drainage specific for both species and that expression was more similar for sites within the same river. Despite limited genetic divergence and limited time for local adaption, round goby exhibited greater differences in levels of gene expression in pairwise comparisons of groups. For Johnny darter, levels of gene expression were more locality specific while expression patterns in round goby showed more evidence of a response to transient environmental conditions. I also examined the gene expression manifestation of generalist strategy in round goby. A generalist species could have a static strategy that works well in many environments or vary its strategy in different environments. As plasticity is energetically costly, a static strategy could be advantageous for an invader. However, I found that gene expression in round goby varies across localities and different environmental conditions rather than having a static gene expression profile across multiple environments. These findings suggest a more diverse gene expression profile as compared to Johnny darter, adding to evidence that plasticity may play an important role in the success of invasive species (Lambrinos 2004, Davidson et al. 2011)

In chapter three, I identified groups of genes correlated with environmental variables, characterizing factors that may be driving patterns of gene expression observed in Chapter 2. In

this exploratory analysis, I identified factors that may warrant further research to understand how environmental conditions may favor round goby invasion. Plasticity has been found to be most advantageous in stressful environments (Richards et al. 2012, Xu et al. 2014), and this work sought to compare how Johnny darter and round goby compare in genes associated with stressors. The use of co-expression analysis facilitated the comparison of expression patterns of Johnny darter and round goby in correlated to variables measured by C. Krabbenhoft for her doctoral dissertation.

Johnny darter and round goby were markedly different in their gene expression associations with stressors in the environment. The variation in groups of genes correlated with temperature, salinity, pH, and copper suggested differing physiological responses to stressors for these two species. For Johnny darter, genes correlated with copper concentration were consistent with known effects of copper contamination on fish while round goby did not exhibit similar correlations. As Krabbenhoft (2019) identified copper contamination as a predictor of round goby invasion, this factor and potential contamination by other heavy metals should be a target for further research. Additionally, the differences in gene groups correlated with temperature, salinity, and pH are consistent with differences between the two species with round goby naturally inhabiting locations characterized by variation in these factors. Adaptation to a variable environment may have made round goby well suited to tolerate stressors present in heavily impacted streams like these characterized here.

This work provides some indication that round goby may be more tolerant to stressors and variation in the environment than Johnny darter. Identifying these factors is an important step in informing more targeted studies to determine how stressors vary in their impact on organisms and the role these may play in invasion success. As round goby continues to expand its range, understanding the role of the environment in the interaction between round goby and native species

will be valuable for making informed management decisions. A better understanding of how an invasive and native species interact with the environment contributes to the broader conservation goal of understanding what makes invasive species successful.

APPENDIX

Table S1: RNA integrity numbers (RIN) and sequencing output for Johnny darter and round goby liver tissues.

Sample ID	RIN	Total Reads
JD15_001	8.7	16248731
JD15_002	6.9	20915895
JD15_003	9.2	19646015
JD15_004	9.2	14282147
JD15_005	9.5	16833967
JD15_006	7.2	14427793
JD15_007	9.4	21079881
JD15_008	9.0	19146022
JD15_009	9.4	20436118
JD15_010	8.9	20049481
JD15_011	9.3	15431200
JD15_012	8.9	15825941
JD15_013	9.1	15924936
JD15_014	9.0	15105560
JD15_015	9.1	17289211
JD15_016	9.3	15896149
JD15_017	8.7	19559660
JD15_018	8.7	16422545
JD15_019	8.3	12894889
JD15_020	9.0	13928627
JD15_021	7.8	14793069
JD15_022	8.7	13478420
JD16_001	8.6	11789063
JD16_002	9.2	15610068
JD16_003	8.8	27120376
JD16_004	8.6	10329989
JD16_005	8.9	17106137
JD16_006	8.5	24985750
JD16_007	8.0	15263387
JD16_008	8.2	14923903
JD16_009	8.5	15884307
JD16_010	8.1	14157044
JD16_011	9.0	14279770
JD16_012	8.9	16164213
JD16_013	7.7	17279452
JD16_014	9.3	16310305
JD16_015	9.2	16419423
JD16_016	9.4	14410682
JD16_017	8.3	15327077
JD16_018	7.0	13845440
JD16_019	8.4	14178044
JD16_020	5.6	14958561
JD16_021	7.8	17451737
JD16_022	8.6	20901138
JD16_023	8.7	19498926
JD17_001	9.3	16828770
JD17_002	9.2	14851990
JD17_003	9.6	13224873
JD17_005	7.9	14282741
JD17_006	8.2	17501559
JD17_007	9.0	23607075
JD17_008	9.5	12440182
	Total	860548239

Sample	RIN	Total Reads
RG15_001	8.8	13659483
RG15_002	9.2	14542429
RG15_003	8.1	12406386
RG15_004	9.6	11033705
RG15_005	6.3	14966204
RG15_006	6.2	11386463
RG15_007	7.8	13313047
RG15_008	8.7	15333636
RG15_009	9.4	16138062
RG15_010	9.1	19274103
RG15_011	9.6	14996115
RG15_012	9.2	31133198
RG16_001	7.1	12641523
RG16_002	7.6	13292598
RG16_003	8.2	14922440
RG16_004	8.8	14756868
RG16_005	8.1	15799533
RG16_006	7.7	15450773
RG16_007	7.1	14217313
RG16_008	8.3	9049955
RG16_009	8.8	14037370
RG16_010	9.1	14877684
RG16_011	8.7	15092362
RG16_012	9.2	15939205
RG16_013	8.9	8788084
RG16_014	8.1	17657680
RG17_001	9.3	23758997
RG17_003	9.5	15001446
RG17_004	8.9	14665471
RG17_005	9.7	11947229
RG17_006	8.9	14918533
RG17_007	8.3	21859996
RG17_009	9.5	20652281
	Total	507510172

