

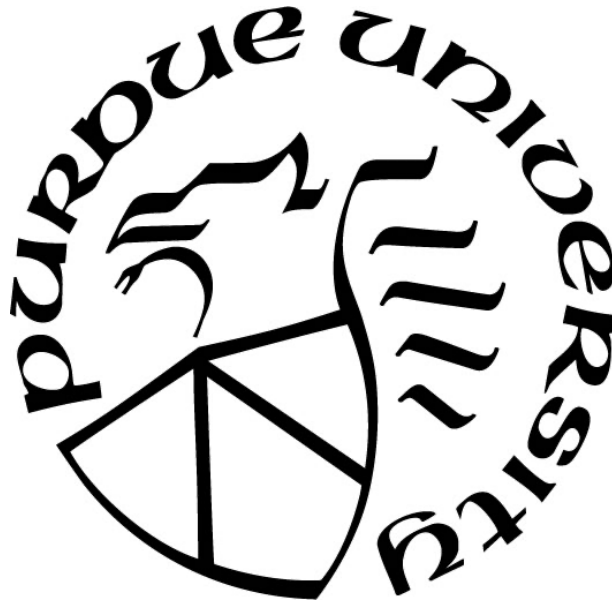
**POPULATION CONNECTIVITY IN LAKE MICHIGAN
YELLOW PERCH**

by
Claire E. Schraidt

A Thesis

*Submitted to the Faculty of Purdue University
In Partial Fulfillment of the Requirements for the degree of*

Master of Science



Department of Forestry and Natural Resources
West Lafayette, Indiana
December 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Mark Christie, Chair

Department of Forestry and Natural Resources

Department of Biological Sciences

Dr. Tomas Höök

Department of Forestry and Natural Resources

Dr. Paris Collingsworth

Department of Forestry and Natural Resources

Approved by:

Dr. Robert G. Wagner

Dedicated to my parents.

Thank you for supporting every foolish venture.

...

And to Gary, obviously.

ACKNOWLEDGMENTS

I'd like to first thank my committee members, Tomas Höök and Paris Collingsworth, for their feedback and insight on this project. Additionally, I would like to express my gratitude to the state agencies assisted in the monumental task of sampling across Lake Michigan. To the individuals at the Indiana Department of Natural Resources, the Michigan Department of Natural Resources, the Wisconsin Department of Natural Resources, the Illinois Natural History Survey, and the Grand Traverse Band of Ottawa and Chippewa—thank you. This would not have been possible without your help.

I would also like to express my gratitude to my advisor, Mark Christie. Thank you for giving me a chance and supporting my growth as both a person and a scientist. I appreciate your mentorship and that you always treated me as both a scientist and a human being in equal measure. I feel so incredibly lucky to have landed in your lab, and no words will ever adequately express my appreciation for both you, and every person in the Christie Lab. I would not have made it without you all.

TABLE OF CONTENTS

LIST OF TABLES.....	6
LIST OF FIGURES	7
ABSTRACT.....	8
CHAPTER 1. POPULATION CONNECTIVITY IN LAKE MICHIGAN YELLOW PERCH.	9
1.1 Introduction.....	9
1.2 Materials and Methods.....	12
1.2.1 Sample Collection.....	12
1.2.2 DNA extraction.....	12
1.2.3 RAD library prep and sequencing	12
1.2.4 Read processing and SNP filtering.....	13
1.2.5 Summary statistics	14
1.2.6 Identifying population structure	14
1.2.7 Isolation by distance	15
1.2.8 Oceanographic Distance	15
1.3 Results.....	17
1.3.1 Sequencing and genotyping.....	17
1.3.2 Summary statistics	17
1.3.3 Identifying population structure	17
1.3.4 Isolation by distance	18
1.4 Discussion.....	18
1.5 References.....	20
1.6 Figures and Tables	29

LIST OF TABLES

Table 1.1. Summary statistics for all populations ($n = 26$). Number of individuals sampled (N) and summary statistics including observed heterozygosity (H_o), expected heterozygosity (H_e), allelic richness, (Ar), and % missing data.....	30
Table 1.3. Pairwise F_{ST} 's across all populations. Values that did not differ from zero based on 95% CI's were excluded.....	36
Table 1.3. A summary of relevant FVCOM models utilized in connectivity matrices.....	38

LIST OF FIGURES

Figure 1.1. (a) A map of sample locations across Lake Michigan including Green Bay (orange) and the main basin (blue). (b) Principal Coordinates Analysis showing relatedness between individual populations based on pairwise F_{ST} . (c) Results of individuals clustering analysis with STRUCTURE ($K = 2$). Colored lines correspond to an individual's estimated proportion membership to each cluster, and individuals are grouped by population (bold black lines). 29

Figure 1.2. (a) Isolation by distance (IBD) plot of F_{ST} values against geographic (Euclidean) distance among Green Bay populations. (b) Regression analysis of proportion Green Bay ancestry and distance from the mouth of Green Bay. 34

Figure 1.3. (a) Isolation by distance (IBD) plot of F_{ST} values against geographic (Euclidean) distance among main basin populations. (b) Regression analysis of genetic distance (F_{ST}) and dispersal likelihood, showing decreased genetic distance with increased probability of dispersal. 34

ABSTRACT

Yellow perch (*Perca flavescens*) are an ecologically and economically important species in the Great Lakes. In Lake Michigan, there is substantial interest in restoring Yellow perch to their historic abundance. Populations in Lake Michigan began to decline dramatically in the late 1980s and, despite management efforts, have not rebounded. Understanding stock structure is imperative for implementing successful management strategies, but assessing population structure in yellow perch is particularly challenging. Yellow perch disperse on surface currents during a 30-40 day larval period where dispersal is difficult to observe directly and varies interannually. In order to better understand yellow perch population structure and connectivity, we sequenced 960 individuals, collected at 20 sampling locations across Lake Michigan using restriction site-associated DNA sequencing (RADseq). We used 3337 single nucleotide polymorphisms (SNPs) to observe genetic differences between populations and paired these findings with a Lagrangian particle tracking model to explain patterns of connectivity and gene flow. We showed that Green Bay and the main basin of Lake Michigan are genetically distinct populations. Within these two genetic groups, drivers for population structure appear to be very different. Green Bay shows distinct populations across its 150 kilometers, consistent with isolation by distance. These populations show lower global allelic richness and heterozygosity than the main basin. In contrast, the main basin shows low but significant genetic distance (measured as pairwise F_{ST}) and higher allelic richness and observed heterozygosity, appearing to behave more like a marine system. To validate this observation, we paired these findings with a Lagrangian particle tracking model to explain patterns of connectivity and gene flow and found that distances derived from these particle tracking models were significantly correlated with the genetic distances observed between main basin populations.