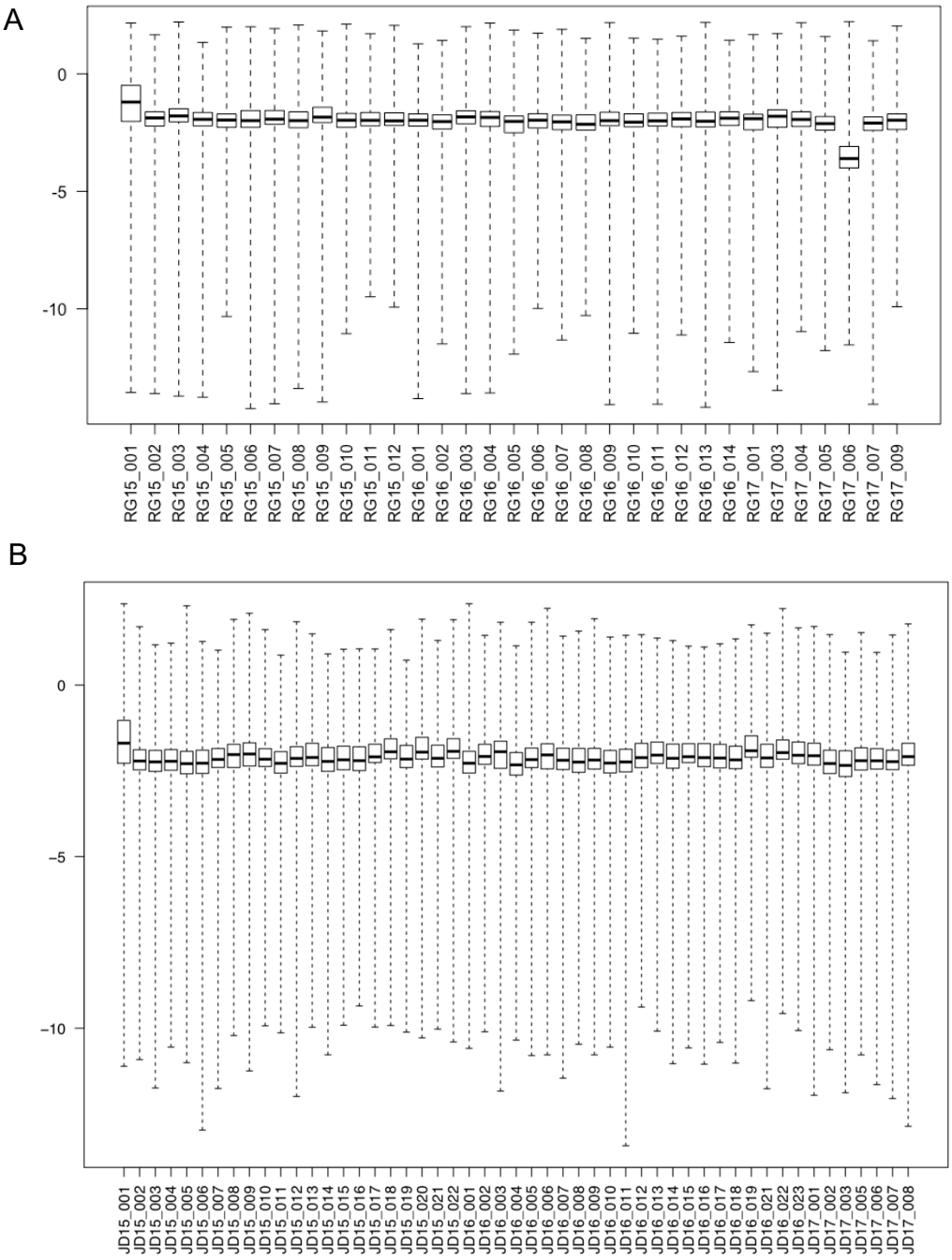


Figure S2. Boxplot of Cook's distance as a measure of each sample's influence on the fitted averages for each gene for A) round goby B) Johnny darter. Round goby sample RG17_006 was judged to be an outlier and removed from the analysis.

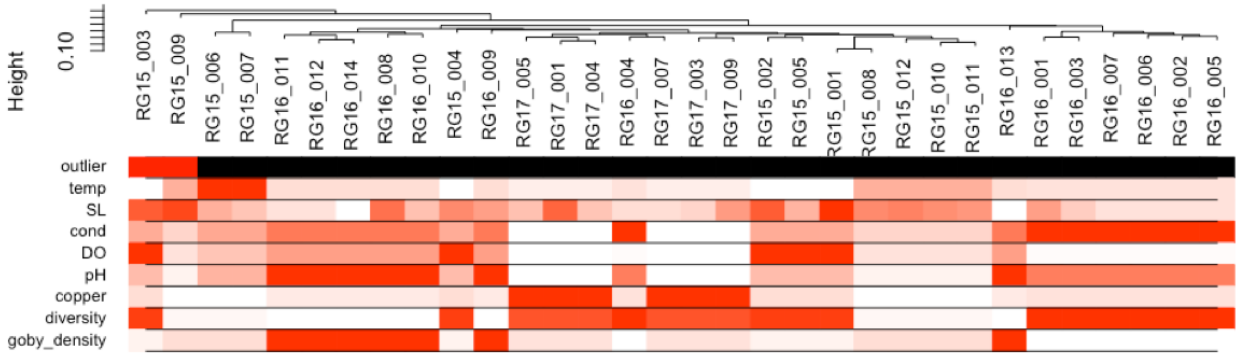


Figure S3. Distance tree of overall expression patterns for round goby. The heatmap shows gradations of red represent the value of the trait for each individual, i.e. for temperature red indicates higher temperature and white indicates cooler temperatures. RG15_003 and RG15_009 were identified as outliers and removed.

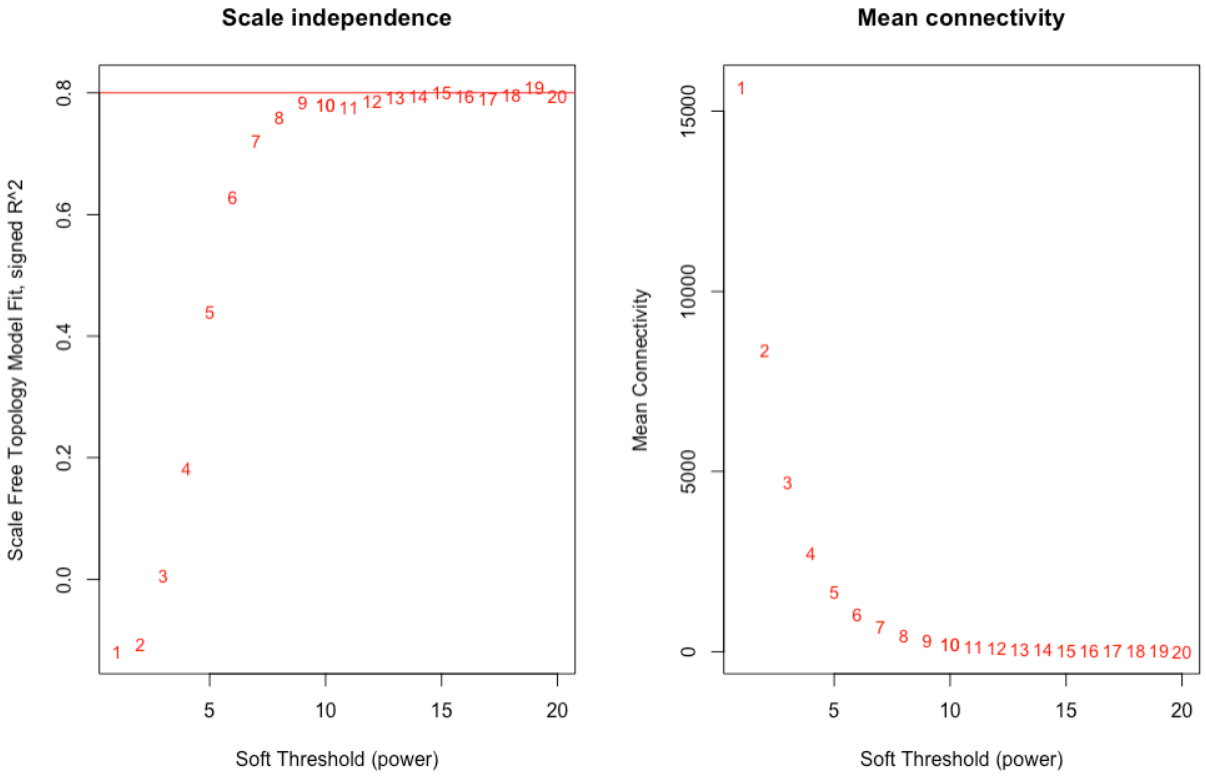


Figure S4. Plots of soft threshold power for round goby. The lowest power which reaches a scale free topology of 0.8 is chosen which was 15.

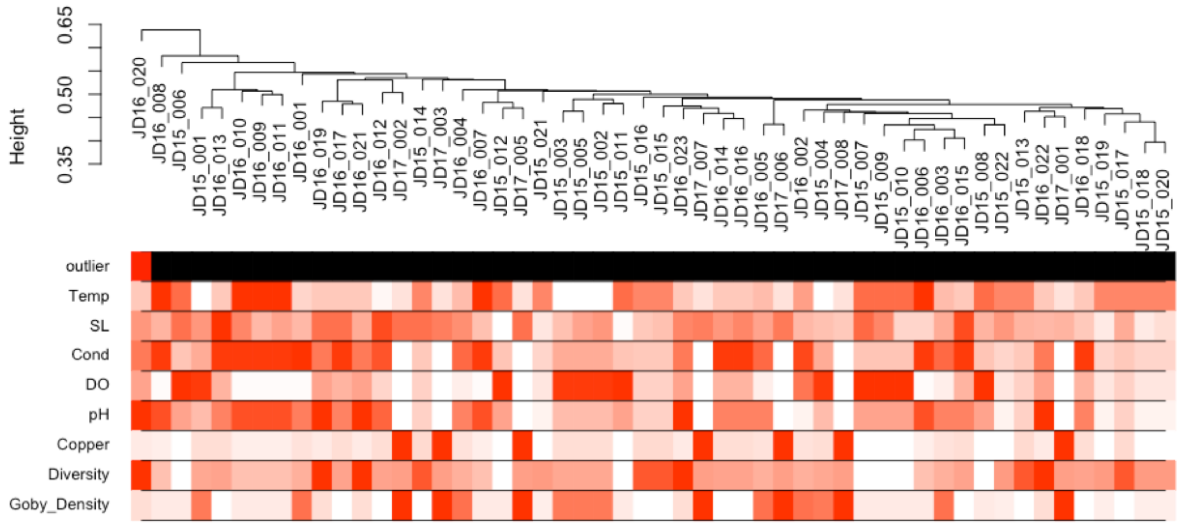


Figure S5. Distance tree of overall expression patterns for Johnny darter and traits. The heatmap indicates scale for the traits measures i.e. for temperature red indicates higher temperature and white indicates lower temperatures. JD16_020 was noted as an outlier and removed.

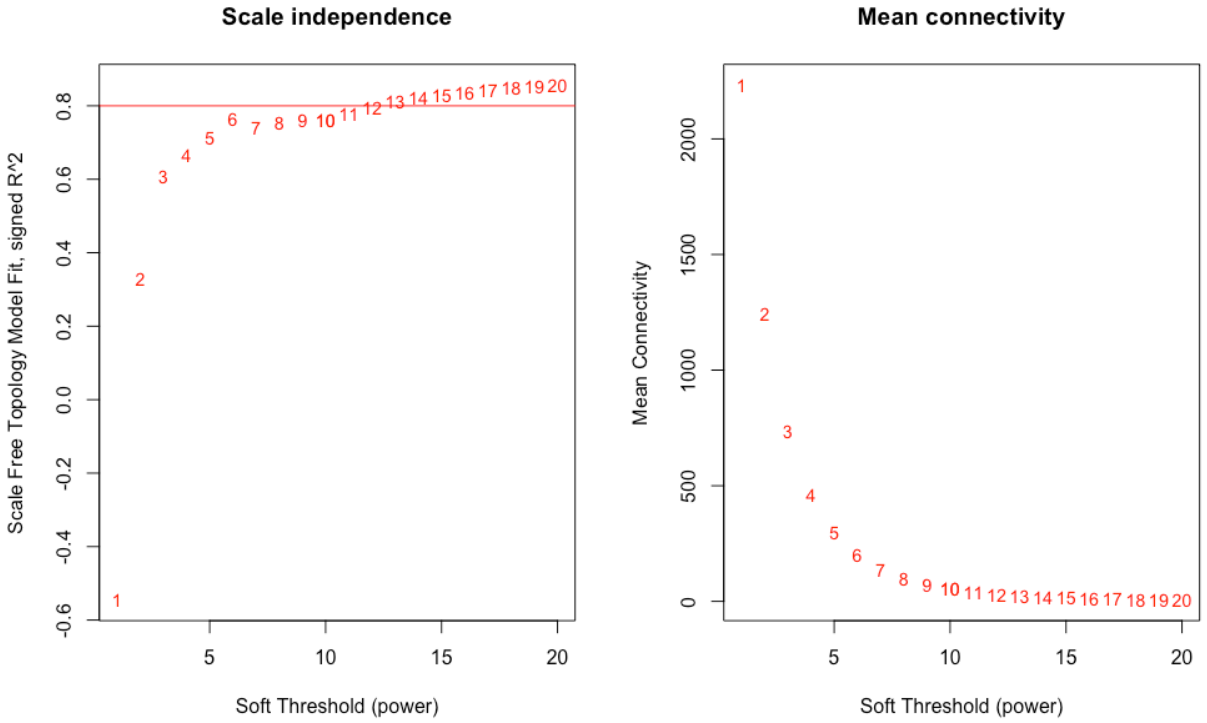


Figure S6. Plots of soft threshold power for Johnny darter. The lowest power which reaches a scale free topology of 0.8 is chosen which was for 12.

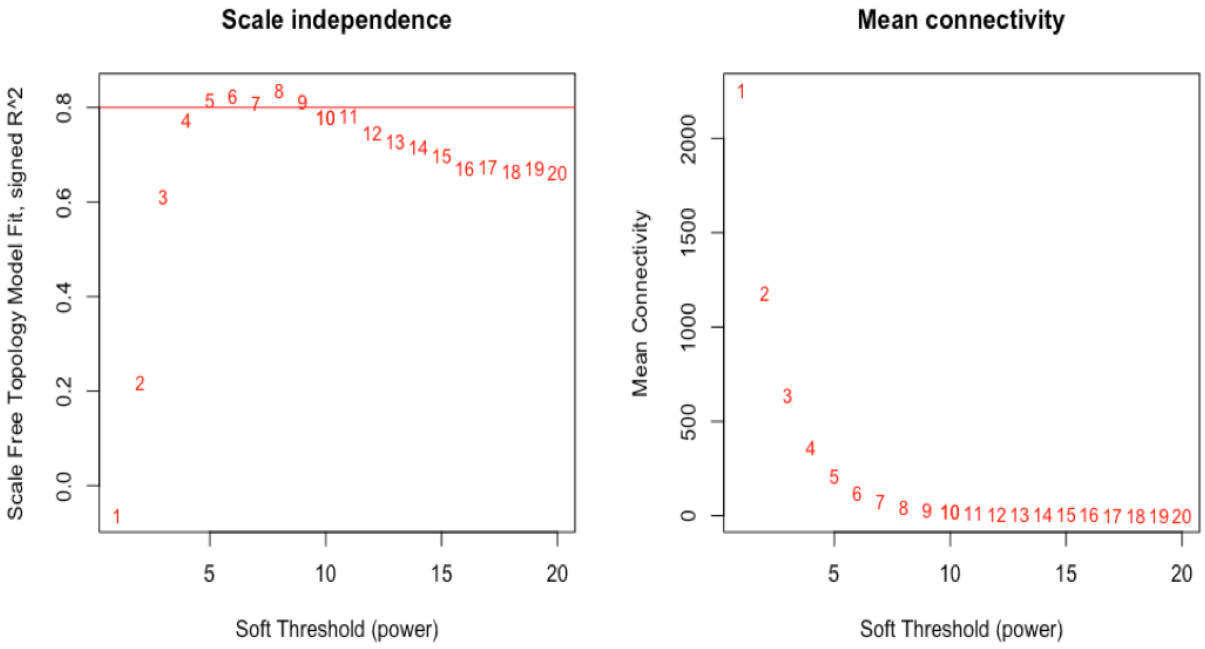


Figure S7. Plots of soft threshold power for Johnny darter reciprocal blast gene set. The lowest power which reaches a scale free topology of 0.8 is chosen which was 5.