CHAPTER 1. POPULATION CONNECTIVITY IN LAKE MICHIGAN YELLOW PERCH

1.1 Introduction

Understanding patterns of population connectivity is imperative for implementing successful species conservation and management strategies (Allendorf et al., 2010; Fraser & Bernatchez, 2001). In many fishes, dispersal plays a fundamental role in structuring populations, but quantifying the direction and magnitude of dispersal can be challenging (Hedgecock et al., 2007). Many freshwater fishes and over 95% of all marine fishes have a pelagic larval stage (Nelson 2006). Larval fish can disperse vast distances—sometimes hundreds of kilometers from where they’ve spawned—on surface currents (Cowen & Sponaugle, 2009). Tracking these larvae, which are often small and transparent, is not feasible using traditional sampling methods (Bradbury & Snelgrove, 2001; Levin, 2006) and, because larval transport cannot be observed *in situ*, integrative approaches are necessary to reconstruct how individuals move and interact across space and time (Burgess et al., 2014; Weersing & Toonen, 2009). Genomic tools allow us to look at patterns of connectivity post-dispersal and identify gene flow among populations. In doing so, we can better define management units and develop informed strategies for population recovery and persistence (Bernatchez et al., 2017; Johnson et al., 2018). However, while genomic tools can reveal discrete population structure, it cannot explain what drives population structure. In order to relate pattern to process, researchers pair other analytical approaches with genetics (Selkoe et al., 2016). These multi-pronged approaches are common in marine systems and include otolith microchemistry (Collins et al., 2013; Feyrer et al., 2007; Hoey et al., 2020), investigating boundaries based on environmental features such as depth, turbidity, and temperature (Johansson et al., 2015; Lawlor & Arellano, 2020; Roy et al., 2012), and modeling currents to estimate connectivity (Alberto et al., 2011; White et al., 2010). While the Great Lakes are a freshwater system, in many ways, it can behave similarly to a marine environment. Many fishes in the Great Lakes also undergo a larval dispersal stage. Moreover, due to their size, Great Lakes’ physical processes occur on spatial and temporal scales more comparable to the coastal ocean (Beletsky et al., 1999), making it an appealing system for applying the aforementioned techniques.

Yellow perch (*Perca flavescens*) are an ecologically and economically important Great Lakes fish. Ecologically, they play an important role in nutrient cycling in the near-shore environment. They are an abundant near-shore species in many areas of the Great Lakes, serve as important predators of small fishes and invertebrates, and are an important prey item for species such as walleye (*Sander vitreus*), northern pike (*Esox Lucius*) and lake trout (*Salvelinus namaycush*) (Evans, 1986). Historically, yellow perch supported a lucrative commercial fishery in Lake Michigan with catches reaching 1.1 million kg annually. Yellow perch have also been a popular fish for recreational anglers—at the peak of this fishery, yellow perch comprised 85% of recreational harvest (Francis et al., 1996). Yellow perch populations began to decline in the late 1980's and early 1990's, which prompted the complete closure of commercial fisheries save for Green Bay and Grand Traverse Bay, reduced bag limits and seasonal spawning closures (Clapp & Dettmers, 2004). Despite these management efforts yellow perch populations have not rebounded. Annual surveys conducted across management jurisdictions and reported by the Lake Michigan yellow perch task force show that the current adult abundance remains well below historically observed abundance in all areas of the lake (GLFC Lake Michigan Committee, 2018). The cause of yellow perch decline in Lake Michigan is not fully understood and likely has many contributing factors including invasive dreissenid mussels filtering out nearshore nutrients (Goto et al., 2020; Hecky et al., 2004; Mida et al., 2010), competition with and predation by invasive alewife (*Alosa pseudoharengus*) (Evans, 1986; Janssen & Luebke, 2004), and low adult abundance (i.e., spawning stock biomass) following overexploitation (Wilberg et al., 2005). Annual surveys of yellow perch show high variability in recruitment and year class strength, and low production across year and region (GLFC Lake Michigan Committee, 2018). Studies suggest that recruitment might be limited by declining zooplankton abundance (Dettmers et al., 2003), larval predation (Madenjian et al., 2008), and adult abundance (Clapp & Dettmers, 2004; Wilberg et al., 2005). However, nearshore dynamics might only partially explain interannual variability in yellow perch recruitment.

Similar to many marine fishes, yellow perch have a pelagic larval stage that lasts between thirty and forty days (Whiteside et al., 1985). Especially during their early development, yellow perch are poor swimmers (Dettmers et al., 2005) and are likely to be transported by surface currents. The dynamic circulation patterns found in Lake Michigan coupled with the yellow perch pelagic larval duration, likely influences recruitment success. Previous studies have used biophysical models to predict the extent of larval dispersal throughout Lake Michigan. Beletksy et

al., (2007) suggested that summer current patterns could result in larval yellow perch dispersing from western to eastern shores but are unlikely to result in movement between the southern and northern basins. Because Lake Michigan sits on a steep latitudinal gradient (494 km north to south), lack of mixing could lead to genetically distinct, locally adapted populations. Additionally, Lake Michigan's two large bays, Green Bay and Grand Traverse Bay, (Figure 1.1a), often experience their own, weaker, regional currents and mix very little with the main basin (Beletsky & Schwab, 2008) and this could present a significant barrier to gene flow.

Previous genetic studies have suggested that the yellow perch in the Great Lakes share an evolutionary history stemming from a single colonization event, with most populations across the Great Lakes sharing a single mtDNA haplotype (Sepulveda-Villet et al., 2009). However, additional studies have shown complex population structure in yellow perch in Lake Erie (Sepulveda-Villet & Stepien, 2011) and there is no reason to anticipate population dynamics in Lake Michigan would be any less complex. A microsatellite study on yellow perch in Lake Michigan (Miller, 2003) revealed moderate genetic structure between populations in Green Bay and the southern region of the main basin. These populations were found to be genetically similar, which matches predictions made with Beletsky et al.'s (2007) biophysical model. However, this study did not include populations from the northern portion of Lake Michigan. Without sampling geographically representative populations throughout the entire basin, it is challenging to fully understand how dispersal might affect gene flow across the entirety of the lake.

Here, we build on previous work by sequencing 960 yellow perch from 20 geographically representative locations across Lake Michigan (Figure 1.1a). Using restriction site associated DNA sequencing (RADseq) we identified population structure across the lake and paired these results with Lagrangian particle tracking model. By coupling genetic data with biophysical processes we were able to better understand how lake currents influence larval dispersal patterns, and more broadly, population connectivity across the lake.

1.2 Materials and Methods

1.2.1 Sample Collection

Yellow perch were collected from 20 sites circumscribing Lake Michigan and Green Bay in 2018 and 2019 (Figure 1.1a, Table 1).

In 2018, adult individuals, defined as year one or older ($>100\text{mm}$), were sampled using 12 hour overnight multi-mesh gill net sets. Net location and depth varied by site and were informed by agency correspondence and local fishing reports. Sampling during the 2018 season began in southern Lake Michigan in early March and continued through the end of July at the northernmost sites in order to correspond with regional spawning times. Young of year ($<100\text{ mm}$) were sampled in September using beach seines. The 2019 sampling regime was greatly assistance by several agencies—the Michigan Department of Natural Resources, the Indiana Department of Natural Resources, the Wisconsin Department of Natural Resources, the Illinois Natural History Survey, and the Grand Traverse Band of Ottawa and Chippewa. Agency sampling methods consisted of bottom trawls, gill netting, and creel surveys, and took place across both 2018 and 2019. 1376 individuals were sampled over two sampling seasons, and a subset of 960 were selected for sequencing (Table 1.1). Tissues were stored at -4°C upon arrival at Purdue University and held until DNA extraction.

1.2.2 DNA extraction

DNA was isolated from fin tissue with Qiagen DNeasy® Blood & Tissue Kits. Following an overnight (~ 14 hour) tissue digestion with Proteinase K incubated at 56°C , DNA was extracted in plates using standard kit protocols, and eluted from the silica membrane with $200\mu\text{L}$ Tris Low-EDTA buffer. Extractions were quantified using a Quant-it™ PicoGreen® dsDNA Assay (Invitrogen, Waltham, MA), and DNA was normalized to a quantity of 200ng or approximately $20\text{ng}/\mu\text{L}$.