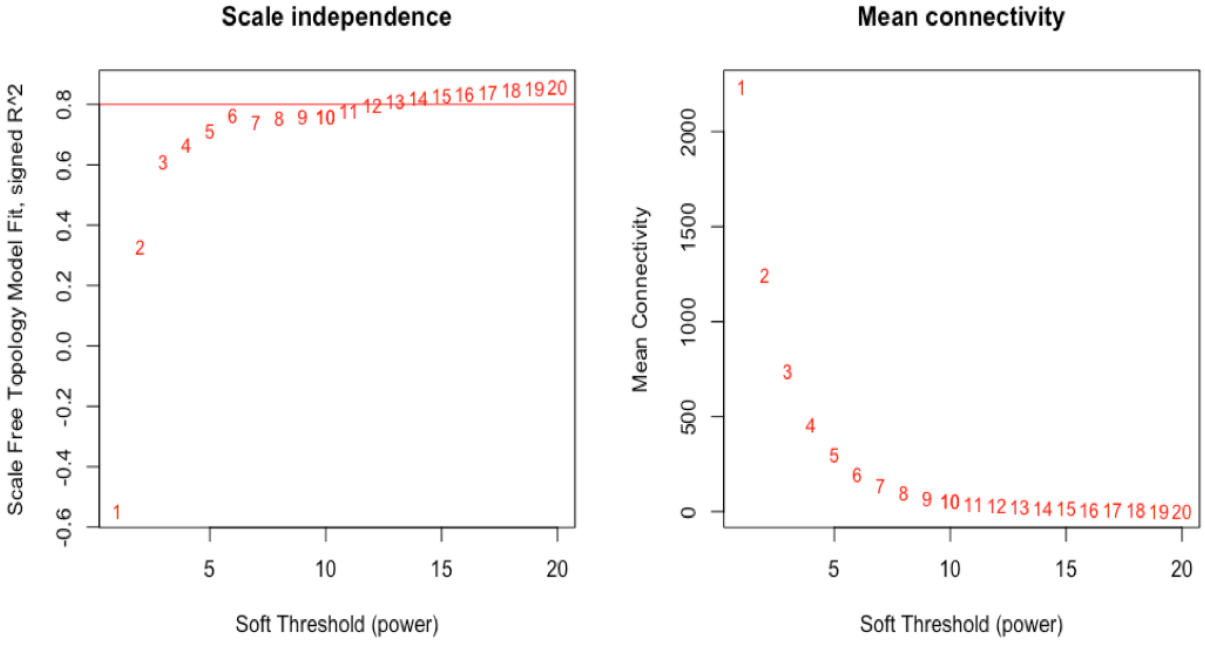


Figure S8. Plots of soft threshold power for round goby for the reciprocal gene set. The lowest power which reaches a scale free topology of 0.8 is chosen which was 12.

REFERENCES

Airoldi, L., Turon, X., Perkol-Finkel, S. and Rius, M. 2015. Corridors for aliens but not for natives: effects of marine urban sprawl at a regional scale. Diversity and Distributions 21: 755-768.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. Journal of molecular biology 215: 403-410.

Aluru, N., and M. M. Vijayan. 2009. Stress transcriptomics in fish: A role for genomic cortisol signaling. General and Comparative Endocrinology 164:142-150.

Anders, S. and Huber, W. 2010. Differential expression analysis for sequence count data. Genome biology 11:1.

Andersson, T. and Förlin, L. 1992. Regulation of the cytochrome P450 enzyme system in fish. Aquatic Toxicology 24: 1-19.

Avise, J.C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos 62-76.

Azour, F., van Deurs, M., Behrens, J., Carl, H., Hüssy, K., Greisen, K., Ebert, R. and Møller, P.R. 2015. Invasion rate and population characteristics of the round goby *Neogobius melanostomus*: effects of density and invasion history. Aquatic Biology 24: 41-52.

Bahamonde, P.A., McMaster, M.E., Servos, M.R., Martyniuk, C.J. and Munkittrick, K.R. 2015. Molecular pathways associated with the intersex condition in rainbow darter (*Etheostoma caeruleum)* following exposures to municipal wastewater in the Grand River basin, ON, Canada. Part B. Aquatic Toxicology 159: 302-316.

Balshine, Sigal, Aikta Verma, Virginia Chant, and Tys Theysmeyer. 2005. Competitive interactions between round gobies and logperch. Journal of Great Lakes Research 31: 68-77.

Bailey, R.M. and Smith, G.R. 1981. Origin and geography of the fish fauna of the Laurentian Great Lakes basin. Canadian Journal of Fisheries and Aquatic Sciences 38: 1539-1561.

Barber, L.B., Brown, G.K., Nettesheim, T.G., Murphy, E.W., Bartell, S.E. and Schoenfuss, H.L. 2011. Effects of biologically-active chemical mixtures on fish in a wastewater-impacted urban stream. Science of the Total Environment 409: 4720-4728.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. The Quarterly Review of Biology 74: 21-45.

Bolger, A.M., Lohse, M. and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 170.

Boltana, S., Rey, S., Roher, N., Vargas, R., Huerta, M., Huntingford, F.A., Goetz, F.W., Moore, J., Garcia-Valtanen, P., Estepa, A. and MacKenzie, S. 2013. Behavioural fever is a synergic signal amplifying the innate immune response. Proceedings of the Royal Society B: Biological Sciences 280: 20131381.

Bossu, C.M. and Near, T.J. 2009. Gene trees reveal repeated instances of mitochondrial DNA introgression in orangethroat darters (Percidae*: Etheostoma*). Systematic biology 58: 114-129.

Bowley LA, Alam F, Marentette JR, Balshine S, Wilson JY. 2010. Characterization of vitellogenin gene expression in round goby (*Neogobius melanostomus*) using a quantitative polymerase chain reaction assay. Environmental Toxicology and Chemistry. 29: 2751-60.

Bradley, C. A., and S. Altizer. 2007. Urbanization and the ecology of wildlife diseases. Trends in Ecology and Evolutionary Biology 22:95-102

Bronnenhuber, J.E., Dufour, B.A., Higgs, D.M. and Heath, D.D. 2011. Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round goby, *Neogobius melanostomus*, in Great Lakes tributaries. Molecular ecology, 20: 1845-1859.

Brown, J.E. and Stepien, C.A., 2009. Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. Molecular Ecology, 18: 64-79.

Buckley, B. A., & Hofmann, G. E. 2002. Thermal acclimation changes DNA-binding activity of heat shock factor 1 (HSF1) in the goby *Gillichthys mirabilis*: implications for plasticity in the heat-shock response in natural populations. Journal of Experimental Biology 205: 3231-3240.

Busby, M.A., Stewart, C., Miller, C.A., Grzeda, K.R. and Marth, G.T. 2013. Scotty: a web tool for designing RNA-Seq experiments to measure differential gene expression. Bioinformatics, 29: 656-657.

Cardinale, B.J. 2011. Biodiversity improves water quality through niche partitioning. Nature, 472: 86.

Carroll, S.B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134: 25-36.

Cheviron, Z.A. and Swanson, D.L. 2017. Comparative transcriptomics of seasonal phenotypic flexibility in two North American songbirds. Integrative and Comparative Biology 57: 1040-1054.

Cheviron, Z.A., Whitehead, A. and Brumfield, R.T. 2008. Transcriptomic variation and plasticity in rufous-collared sparrows (*Zonotrichia capensis*) along an altitudinal gradient. Molecular ecology 17: 4556-4569.

Clarke, A. and Johnston, N.M. 1999. Scaling of metabolic rate with body mass and temperature in teleost fish. Journal of Animal Ecology 68: 893-905.

Cockerill, K., Anderson Jr, W.P., Harris, F.C. and Straka, K. 2017. Hot, Salty Water: A Confluence of Issues in Managing Stormwater Runoff for Urban Streams. Jawra Journal of the American Water Resources Association 53: 707-724.

Copeland, P.A., Sumpter, J.P., Walker, T.K. and Croft, M. 1986. Vitellogenin levels in male and female rainbow trout (Salmo gairdneri Richardson) at various stages of the reproductive cycle. Comparative biochemistry and physiology. B, Comparative biochemistry 83: 487-493.

Corkum, L.D., Sapota, M.R. and Skora, K.E. 2004. The round goby, *Neogobius melanostomus*, a fish invader on both sides of the Atlantic Ocean. Biological invasions 6: 173-181.

Coulter, D.P., Murry, B.A., Webster, W.C. and Uzarski, D.G., 2011. Effects of dreissenid mussels, chironomids, fishes, and zooplankton on growth of round goby in experimental aquaria. Journal of Freshwater Ecology, 26: 155-162.

Crane, D.P. and Einhouse, D.W. 2016. Changes in growth and diet of smallmouth bass following invasion of Lake Erie by the round goby. Journal of Great Lakes Research, 42: 405-412.

Crooks, J.A. 2002. Characterizing ecosystem-level consequences of biological invasions: the role of ecosystem engineers. Oikos, 97: 153-166.

Cross, E.E. and Rawding, R.S. 2009. Acute thermal tolerance in the round goby (*Neogobius melanostomus*). Journal of Thermal Biology 34: 85-92.

Culumber, Z. W., Schumer, M., Monks, S., & Tobler, M. 2015. Environmental heterogeneity generates opposite gene-by-environment interactions for two fitness-related traits within a population. Evolution 69: 541-550.

Davidson, A.M., Jennions, M. and Nicotra, A.B. 2011. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. Ecology letters 14: 419-431

Dabney, A., Storey, J.D. and Warnes, G.R. 2010. qvalue: Q-value estimation for false discovery rate control. R package version, 1(0).

DeWitt, T.J., Sih, A. and Wilson, D.S. 1998. Costs and limits of phenotypic plasticity. Trends in ecology & evolution 13: 77-81.

Drouillard, K.G., Feary, D.A., Sun, X., O'Neil, J.A., Leadley, T. and Johnson, T.B. 2018. Comparison of thermal tolerance and standard metabolic rate of two Great Lakes invasive fish species. Journal of Great Lakes research 44: 476-481.

Dowling, T.E., Broughton, R.E. and DeMarais, B.D. 1997. Significant role for historical effects in the evolution of reproductive isolation: evidence from patterns of introgression between the cyprinid fishes, Luxilus cornutus and Luxilus chrysocephalus. Evolution:1574-1583.

Dubs, D. O. L. and L.D. Corkum. 1996. Behavior interactions between round gobies (*Neogobius melanostomus*) and mottled sculpin (*Cottus bairdii*). Journal of Great Lakes Research 22: 838-844.

Dufour, B.A., Hogan, T.M. and Heath, D.D. 2007. Ten polymorphic microsatellite markers in the invasive round goby (*Neogobius melanostomus*) and cross-species amplification. Molecular Ecology Notes 7:1205-1207.

Ekblom R, Galindo J. 2011 Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 10:1-5.

Excoffier, L. and Lischer, H.E. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564-567.

Eyckmans, M., Celis, N., Horemans, N., Blust, R. and De Boeck, G. 2011. Exposure to waterborne copper reveals differences in oxidative stress response in three freshwater fish species. Aquatic Toxicology 103: 112-120.

Farrell, A.P. 2002. Cardiorespiratory performance in salmonids during exercise at high temperature: insights into cardiovascular design limitations in fishes. Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology 132: 797-810.

Feder, M. and G. Hoffmann. 1999. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and Ecological Physiology. Annual Review of Physiology 61: 243-282.

Filteau, M., Pavey, S.A., St-Cyr, J. and Bernatchez, L. 2013. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. Molecular Biology and Evolution 30: 1384-1396.

FOTR Friends of the Rouge and Alliance of Rouge Communities. 2018. Report on Fish Surveys by Friends of the Rouge. https://drive.google.com/file/d/15_GU3HSSZBnZZLJ31NYMpj0j-37Nt7Ob/view

Gach, M. 1996. Geographic variation in mitochondrial DNA and biogeography of *Culaea inconstans* (Gasterosteidae). Copeia 3: 563-575.

Ghedotti, M.J., Smihula, J.C. and Smith, G.R. 1995. Zebra mussel predation by round gobies in the laboratory. Journal of Great Lakes Research 21: 665-669.

Gene Ontology Consortium. 2004. The Gene Ontology (GO) database and informatics resource. Nucleic acids research 32: D258-D261.

Gendron, A.D., D.J. Marcogliese, and M. Thomas. Invasive species are less parasitized than native competitors, but for how long? The case of the round goby in the Great Lakes-St. Lawrence Basin. Biological invasions 14: 367-384.

Great Lakes Aquatic Nonindigenous Species Information System (GLANSIS). http://www.glerl.noaa.gov/res/Programs/glansis/glansis.html.

Goetz, F. W., & MacKenzie, S. 2008. Functional genomics with microarrays in fish biology and fisheries. Fish and Fisheries 9: 378-395.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9. 3).

Green SJ, Akins JL, Maljković A, Côté IM. 2012. Invasive Lionfish Drive Atlantic Coral Reef Fish Declines. PLoS ONE e32596.

Gutowsky, L.F.G. and Fox, M.G. 2012. Intra-population variability of life-history traits and growth during range expansion of the invasive round goby, *Neogobius melanostomus*. Fisheries Management and Ecology 19: 78-88.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M. and MacManes, M.D. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature protocols 8: 1494-1512.

Hall, T., Biosciences, I. and Carlsbad, C. 2011. BioEdit: an important software for molecular biology. GERF Bull Biosci 2: 60-61.

Haverinen, J. and Vornanen, M. 2009. Responses of action potential and K+ currents to temperature acclimation in fish hearts: phylogeny or thermal preferences? Physiological and Biochemical Zoology 82: 468-482.

Heithaus, M.R. and Laushman, R.H. 1997. Genetic variation and conservation of stream fishes: influence of ecology, life history, and water quality. Canadian Journal of Fisheries and Aquatic Sciences 54: 1822-1836.

He, X., Chaganti, S.R. and Heath, D.D. 2018. Population-specific responses to interspecific competition in the gut microbiota of two Atlantic salmon (*Salmo salar*) populations. Microbial ecology 75: 140-151.

Heckman, K.L., Near, T.J. and Alonzo, S.H. 2009. Phylogenetic relationships among *Boleosoma* darter species (Percidae: *Etheostoma*). Molecular Phylogenetics and Evolution 53: 249-257.

Hedrick, P.W. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution 53: 313-318.

Hitt, N.P. and Angermeier, P.L. 2008. Evidence for fish dispersal from spatial analysis of stream network topology. Journal of the North American Benthological Society 27: 304-320.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics: 65-70.