1.2.3 RAD library prep and sequencing

Libraries for restriction site-associated DNA (RAD) sequencing were prepared following the BestRAD protocol (Ali et al., 2016). Normalized DNA was digested with the restriction enzyme

SbfI followed by ligation with indexed adaptors. Barcoded libraries were pooled into master libraries of 96 individuals and fragmented to ~300-500bp with 12 30s cycles in a Q500 sonicator (Qsonica, Newtown, CT). Fragmented DNA was bound to Dynabeads™ M-280 Streptavidin magnetic beads (Invitrogen) and washed with buffer to remove non-target fragments. Following purification with AMPure XP beads (Beckman Coulter, Brea, CA), master libraries were input into the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® at the End Prep step for ligation of master library barcodes, a 250-bp insert size-selection, and a 12-cycle PCR enrichment. Successful size-selection and enrichment were confirmed with visualization of products on a 2% agarose E-Gel (Invitrogen). Products underwent a final AMPure XP purification clean-up followed by quantification with a Qubit® 2.0 Fluorometer. A total of 10 master libraries, each containing 96 individually barcoded samples, were sent to Novogene (Sacramento, CA) for PE150 sequencing on one lane of the Illumina NovaseqS4 platform.

1.2.4 Read processing and SNP filtering

Raw Illumina RAD sequence reads were processed using the STACKS v2.54 (Rochette et al., 2019) software pipeline. Reads were cleaned and demultiplexed by barcode using the STACKS subprogram *process_radtags*. Sequences were demultiplexed by barcode, filtered for illumina quality score and enzyme cut-site, and trimmed to 140 base pairs to reduce tail-end sequencing errors (parameter flags = --filter_illumina, --bestrad, -t 140). The resulting filtered, individually assigned reads were aligned to the yellow perch reference genome (*P. flavescens* PFLA_1.0 assembly, GenBank accession GCA_004354835.1; Feron et al., 2020) with bowtie2 (Langmead et al., 2019, Langmead & Salzberg, 2012) (parameter flag = --very-sensitive). SNPS were called from reference aligned paired end reads with the STACKS subprogram *gstacks* (parameter flag = --rm-unpaired reads) and individuals were genotyped at each identified SNP. The *gstacks* output files, which contain consensus sequences at each loci identified as well as individual genotyping data, were filtered through STACKS subprogram *populations*. SNPS that genotyped in less than 30% of individuals were discarded (parameter flags = -r 0.3) and results were exported in variant call format (VCF). In a final filtering step, SNPs were filtered using vcftools v0.1.9 (Danecek et al., 2011). Filtering consisted of first removing SNPs that genotyped in fewer than 80% of individuals, then removing outlier SNPs with regard to mean read depth across individuals (--min-meanDP 10, --max-meanDP 60) as recommended in O'Leary et al., (2018). In the final filtering

step, individuals with > 50% missing data were identified and removed (mean % missingness was 7.7%). The resulting vcf file was converted to GENEPOP and STRUCTURE format for downstream analysis using PGDSpider (Lischer & Excoffier, 2012).

1.2.5 Summary statistics

Population level summary statistics including observed heterozygosity (H_o), expected heterozygosity (H_e), and allelic richness (Ar) were calculated using the in the R package *hierfstat* v0.04.22 (Goudet, 2005), which returns loci and population level statistics using equations as defined in Nei, 1987. A Bartlett test of homogeneity of variances and a paired t-test between H_o and H_e was performed on each population to determine whether or not observed heterozygosity significantly differed from expected heterozygosity. ANOVAs were used to determine whether global genetic diversity estimates (H_o and Ar) between Green Bay and the Main basin statistically differed from one another.

1.2.6 Identifying population structure

A Bayesian clustering method was implemented in STRUCTURE v2.3.4 (Pritchard et al., 2000) to determine population structure present among all populations. STRUCTURE was also used to examine potential fine-scale structure in the main basin, as the populations in the main basin did not follow any discernible pattern of isolation by distance, nor show any demonstrable structure in either principal components or principal coordinates analysis. For STRUCTURE analysis with all 26 populations, optimal K was determined using the delta K method (Evanno et al., 2005). Runs consisted of an initial burn in period of 50,000 Markov Chain Monte Carlo (MCMC) iterations followed by 50,000 iterations for each inferred cluster. Analyses were performed with K = 1-30 clusters and replicated five times for each K. For main basin STRUCTURE analyses, we employed admixture and correlated allele frequency models, as they are considered to be most appropriate when subtle population structure is expected (Falush et al., 2003; Hubisz et al., 2009). Analyses were performed for K = 1-25 and repeated five times for each K.

Pairwise F_{ST} between populations and 95% confidence intervals were calculated in *hierfstat* v0.5.7 (Goudet, 2005). Non-zero pairwise F_{ST} 's based on confidence intervals were reported in Table 1.2 and exported for analysis in GenAlEx v6.5 (Peakall & Smouse, 2006, 2012)

where principal coordinate analysis was performed for the in order to visualize structure among populations (Figure 1.1b).

1.2.7 Isolation by distance

Pairwise F_{ST} and Euclidean distances (defined as the straight-line distance between two geographic points) was compared using non parametric permutational testing (Smouse & Peakall, 1999). Because both genetic and geographic distance matrices are not independent, the assumptions of traditional linear regressions can be violated. Instead, we shuffled the y matrix (genetic distance) and assessed whether the relationship between the true matrices (R_{xy}) varies significantly from the relationship between random shuffles. A total of 999 randomization permutations were performed in order to estimate what values would be anticipated if the matrices were unrelated and these values were used to calculate the modified P value (Peakall & Smouse, 2012).

1.2.8 Oceanographic Distance

For the biophysical model, we used a Lagrangian particle tracking model developed previously to study transport of larval cod (Churchill et al., 2011; Huret et al., 2007, <http://fvcom.smast.umassd.edu/>). Vertical mixing was implemented using turbulent diffusivity output from FVCOM and a random-walk scheme for spatially varying vertical diffusivity. Rowe et al. (2016) recently modified the code to include a vertical floating/sinking/swimming velocity (Gräwe, 2011). Particles were assumed to be neutrally buoyant in the preliminary simulation, but a vertical swimming velocity of 0.0003 was also included. The Lagrangian particle tracking simulations were forced by output from a FVCOM simulation of Lake Michigan-Huron developed by Anderson and Schwab (E. J. Anderson & Schwab, 2013) to simulate currents in the Straits of Mackinac. Horizontal grid resolution varied with finer resolution nearshore and in regions with complex coastlines (100 m in the Straits to 2.5 km in the center of the lakes), and each horizontal grid included 20 sigma layers in the vertical. The FVCOM is an unstructured grid, finite-volume, free surface, three-dimensional primitive equation ocean model that solves the momentum, continuity, temperature, salinity, and density equations (Chen et al., 2003). The unstructured grid of FVCOM conforms to complex coastline morphologies and allows for increased grid resolution in regions of interest. Turbulence closure was implemented through the MY-2.5 scheme for

vertical mixing (Galperin et al., 1988), and the Smagorinsky scheme for horizontal mixing (Smagorinsky, 1963). FVCOM has been implemented for the Great Lakes yielding accurate predictions of temperature, water levels, and currents (E. Anderson et al., 2010; E. J. Anderson et al., 2015; E. J. Anderson & Schwab, 2013; Bai et al., 2013) and is being used for NOAA's next-generation Great Lakes Operational Forecast System (GLOFS).

To generate a connectivity matrix, the probability of transport from region i to region f was calculated as N_{if}/N_i , where N_{if} is the number of particles initiated in region i that are within region f at the end of the simulation, and N_i is the total number of particles that were initiated in region i . Connectivity matrices were developed for six years (2014-2019). The sensitivity of the connectivity matrices to model assumptions can be evaluated by considering scenarios of, 1.) swimming behavior, and 2.) horizontal diffusion (Table 1.3). Simple behavior scenarios were tested, including passive particle, upward swimming, and downward swimming. Realistic estimates of swimming velocity were implemented as a deterministic vertical velocity in the vertical random walk turbulence scheme; thus representing the combined effects of turbulence and directed swimming. Scenarios including horizontal diffusion were tested, using literature values (Okubo, 1971), and using existing data from drifter deployments in Lake Michigan-Huron (www.nefsc.noaa.gov/drifter) resulting in a final horizontal diffusion coefficient of $5.6 \text{ m}^2/\text{s}$. These models were run for a larval duration of 50 days, with bi-weekly release dates ranging from late May to late July (Table 1.3).