Homola, J.J., Ruetz III, C.R., Kohler, S.L. and Thum, R.A. 2016. Complex postglacial recolonization inferred from population genetic structure of mottled sculpin *Cottus bairdii* in tributaries of eastern Lake Michigan, USA. Journal of Fish Biology 89: 2234-2250.

Hontela, A., 1998. Interrenal dysfunction in fish from contaminated sites: in vivo and in vitro assessment. Environmental Toxicology and Chemistry 17: 44-48.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protocols 4: 44-57

Huntingford, F. A., C. Adams, V. A. Braithwaite, S. Kadri, T. G. Pottinger, P. Sandoe, and J. F. Turnbull. 2006. Current issues in fish welfare. Journal of Fish Biology 68: 332-372.

Ingersoll, C.G. and Claussen, D.L. 1984. Temperature selection and critical thermal maxima of the fantail darter, *Etheostoma flabellare*, and Johnny darter, *E. nigrum*, related to habitat and season. In Environmental Biology of Darters 95-102.

Iwama, G. K., Thomas, P. T., Forsyth, R. B., & Vijayan, M. M. 1998. Heat shock protein expression in fish. Reviews in Fish Biology and Fisheries 8: 35-56.

Jaffe, A.E., Hyde, T., Kleinman, J., Weinbergern, D.R., Chenoweth, J.G., McKay, R.D., Leek, J.T. and Colantuoni, C. 2015. Practical impacts of genomic data "cleaning" on biological discovery using surrogate variable analysis. BMC bioinformatics 16: 372.

Jakubčinová, K., Haruštiaková, D., Števove, B., Švolíková, K., Makovinská, J. and Kováč, V. 2018. Distribution patterns and potential for further spread of three invasive fish species (*Neogobius melanostomus, Lepomis gibbosus* and *Pseudorasbora parva*) in Slovakia. Aquatic Invasions 13.

Janssen, J and D. J. Jude. 2001. Recruitment failure of mottled sculpin *Cottus bairdii*in Calumet harbor, Southern Lake Michigan, induced by the newly introduced round goby *Neogobius melanostomus*. Journal of Great Lakes Research 27: 319-328.

Jobling, S., Nolan, M., Tyler, C.R., Brighty, G. and Sumpter, J.P. 1998. Widespread sexual disruption in wild fish. Environmental science & technology 32: 2498-2506.

Jude, D.J., Reider, R.H. and Smith, G.R.1992. Establishment of Gobiidae in the Great Lakes basin. Canadian Journal of Fisheries and Aquatic Sciences49: 416-421.

Karsiotis, S.I., Pierce, L.R., Brown, J.E. and Stepien, C.A. 2012. Salinity tolerance of the invasive round goby: experimental implications for seawater ballast exchange and spread to North American estuaries. Journal of Great Lakes Research 38: 121-128.

Karatayev, A.Y., Burlakova, L.E., Padilla, D.K., Mastitsky, S.E. and Olenin, S. 2009. Invaders are not a random selection of species. Biological Invasions 11.

Kornis, M.S., N. Mercado-Silva, and M.J. VanderZanden. 2012. Twenty years of invasion: a review of round goby *Neogobius melanostomus* biology, spread and ecological implications. Journal of Fish Biology 80: 235-285

Kornis, M. S., Sharma, S., & Jake Vander Zanden, M. 2013. Invasion success and impact of an invasive fish, round goby, in Great Lakes tributaries. Diversity and Distributions 19: 184-198.

Krause, R.J., Grant, J.W., McLaughlin, J.D. and Marcogliese, D.J. 2010. Do infections with parasites and exposure to pollution affect susceptibility to predation in Johnny darters (*Etheostoma nigrum*)? Canadian Journal of Zoology, 88: 218-1225.

Krabbenhoft, Corey, "Drivers And Impacts Of The Invasive Round Goby (*Neogobius Melanostomus*) In Michigan Tributaries To The Great Lakes" (2019). Wayne State University Dissertations. 2171.

Krueger, F. 2015. Trim galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution 35: 1547-1549.

Kvamme, B.O., Gadan, K., Finne-Fridell, F., Niklasson, L., Sundh, H., Sundell, K., Taranger, G.L. and Evensen, Ø. 2013. Modulation of innate immune responses in Atlantic salmon by chronic hypoxia-induced stress. Fish and Shellfish Immunology 34: 55-65.

Lafontaine, P. and Dodson, J.J. 1997. Intraspecific genetic structure of white sucker (*Catostomus commersoni*) in northeastern North America as revealed by mitochondrial DNA polymorphism. Canadian Journal of Fisheries and Aquatic Sciences 54: 555-565.

Lambrinos, J.G. 2004. How interactions between ecology and evolution influence contemporary invasion dynamics. Ecology 85: 2061-2070.

Langfelder, P. and Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9: 559.

Langmead, B. and Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature methods, 9: 357-359.

Lauer, T.E., P. Allen, and T.S. McComish. 2004. Changes in mottled sculpin and Johnny darter trawl catches after the appearance of round gobies in the Indiana waters of Lake Michigan. Transactions of the American Fisheries Society 131: 185-189.

Lee, C.E., 2002. Evolutionary genetics of invasive species. Trends in ecology & evolution 17: 386-391.

Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E. and Storey, J.D. 2012. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28: 882-883.

Leidy, R.A.1992. Microhabitat selection by the Johnny darter, *Etheostoma nigrum* Rafinesque, in a Wyoming stream. The Great Basin Naturalist. 68-74.

Leigh, J.W. and Bryant, D. 2015. popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution 6: 1110-1116.

Lenz, M., da Gama, B.A., Gerner, N.V., Gobin, J., Gröner, F., Harry, A., Jenkins, S.R., Kraufvelin, P., Mummelthei, C., Sareyka, J. and Xavier, E.A. 2011. Non-native marine invertebrates are more tolerant towards environmental stress than taxonomically related native species: results from a globally replicated study. Environmental Research 111: 943-952.

Li, B. and Dewey, C.N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics*

Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 27: 2987-93.

Livak, K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. methods. 25: 402-408.

Lockwood, J.L., Cassey, P. and Blackburn, T. 2005. The role of propagule pressure in explaining species invasions. Trends in Ecology & Evolution 20: 223-228.

Love, M.I., Huber, W. and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15:550.

Lushchak, V.I. and Bagnyukova, T.V. 2006. Effects of different environmental oxygen levels on free radical processes in fish. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 144: 283-289.

Lydy, M.J. and Wissing, T.E. 1988. Effect of sublethal concentrations of copper on the critical thermal maxima (CTMax) of the fantail (*Etheostoma flabellare*) and Johnny (*E. nigrum*) darters. Aquatic Toxicology 12: 311-321.

MacGuigan, D.J. and Near, T.J. 2018. Phylogenomic Signatures of Ancient Introgression in a Rogue Lineage of Darters (Teleostei: Percidae*)*. Systematic Biology 68: 329-346.

Malone, Margaret Ann. 2016. Early Round Goby (*Neogobius melanostomus*) Invasion Into Lake Michigan Tributaries and Competitive Interactions with Two Native Benthic Fishes. Master's Theses.

Marentette, J. R., Gooderham, K. L., McMaster, M. E., Ng, T., Parrott, J. L., Wilson, J. Y., and Balshine, S. 2010. Signatures of contamination in invasive round gobies (Neogobius melanostomus): A double strike for ecosystem health? Ecotoxicology and environmental safety, 73: 1755-1764.