Because some sampling sites fell within the same region utilized for estimating particle transport, the dynamics of transport between them could not be resolved. Thus, we selected one population from each region for analyses pairing oceanic and genetic distance.

Lagrangian particle transport models are inherently asymmetrical, allowing reciprocal exchange between populations where the probability of transport from point a to point b might have a different value than the probability of transport from point b to point a . To match our pairwise F_{ST} matrices, triangular distance matrices were created from asymmetrical, square matrix model outputs by averaging across the upper and lower probability matrices.

Models were generated for multiple release dates (May through July) from 2014 to 2019. Matrices based on mean triangular distance matrices were generated for each release date across

the six years for which the models were run. We tested the explanatory power of these models by regressing them with the genetic distance (pairwise F_{ST}) using the same statistical methods as the Euclidean isolation by distance analysis.

1.3 Results

1.3.1 Sequencing and genotyping

960 individuals were RAD sequenced, producing over 5 billion reads that resulted in an average of 5,463,100 paired-end reads per sample. Following filtering, 935 individuals represent from 26 sampling sites, year, and age class were genotyped at 3337 loci (Table .1). Mean read depth of loci across individuals was 29x and mean missingness per individual was 7.7%.

1.3.2 Summary statistics

The estimates of global genetic diversity for heterozygosity and allelic richness were $H_o = 0.449$, $H_e = 0.451$, and $Ar = 1.958$, respectively. Genetic diversity estimates were also calculated for each population (Table 1). In order to better understand processes driving genetic differentiation between Green Bay and the main basin, an ANOVA was run to determine if H_o and Ar varied significantly between the main basin and Green Bay. H_o was not found to be significantly different between groups, ($p = 0.06$). However, allelic richness was significantly different between Green Bay and the main basin ($p < 0.0001$), with values ranging from 1.490 – 1.939 for Green Bay and 1.959 – 1.973 for the main basin (Table 1).

1.3.3 Identifying population structure

Both mean likelihood values ($L(K)$) and ΔK suggested two clusters ($k=2$) across the entire basin. This clustering reveals a distinct genetic split between Green Bay populations and Main basin samples (Figure 1.1c). When the main basin samples were run separately to determine fine-scale population structure, analysis revealed no optimal K , suggesting minimal population structure in across Main basin samples.

Populations in Green Bay, particularly south and central, are largely isolated from the main basin. The only connection to the main basin is at the northern end of that bay, about 100 km from

the southernmost site sampled. There is little exchange of current into and out of the bay (Beletsky & Schwab, 2008), suggesting that larval recruits are unlikely to be exchanged between Green Bay and the main basin. The STRUCTURE results (Figure 1.1c) show a larger proportion of Green Bay ancestry (blue) in the southern and central populations, supporting the idea of decreasing exchange rates of individuals with increasing distance from the mouth of the bay. To quantify this, we regressed proportion Green Bay ancestry with distance from the mouth of Green Bay and found a significant linear relationship between proportion main basin ancestry and distance from the mouth (Figure 1.2b). Distance to mouth was calculated by using the straight-line distance between population coordinates and a point in the center of the mouth of Green Bay.

1.3.4 Isolation by distance

There is a strong linear relationship between F_{ST} and Euclidean distances in the Green Bay dataset ($R^2 = 0.6213$, Mantel $r = 0.788$, $p = 0.004$). Inversely, the Main basin dataset showed no relationship between F_{ST} and Euclidean distance ($R^2 = 0.0009$, Mantel $r = 0.030$, $p = 0.363$). However, a reduced, linearized F_{ST} matrix did correlate with the probability matrix derived from oceanographic current ($R_{xy} = -0.499$, $p = 0.014$). The model averaged from 2014 - 2019 for the mid-June release dates had a significant relationship to the genetic connectivity matrix for the main basin, showing the potential for oceanographic data to resolve fine scale patterns of population connectivity (Figure 1.3b).

1.4 Discussion

Comparisons of population structure show a distinct genetic split between Green Bay and the Main basin, supporting previous genetic analyses of yellow perch in Lake Michigan (Miller, 2003). When the Main basin and Green Bay are inspected separately, two different patterns of genetic differentiation emerge. Green Bay shows clear population differentiation between sites that fit a traditional model of isolation by distance. Sites across Green Bay cluster spatially (Fig 1B), and vary little cross years (SGB18, SGB19) and age classes (BNDYO, BND19 and LBDYO, LBD19). This spatial clustering is supported by the strong linear relationship between genetic distance and Euclidean distance (Figure 1.2a). Alternatively, the Main basin population structure appears to be panmictic with genetic differentiation remaining relatively low between sites and no significant

relationship to Euclidean distance (Figure 1.3a). Overall, Green Bay showed a much stronger signal for population differentiation (global $F_{st} = 0.010$) compared to the Main basin (global $F_{st} = 0.0031$), over a much smaller geographic range.

Summer currents in Green Bay are relatively weak, particularly compared to Main basin currents (Beletsky & Schwab, 2008). A weak, anticyclonic gyre often forms in northern Green Bay, allowing for some mixing with the Main basin, while the south remains relatively isolated (Beletsky et al., 1999; Beletsky & Schwab, 2001). Adult yellow perch have modest home ranges and high spawning site fidelity (Glover et al., 2008), suggesting that they do not disperse after larval settlement. The patterns observed in Green Bay, i.e. strong genetic partitioning based on geographic distance in the absence of strong current, support the idea that larval dispersal is the primary driver of population connectivity in Lake Michigan. The increased admixture at northern Green Bay sites (quantified by proportion ancestry) track with the current models, which show slight mixing in the northern portion of Green Bay that does not reach the central and southern sites.

While Miller (2003) did not detect genetic differentiation between southern main basin populations, we were able to detect low but significant population structure in the main basin. This, coupled with high allelic richness (global $F_{st} = 0.0031$, $p = 0.004$; mean $A_r = 1.965$) is similar to patterns of genetic diversity in marine fishes (DeWoody & Avise, 2000; Martinez et al., 2018). While STRUCTURE was unable to detect population structure among Main basin populations, and traditional isolation by distance yielded no significant correlation, patterns of connectivity were correlated with derived oceanic distances. In this scenario, the ability of oceanographic distances to resolve genetic isolation by distance patterns demonstrates the value of considering dispersal in estimating population connectivity. Additionally, the juxtaposition between population structure in Green Bay and the Main basin underscore the necessity of considering current as a driver for population structure in species with pelagic larval dispersal stages in large lakes.

While our derived ocean distance explained some of the genetic differentiation amongst main basin populations, it is likely that myriad additional environmental factors play into population dynamics of yellow perch across Lake Michigan. At higher dispersal levels, it becomes

increasingly difficult to detect barriers of dispersal, arising from the system becoming more panmictic and populations less differentiated (Wang, 2013). Our models, which broadly estimate dispersal patterns across multiple years, likely do not capture fine scale environmental heterogeneity that could influence annual variance in recruitment, survival, and admixture—particularly on an annual basis (Nanninga et al., 2014; Riginos & Liggins, 2013). Moreover, along Lake Michigan’s steep longitudinal gradient, the environment varies quite dramatically. Water temperature, substrate composition, anthropogenic disturbance, and many other environmental factors may influence local adaptation. Studies have shown that, despite near panmixia, many marine species show signatures of local adaptation across the species range (Clarke et al., 2010; Hoey & Pinsky, 2018; Wilder et al., 2020). While SNPs used in this analysis were putatively neutral, identifying adaptive loci in this dataset could uncover further population differentiation.

Our study provides the first comprehensive, lake-wide investigation of yellow perch in Lake Michigan. We identified two regionally differentiated populations, with unique drivers for population structure: isolation by distance, and isolation by environment. This study provides the foundation for better understanding the ecology and evolution of yellow perch in Lake Michigan, which is vital for meeting conservation management goals, and rebuilding declining stocks.

1.5 References

- Alberto, F., Raimondi, P. T., Reed, D. C., Watson, J. R., Siegel, D. A., Mitarai, S., Coelho, N., & Serrão, E. A. (2011). Isolation by oceanographic distance explains genetic structure for *Macrocystis pyrifera* in the Santa Barbara Channel. *Molecular Ecology*, 20(12), 2543–2554. <https://doi.org/10.1111/j.1365-294X.2011.05117.x>
- Ali, O. A., O’Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD Capture (Rapture): Flexible and Efficient Sequence-Based Genotyping. *Genetics*, 202(2), 389–400. <https://doi.org/10.1534/genetics.115.183665>
- Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11(10), 697–709. <https://doi.org/10.1038/nrg2844>

- Anderson, E. J., Bechle, A. J., Wu, C. H., Schwab, D. J., Mann, G. E., & Lombardy, K. A. (2015). Reconstruction of a meteotsunami in Lake Erie on 27 May 2012: Roles of atmospheric conditions on hydrodynamic response in enclosed basins. *Journal of Geophysical Research: Oceans*, 120(12), 8020–8038. <https://doi.org/10.1002/2015JC010883>
- Anderson, E. J., & Schwab, D. J. (2013). Predicting the oscillating bi-directional exchange flow in the Straits of Mackinac. *Journal of Great Lakes Research*, 39(4), 663–671. <https://doi.org/10.1016/j.jglr.2013.09.001>
- Anderson, E., Schwab, D., & Lang, G. (2010). Real-Time Hydraulic and Hydrodynamic Model of the St. Clair River, Lake St. Clair, Detroit River System. *Journal of Hydraulic Engineering-Asce - J HYDRAUL ENG-ASCE*, 136. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000203](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000203)
- Bai, X., Wang, J., Schwab, D. J., Yang, Y., Luo, L., Leshkevich, G. A., & Liu, S. (2013). Modeling 1993–2008 climatology of seasonal general circulation and thermal structure in the Great Lakes using FVCOM. *Ocean Modelling*, 65, 40–63. <https://doi.org/10.1016/j.ocemod.2013.02.003>
- Beletsky, D., Saylor, J. H., & Schwab, D. J. (1999). Mean Circulation in the Great Lakes. *Journal of Great Lakes Research*, 25(1), 78–93. [https://doi.org/10.1016/S0380-1330\(99\)70718-5](https://doi.org/10.1016/S0380-1330(99)70718-5)
- Beletsky, D., & Schwab, D. (2008). Climatological circulation in Lake Michigan. *Geophysical Research Letters*, 35(21). <https://doi.org/10.1029/2008GL035773>
- Beletsky, D., & Schwab, D. J. (2001). Modeling circulation and thermal structure in Lake Michigan: Annual cycle and interannual variability. *Journal of Geophysical Research: Oceans*, 106(C9), 19745–19771. <https://doi.org/10.1029/2000JC000691>
- Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E., Martinsohn, J. T., Miller, K. M., Naish, K. A., Ovenden, J. R., Primmer, C. R., Young Suk, H., Therkildsen, N. O., & Withler, R. E. (2017). Harnessing the Power of Genomics to Secure the Future of Seafood. *Trends in Ecology & Evolution*, 32(9), 665–680. <https://doi.org/10.1016/j.tree.2017.06.010>
- Bradbury, I. R., & Snelgrove, P. V. R. (2001). Contrasting larval transport in demersal fish and benthic invertebrates: The roles of behaviour and advective processes in determining spatial pattern. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(4), 641–644. <https://doi.org/10.1139/f01-031>

- Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A. K., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W., & Botsford, L. W. (2014). Beyond connectivity: How empirical methods can quantify population persistence to improve marine protected-area design. *Ecological Applications*, 24(2), 257–270. <https://doi.org/10.1890/13-0710.1>
- Chen, C., Liu, H., & Beardsley, R. C. (2003). An Unstructured Grid, Finite-Volume, Three-Dimensional, Primitive Equations Ocean Model: Application to Coastal Ocean and Estuaries. *Journal of Atmospheric and Oceanic Technology*, 20(1), 159–186. [https://doi.org/10.1175/1520-0426\(2003\)020<0159:AUGFVT>2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2)
- Churchill, J. H., Runge, J., & Chen, C. (2011). Processes controlling retention of spring-spawned Atlantic cod (*Gadus morhua*) in the western Gulf of Maine and their relationship to an index of recruitment success. *Fisheries Oceanography*, 20(1), 32–46. <https://doi.org/10.1111/j.1365-2419.2010.00563.x>
- Clapp, D. F., & Dettmers, J. M. (2004). Yellow Perch Research and Management in Lake Michigan: Evaluating Progress in a Cooperative Effort, 1997–2001. *Fisheries*, 29(11), 11–19. [https://doi.org/10.1577/1548-8446\(2004\)29\[11:YPRAMI\]2.0.CO;2](https://doi.org/10.1577/1548-8446(2004)29[11:YPRAMI]2.0.CO;2)
- Clarke, L. M., Munch, S. B., Thorrold, S. R., & Conover, D. O. (2010). High connectivity among locally adapted populations of a marine fish (*Menidia menidia*). *Ecology*, 91(12), 3526–3537. <https://doi.org/10.1890/09-0548.1>
- Collins, S. M., Bickford, N., McIntyre, P. B., Coulon, A., Ulseth, A. J., Taphorn, D. C., & Flecker, A. S. (2013). Population Structure of a Neotropical Migratory Fish: Contrasting Perspectives from Genetics and Otolith Microchemistry. *Transactions of the American Fisheries Society*, 142(5), 1192–1201. <https://doi.org/10.1080/00028487.2013.804005>
- Cowen, R. K., & Sponaugle, S. (2009). Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, 1(1), 443–466. <https://doi.org/10.1146/annurev.marine.010908.163757>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, 1000 Genomes Project Analysis. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>

- Dettmers, J. M., Janssen, J., Pientka, B., Fulford, R. S., & Jude, D. J. (2005). Evidence across multiple scales for offshore transport of yellow perch (*Perca flavescens*) larvae in Lake Michigan. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(12), 2683–2693. <https://doi.org/10.1139/f05-173>
- Dettmers, J. M., Raffenberg, M. J., & Weis, A. K. (2003). Exploring Zooplankton Changes in Southern Lake Michigan: Implications for Yellow Perch Recruitment. *Journal of Great Lakes Research*, 29(2), 355–364. [https://doi.org/10.1016/S0380-1330\(03\)70439-0](https://doi.org/10.1016/S0380-1330(03)70439-0)
- DeWoody, J. A., & Avise, J. C. (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, 56(3), 461–473. <https://doi.org/10.1111/j.1095-8649.2000.tb00748.x>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Evans, M. S. (1986). Recent Major Declines in Zooplankton Populations in the Inshore Region of Lake Michigan: Probable Causes and Implications. *Canadian Journal of Fisheries and Aquatic Sciences*, 43(1), 154–159. <https://doi.org/10.1139/f86-017>
- Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics*, 164(4), 1567–1587.
- Feron, R., Zahm, M., Cabau, C., Klopp, C., Roques, C., Bouchez, O., Ech , C., Vali re, S., Donnadi u, C., Haffray, P., Bestin, A., Morvezen, R., Acloque, H., Euclide, P. T., Wen, M., Jouano, E., Scharl, M., Postlethwait, J. H., Schraidt, C., ... Guiguen, Y. (2020). Characterization of a Y-specific duplication/insertion of the anti-Mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, *Perca flavescens*. *Molecular Ecology Resources*, 20(2), 531–543. <https://doi.org/10.1111/1755-0998.13133>
- Feyrer, F., Hobbs, J., Baerwald, M., Sommer, T., Yin, Q., Clark, K., May, B., & Bennett, W. (2007). Otolith Microchemistry Provides Information Complementary to Microsatellite DNA for a Migratory Fish. *Transactions of the American Fisheries Society*, 136(2), 469–476. <https://doi.org/10.1577/T06-044.1>