Marentette, J.R., Tong, S. and Balshine, S. 2013. The cortisol stress response in male round goby (Neogobius melanostomus): effects of living in polluted environments? Environmental biology of fishes 96: 723-733.

Marsh, P.C. and Pacey, C.A. 2005. Immiscibility of native and nonnative fishes. Restoring native fish to the lower Colorado River: interactions of native and nonnative fishes. US Fish and Wildlife Service, Albuquerque, NM, and US Bureau of Reclamation, Boulder City, NV59-63.

MacInnis, A.J. and Corkum, L.D. 2000. Fecundity and reproductive season of the round goby *Neogobius melanostomus* in the upper Detroit River. Transactions of the American Fisheries Society 121:136-144.

McCallum, E.S., Du, S.N., Vaseghi-Shanjani, M., Choi, J.A., Warriner, T.R., Sultana, T., Scott, G.R. and Balshine, S. 2017. In situ exposure to wastewater effluent reduces survival but has little effect on the behaviour or physiology of an invasive Great Lakes fish. Aquatic Toxicology 184: 37-48.

McIntyre, L.M., Lopiano, K.K., Morse, A.M., Amin, V., Oberg, A.L., Young, L.J. and Nuzhdin, S.V. 2011. RNA-seq: technical variability and sampling. BMC genomics 12.

Michael Lewis, C.F., Karrow, P.F., Blasco, S.M., McCarthy, F.M., King, J.W., Moore Jr, T.C. and Rea, D.K. 2008. Evolution of lakes in the Huron basin: deglaciation to present. Aquatic Ecosystem Health & Management 11: 27-136.

Meynard, C.N., Mouillot, D., Mouquet, N. and Douzery, E.J. 2012. A phylogenetic perspective on the evolution of mediterranean teleost fishes. PLoS One e36443

Mohammed, R.S., Reynolds, M., James, J., Williams, C., Mohammed, A., Ramsubhag, A., Van Oosterhout, C. and Cable, J. 2016. Getting into hot water: sick guppies frequent warmer thermal conditions. Oecologia 181: 911-917.

Müllner, D. 2013. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. Journal of Statistical Software 53: 1-18.

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

Neumann, J. 1991. Climate of the Black Sea region around 0 CE. Climatic Change 18: 453-465.

Newman, D. and Tallmon, D.A. 2001. Experimental evidence for beneficial fitness effects of gene flow in recently isolated populations. Conservation Biology 15: 1054-1063.

Nilsson, G.E. and Renshaw, G.M. 2004. Hypoxic survival strategies in two fishes: extreme anoxia tolerance in the North European crucian carp and natural hypoxic preconditioning in a coral-reef shark. Journal of Experimental Biology 207: 3131-3139.

Nurkse, K., Kotta, J., Orav-Kotta, H. and Ojaveer, H. 2016. A successful non-native predator, round goby, in the Baltic Sea: generalist feeding strategy, diverse diet and high prey consumption. Hydrobiologia 777: 271-281.

Paine, M.D., Dodson, J.J. and Power, G. 1982. Habitat and food resource partitioning among four species of darters (Percidae: *Etheostoma*) in a southern Ontario stream. Canadian Journal of Zoology 60: 1635-1641.

Pettitt-Wade, H., Wellband, K.W., Heath, D.D. and Fisk, A.T. 2015. Niche plasticity in invasive fishes in the Great Lakes. Biological Invasions 17: 2565-2580

Pejchar, L., and H.A. Mooney. 2009. Invasive species, ecosystem services and human well-being. Trends in Ecology and Evolution 24: 497-504.

Piller, K.R., Bart Jr, H.L. and Hurley, D.L. 2008. Phylogeography of the greenside darter complex, *Etheostoma blennioides* (Teleostomi: Percidae): A wide-ranging polytypic taxon. Molecular Phylogenetics and Evolution 46: 974-985.

Poos M, Dextrase AJ, Schwalb AN, Ackerman JD. 2010. Secondary invasion of the round goby into high diversity Great Lakes tributaries and species at risk hotspots: potential new concerns for endangered freshwater species. Biological Invasions.12:1269-84.

Pratt, A.E. and Lauer, T.E. 2013. Habitat use and separation among congeneric darter species. Transactions of the American Fisheries Society 142: 568-577.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Rahel, F.J. 2007. Biogeographic barriers, connectivity and homogenization of freshwater faunas: it's a small world after all. Freshwater Biology 52: 696-710.

Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D. and Betel, D. 2013. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome biology 14.

Ray, J.M., Lang, N.J., Wood, R.M. and Mayden, R.L. 2008. History repeated: recent and historical mitochondrial introgression between the current darter *Etheostoma uniporum* and rainbow darter *Etheostoma caeruleum* (Teleostei: Percidae). Journal of Fish Biology, 72: 418-434.

Ray, J.M., Wood, R.M. and Simons, A.M. 2006. Phylogeography and post-glacial colonization patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). Journal of Biogeography 33: 1550-1558.

Reash, R.J. and Berra, T.M. 1987. Comparison of fish communities in a clean-water stream and an adjacent polluted stream. American Midland Naturalist 301-322.

Reed, D.H. and Frankham, R., 2003. Correlation between fitness and genetic diversity. Conservation biology 17: 230-237.

Richards, C.L., Bossdorf, O., Muth, N.Z., Gurevitch, J. and Pigliucci, M. 2006. Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. Ecology letters 9: 981-993.

Richards, J.G. 2011. Physiological, behavioral and biochemical adaptations of intertidal fishes to hypoxia. Journal of Experimental Biology 214: 191-199.

Scott, W.B. and E. J. Crossman. 1973. Freshwater fishes of Canada. Bulletin 184, Fisheries Research Board of Canada, Ottawa.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270: 467.

Sheng, Y., Hua, Z.Y., Yang, Z., Wei, X.L., Sheng, Y.J., Jia, H.L. and Jun, Q. 2019. Effects of acute hypoxic stress on biochemical parameters, immune regulation and metabolic capacity of the blood in genetically improved farmed tilapia (GIFT, *Oreochromis niloticus*). Journal of Applied Ichthyology 35: 978-986.

Singh, D., Nath, K., Trivedi, S.P. and Sharma, Y.K. 2008. Impact of copper on haematological profile of freshwater fish, *Channa punctatus*. Journal of Environmental Biology 29: 253

Slarkin, M. 1985. Gene flow in natural populations. Annual review of ecology and systematics, 16: 393-430.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 23: 787-792.

Smith, A.D., Houde, A.L.S., Neff, B. and Peres-Neto, P.R. 2017. Effects of competition on fitness-related traits. Oecologia, 183: 701-713.

Stepien, CA, Karsiotis, S.I. Sullivan, T.J. and Klymus, K.E. 2017. Population genetic structure and comparative diversity of smallmouth bass *Micropterus dolomieu*: congruent patterns from two genomes. Journal of Fish Biology 90: 2125-2147.

Stepien, C.A., Murphy, D.J., Lohner, R.N., Sepulveda-Villet, O.J. and Haponski, A.E. 2009. Signatures of vicariance, postglacial dispersal and spawning philopatry: population genetics of the walleye *Sander vitreus*. Molecular Ecology 18: 3411-3428.

Stepien, C. A., D. J. Murphy, and R. M. Strange. 2007. Broad-to fine-scale population genetic patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great Lakes and beyond: an interplay of behaviour and geography. Molecular Ecology 16:1605–1624.

Stepien, C.A. and Tumeo, M.A. 2006. Invasion genetics of Ponto-Caspian gobies in the Great Lakes: a 'cryptic'species, absence of founder effects, and comparative risk analysis. Biological Invasions 8: 61-78.