- Francis, J. T., Robillard, S. R., & Marsden, J. E. (1996). Yellow perch management in Lake Michigan: A multi-jurisdictional challenge. *Fisheries*, 21(2), 18–20.
- Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Molecular Ecology*, 10(12), 2741–2752. <https://doi.org/10.1046/j.0962-1083.2001.01411.x>
- Galperin, B., Kantha, L., Hassid, S., & Rosati, A. (1988). A Quasi-equilibrium Turbulent Energy Model for Geophysical Flows. *Journal of the Atmospheric Sciences*, 45, 55–62. [https://doi.org/10.1175/1520-0469\(1988\)045<0055:AQETEM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<0055:AQETEM>2.0.CO;2)
- GLFC Lake Michigan Committee: Makauskas, D., Clapp, D. 2018. Status of yellow perch in Lake Michigan 2016 - 2017. Summary Report. Lake Michigan Committee of Great Lake Fishery Commission. [http://www.glfc.org/pubs/lake_committees/michigan/yellow_perch/Status%20of%20YEP%20in%20LM%202016-2017%20\(2018\).pdf](http://www.glfc.org/pubs/lake_committees/michigan/yellow_perch/Status%20of%20YEP%20in%20LM%202016-2017%20(2018).pdf)
- Glover, D. C., Dettmers, J. M., Wahl, D. H., & Clapp, D. F. (2008). Yellow perch (*Perca flavescens*) stock structure in Lake Michigan: An analysis using mark–recapture data. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(9), 1919–1930. <https://doi.org/10.1139/F08-100>
- Goto, D., Dunlop, E. S., Young, J. D., & Jackson, D. A. (2020). Shifting trophic control of fishery–ecosystem dynamics following biological invasions. *Ecological Applications*, n/a(n/a). <https://doi.org/10.1002/eap.2190>
- Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Gräwe, U. (2011). Implementation of high-order particle-tracking schemes in a water column model. *Ocean Modelling*, 36(1), 80–89. <https://doi.org/10.1016/j.ocemod.2010.10.002>
- Hecky, R. E., Smith, R. E. H., Barton, D. R., Guildford, S. J., & al, et. (2004). The nearshore phosphorus shunt: A consequence of ecosystem engineering by dreissenids in the Laurentian Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences; Ottawa*, 61(7), 1285–1293.
- Hedgecock, D., Barber, P., & Edmands, S. (2007). Genetic Approaches to Measuring Connectivity. *Oceanography*, 20(3), 70–79. <https://doi.org/10.5670/oceanog.2007.30>

- Hoey, J. A., Fodrie, F. J., Walker, Q. A., Hilton, E. J., Kellison, G. T., Targett, T. E., Taylor, J. C., Able, K. W., & Pinsky, M. L. (2020). Using multiple natural tags provides evidence for extensive larval dispersal across space and through time in summer flounder. *Molecular Ecology*, *mec.15414*. <https://doi.org/10.1111/mec.15414>
- Hoey, J. A., & Pinsky, M. L. (2018). Genomic signatures of environmental selection despite nearpanmixia in summer flounder. *Evolutionary Applications*, *11*(9), 1732–1747. <https://doi.org/10.1111/eva.12676>
- Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*(5), 1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>
- Huret, M., Runge, J. A., Chen, C., Cowles, G., Xu, Q., & Pringle, J. M. (2007). Dispersal modeling of fish early life stages: Sensitivity with application to Atlantic cod in the western Gulf of Maine. *Marine Ecology Progress Series*, *347*, 261–274.
- Janssen, J., & Luebke, M. A. (2004). Preference for Rocky Habitat by Age-0 Yellow Perch and Alewives. *Journal of Great Lakes Research*, *30*(1), 93–99. [https://doi.org/10.1016/S0380-1330\(04\)70332-9](https://doi.org/10.1016/S0380-1330(04)70332-9)
- Johansson, M. L., Alberto, F., Reed, D. C., Raimondi, P. T., Coelho, N. C., Young, M. A., Drake, P. T., Edwards, C. A., Cavanaugh, K., Assis, J., Ladah, L. B., Bell, T. W., Coyer, J. A., Siegel, D. A., & Serrão, E. A. (2015). Seascape drivers of *Macrocystis pyrifera* population genetic structure in the northeast Pacific. *Molecular Ecology*, *24*(19), 4866–4885. <https://doi.org/10.1111/mec.13371>
- Johnson, D. W., Christie, M. R., Pusack, T. J., Stallings, C. D., & Hixon, M. A. (2018). Integrating larval connectivity with local demography reveals regional dynamics of a marine metapopulation. *Ecology*, *99*(6), 1419–1429. <https://doi.org/10.1002/ecy.2343>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Langmead, B., Wilks, C., Antonescu, V., & Charles, R. (2019). Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics*, *35*(3), 421–432. <https://doi.org/10.1093/bioinformatics/bty648>

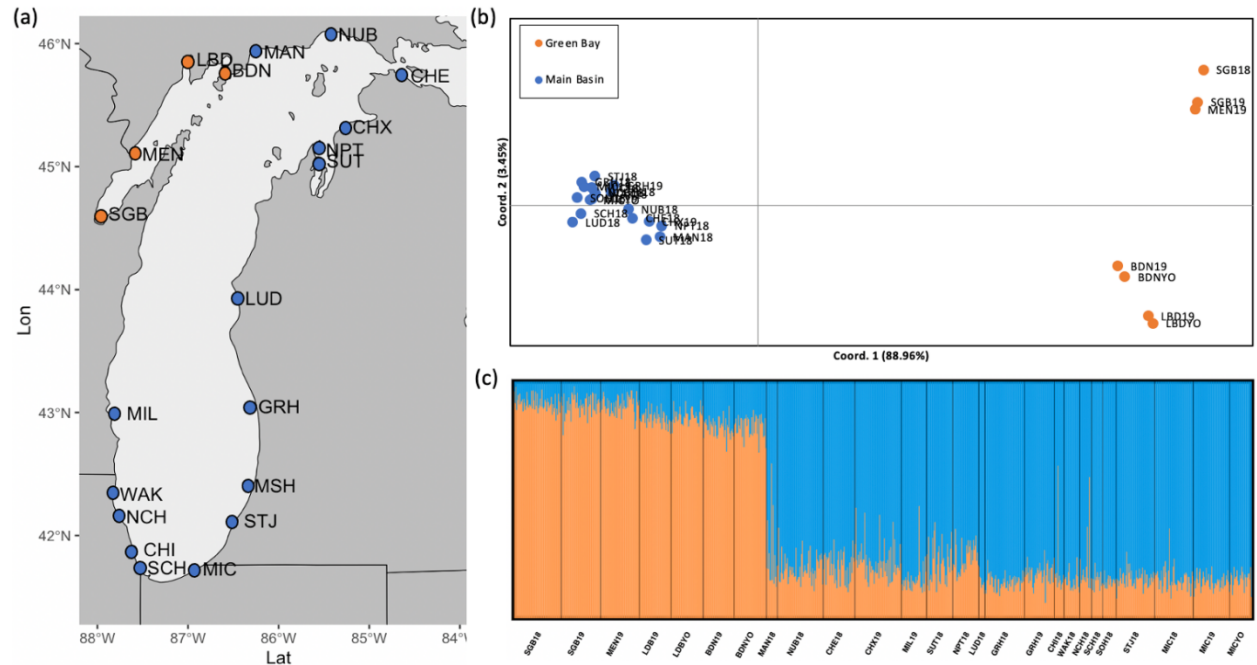
- Lawlor, J. A., & Arellano, S. M. (2020). Temperature and salinity, not acidification, predict near-future larval growth and larval habitat suitability of *Olympia* oysters in the Salish Sea. *Scientific Reports*, 10(1), 13787. <https://doi.org/10.1038/s41598-020-69568-w>
- Levin, L. A. (2006). Recent progress in understanding larval dispersal: New directions and digressions. *Integrative and Comparative Biology*, 46(3), 282–297. <https://doi.org/10.1093/icb/icj024>
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299. <https://doi.org/10.1093/bioinformatics/btr642>
- Madenjian, C. P., O’Gorman, R., Bunnell, D. B., Argyle, R. L., Roseman, E. F., Warner, D. M., Stockwell, J. D., & Stapanian, M. A. (2008). Adverse Effects of Alewives on Laurentian Great Lakes Fish Communities. *North American Journal of Fisheries Management*, 28(1), 263–282. <https://doi.org/10.1577/M07-012.1>
- Martinez, A. S., Willoughby, J. R., & Christie, M. R. (2018). Genetic diversity in fishes is influenced by habitat type and life-history variation. *Ecology and Evolution*, 8(23), 12022–12031. <https://doi.org/10.1002/ece3.4661>
- Mida, J. L., Scavia, D., Fahnenstiel, G. L., Pothoven, S. A., Vanderploeg, H. A., & Dolan, D. M. (2010). Long-term and recent changes in southern Lake Michigan water quality with implications for present trophic status. *Journal of Great Lakes Research*, 36, 42–49. <https://doi.org/10.1016/j.jglr.2010.03.010>
- Miller, L. M. (2003). Microsatellite DNA Loci Reveal Genetic Structure of Yellow Perch in Lake Michigan. *Transactions of the American Fisheries Society*, 132(3), 503–513. [https://doi.org/10.1577/1548-8659\(2003\)132<0503:MDLRGS>2.0.CO;2](https://doi.org/10.1577/1548-8659(2003)132<0503:MDLRGS>2.0.CO;2)
- Nanninga, G. B., Saenz-Agudelo, P., Manica, A., & Berumen, M. L. (2014). Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, 23(3), 591–602. <https://doi.org/10.1111/mec.12623>
- Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press.
- Okubo, A. (1971). Oceanic diffusion diagrams. *Deep Sea Research and Oceanographic Abstracts*, 18(8), 789–802. [https://doi.org/10.1016/0011-7471\(71\)90046-5](https://doi.org/10.1016/0011-7471(71)90046-5)

- O’Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren’t the loci you’re looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, 27(16), 3193–3206. <https://doi.org/10.1111/mec.14792>
- Peakall, R., & Smouse, P. E. (2006). genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28(19), 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155(2), 945–959.
- Riginos, C., & Liggins, L. (2013). Seascape Genetics: Populations, Individuals, and Genes Marooned and Adrift. *Geography Compass*, 7(3), 197–216. <https://doi.org/10.1111/gec3.12032>
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754. <https://doi.org/10.1111/mec.15253>
- Rowe, M. D., Anderson, E. J., Wynne, T. T., Stumpf, R. P., Fanslow, D. L., Kijanka, K., Vanderploeg, H. A., Strickler, J. R., & Davis, T. W. (2016). Vertical distribution of buoyant *Microcystis* blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. *Journal of Geophysical Research: Oceans*, 121(7), 5296–5314. <https://doi.org/10.1002/2016JC011720>
- Roy, D., Hurlbut, T. R., & Ruzzante, D. E. (2012). Biocomplexity in a demersal exploited fish, white hake (*Urophycis tenuis*): Depth-related structure and inadequacy of current management approaches. *Canadian Journal of Fisheries and Aquatic Sciences*, 69(3), 415–429. <https://doi.org/10.1139/f2011-178>
- Selkoe, K., D’Aloia, C., Crandall, E., Iacchei, M., Liggins, L., Puritz, J., von der Heyden, S., & Toonen, R. (2016). A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554, 1–19. <https://doi.org/10.3354/meps11792>

- Sepulveda-Villet, O. J., Ford, A. M., Williams, J. D., & Stepien, C. A. (2009). Population genetic diversity and phylogeographic divergence patterns of the yellow perch (*Perca flavescens*). *Journal of Great Lakes Research*, 35(1), 107–119. <https://doi.org/10.1016/j.jglr.2008.11.009>
- Sepulveda-Villet, O. J., & Stepien, C. A. (2011). Fine-scale population genetic structure of the yellow perch *Perca flavescens* in Lake Erie. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(8), 1435–1453. <https://doi.org/10.1139/f2011-077>
- Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The Basic Experiment. *Monthly Weather Review*, 91(3), 99–164. [https://doi.org/10.1175/1520-0493\(1963\)091<0099:GCEWTP>2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2)
- Smouse, P. E., & Peakall, R. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, 82(5), 561–573. <https://doi.org/10.1038/sj.hdy.6885180>
- Wang, I. J. (2013). Examining the Full Effects of Landscape Heterogeneity on Spatial Genetic Variation: A Multiple Matrix Regression Approach for Quantifying Geographic and Ecological Isolation. *Evolution*, 67(12), 3403–3411. <https://doi.org/10.1111/evo.12134>
- Weersing, K., & Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, 393, 1–12. JSTOR.
- White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences*, 277(1688), 1685–1694. <https://doi.org/10.1098/rspb.2009.2214>
- Whiteside, M. C., Swindoll, C. M., & Doolittle, W. L. (1985). Factors affecting the early life history of yellow perch, *Perca flavescens*. *Environmental Biology of Fishes*, 12(1), 47–56. <https://doi.org/10.1007/BF00007709>
- Wilberg, M. J., Bence, J. R., Eggold, B. T., Makauskas, D., & Clapp, D. F. (2005). Yellow Perch Dynamics in Southwestern Lake Michigan during 1986–2002. *North American Journal of Fisheries Management*, 25(3), 1130–1152. <https://doi.org/10.1577/M04-193.1>
- Wilder, A. P., Palumbi, S. R., Conover, D. O., & Therkildsen, N. O. (2020). Footprints of local adaptation span hundreds of linked genes in the Atlantic silverside genome. *Evolution Letters*, 4(5), 430–443. <https://doi.org/10.1002/evl3.189>

1.6 Figures and Tables

Figure 1.1. (a) A map of sample locations across Lake Michigan including Green Bay (orange) and the main basin (blue). (b) Principal Coordinates Analysis showing relatedness between



individual populations based on pairwise *fst*. (c) Results of individuals clustering analysis with STRUCTURE ($K=2$). Colored lines correspond to an individual's estimated proportion membership to each cluster, and individuals are grouped by population (bold black lines).

Table 1.1. Summary statistics for all populations ($n = 26$). Number of individuals sampled (N) and summary statistics including observed heterozygosity (H_o), expected heterozygosity (H_e), allelic richness, (Ar), and % missing data.