Strayer, D.L., 2009. Twenty years of zebra mussels: lessons from the mollusk that made headlines. Frontiers in Ecology and the Environment 7: 135-141.

Strayer DL. 2010 Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. Freshwater biology 55:152-74.

Sullam, K. E., Essinger, S. D., Lozupone, C. A., O'Conner, M. P., Rosen, G. L., Knight, R. O. B., ... & Russell, J. A. 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. Molecular ecology 21: 3363-3378.

Tang, M. and Boisclair, D.,1995. Relationship between respiration rate of juvenile brook trout (*Salvelinus fontinalis*), water temperature, and swimming characteristics. Canadian Journal of Fisheries and Aquatic Sciences 52: 2138-2145.

Thorlacius, M., Hellström, G. and Brodin, T. 2015. Behavioral dependent dispersal in the invasive round goby *Neogobius melanostomus* depends on population age. Current Zoology 61: 529-542.

Tibbets, C. A., & Dowling, T. E. 1996. Effects of intrinsic and extrinsic factors on population fragmentation in three North American minnows (Teleostei: Cyprinidae). Evolution 50: 1280-1292.

Todd, E.V., Black, M.A. and Gemmell, N.J. 2016. The power and promise of RNA-seq in ecology and evolution. Molecular ecology 25: 1224-1241.

Turner, T.F. and Trexler, J.C. 1998. Ecological and historical associations of gene flow in darters (Teleostei: Percidae). Evolution 52: 1781-1801.

Vander Zanden, M.J, G.J.A. Hansen, S.N. Higgins, and M.S.Kornis. 2010. A pound of prevention, plus a pound of cure: early detection and eradication of invasive species in the Laurentian Great Lakes. Journal of Great Lakes Research 36: 199-205.

Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M. and Morgan, R.P. 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24: 706-723.

Walter, W., Sánchez-Cabo, F. and Ricote, M. 2015. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics 31: 2912-2914.

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Genetics 10: 57-63.

Wang, R.L., Bencic, D.C., Garcia-Reyero, N., Perkins, E.J., Villeneuve, D.L., Ankley, G.T. and Biales, A.D. 2014. Natural variation in fish transcriptomes: Comparative analysis of the fathead minnow (*Pimephales promelas*) and zebrafish (*Danio rerio*). PloS one e114178.

Warren, I.A., Ciborowski, K.L., Casadei, E., Hazlerigg, D.G., Martin, S., Jordan, W.C. and Sumner, S. 2014. Extensive local gene duplication and functional divergence among paralogs in Atlantic salmon. Genome Biology and Evolution 6: 1790-1805

Westbrook, L.J. 2012. Genetic Structure of Rock Bass and Johnny Darters: Implications for Gamefish Management in Wisconsin. Master's Thesis.

Welker, T.L., Mcnulty, S.T. and Klesius, P.H. 2007. Effect of sublethal hypoxia on the immune response and susceptibility of channel catfish, *Ictalurus punctatus*, to enteric septicemia. Journal of the World Aquaculture Society 38: 12-23

Whitehead, A., & Crawford, D. L. 2005. Variation in tissue-specific gene expression among natural populations. Genome biology 6: R13.

Whitehead, A. and Crawford, D.L. 2006. Neutral and adaptive variation in gene expression. Proceedings of the National Academy of Sciences 103: 5425-5430.

Wittkopp, P.J., Haerum, B.K. and Clark, A.G. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430: 85.

Wright, A.J., Wardle, D.A., Callaway, R. and Gaxiola, A. 2017. The overlooked role of facilitation in biodiversity experiments. Trends in Ecology and Evolution 32: 383-390.

Wu, X., Zhang, Y., Xu, S., Chang, Y., Ye, Y., Guo, A., Kang, Y., Guo, H., Xu, H., Chen, L. and Zhao, X. 2019. Loss of Gsdf leads to a dysregulation of Igf2bp3-mediated oocyte development in medaka. General and Comparative Endocrinology 277: 122-129.

Xu, Q., Zhu, C., Fan, Y., Song, Z., Xing, S., Liu, W., Yan, J. and Sang, T. 2016. Scientific Reports 6: 25536.

Zhao, Y., Wang, J., Thammaratsuntorn, J., Wu, J.W., Wei, J., Wang, Y., Xu, J. and Zhao, J. 2015. Comparative transcriptome analysis of Nile tilapia (*Oreochromis niloticus*) in response to alkalinity stress. Genetics and Molecular Research 14: 17916-26.

ABSTRACT

EVOLUTIONARY ECOLOGY OF THE INVASIVE ROUND GOBY (NEOGOBIUS MELANOSTOMUS) AND NATIVE JOHNNY DARTER (ETHEOSTOMA NIGRUM: A GENOMIC PERSPECTIVE

by

ABBY J. WICKS

May 2020

Advisor: Dr. Thomas Dowling

Major: Biological Sciences

Degree: Doctor of Philosophy

Round goby (*Neogobius melanostomus*) is an invasive fish present in all five Great Lakes and is becoming increasingly common in their tributaries. Johnny darter (*Etheostoma nigrum*) is a native species that often coexists with round goby. Here, I use traditional population genetic methods and genomic tools to explore the evolutionary ecology of these species. First, historic factors are addressed as a source of variation in study populations by characterizing patterns of mitochondrial DNA variation throughout Lower Michigan. Round goby populations were largely homogenous and exhibited no evidence of overarching historical genetic structure, consistent with the recent invasion and rapid expansion of this species. Johnny darter exhibited significant differentiation among local populations; however, similarity among mtDNA haplotypes indicated that differences among populations are recent. Therefore, historical factors should have limited influence on patterns of gene expression in study populations.

Response to environmental heterogeneity was assessed by comparing gene expression between these species. Differential expression analysis was applied to fishes sampled from two southeastern Michigan drainages, the Rouge and Clinton rivers, and gene expression profiles

between the two drainages were found to be distinct for these species. Pairwise comparisons among sampling sites and years indicated that there were more differentially expressed genes in round goby than Johnny darter, suggesting that Johnny darter has a more specialized strategy while round goby exhibits more plasticity. Plasticity is often cited as a trait common to successful invaders and the ability to respond to a range of conditions may contribute to the rapid range expansion of round goby. Finally, co-expression analysis was used to identify groups of genes correlated with specific environmental factors for each species. The two species differ markedly in patterns of gene expression correlated to stressors, possibly reflecting round goby's adaption to varying environmental conditions. An ability to more readily respond to environmental stressors may make round goby well-suited to degraded habitats, which are common sites of early invasion.

AUTOBIOGRAPHICAL STATEMENT

I was raised in rural west Michigan and grew up exploring the outdoors, always gravitating toward science in hobbies and in school. As a young college student at Michigan State University I never saw myself in a science career and instead majored in Communication Arts and Sciences, graduating in 2007. After working in the insurance industry for a few years, volunteer experiences with invasive species removal and stream monitoring projects inspired me to return to college for a degree in Environmental Sciences at The University of Michigan-Dearborn. There I had the opportunity to work in a conservation genetics lab and gained an appreciation for the questions that could be addressed using genomic tools, leading me to focus my graduate interests in molecular ecology. My graduate research introduced me to fishes as study species. I have discovered that, not only are fishes beautiful and fascinating, but fishes are also ideal for studying the interaction of species and their environment. I plan to continue with research that focuses on conserving biodiversity as well as exploring the evolutionary relationships among species and their environment.