Site ID	N	H_o	H_e	Ar	%missing
Main Basin					
MIL19	32	0.491	0.451	1.963	7.78
WAK18	20	0.490	0.446	1.964	6.32
NCH18	15	0.464	0.439	1.960	10.82
CHI18	12	0.481	0.442	1.967	2.68
SCH18	14	0.480	0.442	1.963	8.44
MIC18	49	0.494	0.454	1.963	5.77
MIC19	46	0.498	0.452	1.962	8.57
MICYO	28	0.497	0.452	1.964	4.26
STJ18	49	0.504	0.452	1.961	10.57
GRH18	50	0.487	0.454	1.962	4.04
GRH19	38	0.638	0.459	1.968	10.32
SOH18	17	0.502	0.445	1.963	5.12
LUD18	8	0.484	0.425	1.959	7.14
SUT18	32	0.498	0.457	1.970	7.82
NPT18	34	0.518	0.462	1.973	7.56
CHX19	59	0.512	0.460	1.969	10.01
CHE18	40	0.494	0.459	1.969	3.91
NUB18	58	0.494	0.459	1.968	9.92
MAN18	14	0.478	0.449	1.970	5.83

Table 1.1 continued

Green Bay					
BDN19	39	0.480	0.439	1.949	8.41
BDNYO	41	0.476	0.437	1.948	7.55
SGB18	60	0.451	0.421	1.927	5.46

SGB19	50	0.509	0.426	1.931	12.50
LBD19	40	0.464	0.433	1.941	7.59
LBDYO	41	0.469	0.432	1.939	7.56
MEN19	49	0.456	0.425	1.932	7.00

Table 1.2. Pairwise F_{ST} 's across all populations. Values that did not differ from zero based on 95% CI's were excluded.

	BDN19	BDNYO	CHI18	WAK18	NCH18	SCH18	SGB18	SGB19	GRH18	GRH19	MIC18	MIC19	LBD19	LBDYO	MAN18	CHE18	CHX19	MICYO	MEN19	MIL19	STJ18	LUD18	SOH18	SUT18	NPT18
BDNYO	0.000886																								
CHI18	0.080302	0.083696																							
WAK18	0.085694	0.087729	0.001972																						
NCH18	0.085058	0.088456																							
SCH18	0.087246	0.090529																							
SGB18	0.01297	0.014286	0.107875	0.114378	0.113863	0.118692																			
SGB19	0.013243	0.013619	0.106315	0.109853	0.110386	0.116292	0.002657																		
GRH18	0.089172	0.091631	0.001657	0.001534	0.001486		0.11734	0.113624																	
GRH19	0.082003	0.084792	0.008964	0.006318	0.008144	0.006484	0.110203	0.106388	0.002534																
MIC18	0.085971	0.088534	0.002526	0.001225	0.001511		0.113945	0.11015		0.002616															
MIC19	0.088336	0.091027		0.001155			0.117297	0.113105		0.002238															
LBD19	0.007335	0.00768	0.090135	0.094699	0.095151	0.097635	0.017116	0.01273	0.099016	0.092742	0.095549	0.097809													
LBDYO	0.010976	0.010406	0.094632	0.097397	0.098045	0.099958	0.020031	0.013335	0.101969	0.095274	0.098526	0.101756													
MAN18	0.064429	0.06662		0.004452	0.003786	0.002169	0.093669	0.090398	0.003858	0.008334	0.003755	0.003675	0.074065	0.07783											
CHE18	0.072846	0.075471	0.002853	0.003416	0.002814	0.001617	0.101838	0.098084	0.002457	0.003947	0.002692	0.002821	0.08284	0.085788											
CHX19	0.069283	0.071188	0.004913	0.004913	0.00501	0.003568	0.097372	0.092875	0.004236	0.003619	0.00337	0.004543	0.0787	0.081841		0.001165									
MICYO	0.08506	0.087524	0.001621		0.003182		0.115313	0.111367		0.003682	0.000796	0.001112	0.095975	0.099727		0.001625	0.0033								
MEN19	0.012263	0.012718	0.104393	0.108594	0.10871	0.11303	0.002784	0.00184	0.112354	0.105803	0.109708	0.112697	0.011843	0.013156	0.08928	0.096774	0.0926	0.110348							
MIL19	0.086334	0.088474	0.001605	0.001045	0.001281	-0.00078	0.115677	0.111058	0.000677	0.003861		0.001042	0.094861	0.098574	0.003968	0.002576	0.004419		0.110428						
STJ18	0.086177	0.088964	0.004011	0.002181	0.00223	0.001023	0.113899	0.109214	0.000779	0.003207		0.001214	0.095642	0.098896	0.00524	0.003269	0.003833		0.10851	0.001189					
LUD18	0.092757	0.096123	0.003166	0.002844		0.002683	0.124765	0.123346	0.002544	0.012597		0.003245	0.103092	0.10509	0.006119	0.005004	0.00635	0.003414	0.119077	0.004873	0.004102				
SOH18	0.089095	0.091719	0.002625				0.119355	0.116868		0.004493			0.099969	0.103336	0.003105	0.001296	0.004239		0.114219	0.000939	0.001564	0.002349			
SUT18	0.068913	0.071459	0.006047	0.004493	0.004307	0.003797	0.099505	0.09529	0.004194	0.005332	0.004206	0.005062	0.079507	0.08279			4.92E-05	0.003701	0.094881	0.004693	0.005336	0.006856	0.003388		
NPT18	0.067422	0.069384	0.008331	0.007097	0.008009	0.00698	0.095056	0.090959	0.005991	0.006095	0.004841	0.006171	0.076465	0.080138	0.002415	0.002	0.00091	0.004639	0.091076	0.005089	0.006249	0.01347		0.000477	
NUB18	0.073109	0.076395	0.002059	0.001889	0.001155		0.102227	0.097743	0.001384	0.002467	0.000761	0.001721	0.08322	0.086502					0.097176	0.001475	0.001618	0.003281			0.001905

Table 1.3. A summary of relevant FVCOM models utilized in connectivity matrices.

Year	Start Date (Julian)	End Date (Julian)	Vertical Velocity	Horizontal Diffusivity	# Particles	Run Duration
2014	May 25 (145)	July 14 (195)	0.0003 m/s	5.6 m ² /s	224600	50 days
2014	June 01 (152)	July 21 (202)	0.0003 m/s	5.6 m ² /s	224600	50 days
2014	June 08 (159)	July 28 (209)	0.0003 m/s	5.6 m ² /s	224600	50 days
2014	June 15 (166)	Aug 04 (216)	0.0003 m/s	5.6 m ² /s	224600	50 days
2014	June 22 (173)	Aug 11 (223)	0.0003 m/s	5.6 m ² /s	224600	50 days
2014	June 29 (180)	Aug 18 (230)	0.0003 m/s	5.6 m ² /s	224600	50 days
2015	May 24 (144)	July 13 (194)	0.0003 m/s	5.6 m ² /s	224600	50 days
2015	May 31 (151)	July 20 (201)	0.0003 m/s	5.6 m ² /s	224600	50 days
2015	June 07 (158)	July 27 (208)	0.0003 m/s	5.6 m ² /s	224600	50 days
2015	June 14 (165)	Aug 03 (215)	0.0003 m/s	5.6 m ² /s	224600	50 days
2015	June 21 (172)	Aug 10 (222)	0.0003 m/s	5.6 m ² /s	224600	50 days
2015	June 28 (179)	Aug 17 (229)	0.0003 m/s	5.6 m ² /s	224600	50 days
2016	May 29 (150)	July 18 (200)	0.0003 m/s	5.6 m ² /s	224600	50 days
2016	June 05 (157)	July 25 (207)	0.0003 m/s	5.6 m ² /s	224600	50 days
2016	June 12 (164)	Aug 01 (214)	0.0003 m/s	5.6 m ² /s	224600	50 days
2016	June 19 (171)	Aug 08 (221)	0.0003 m/s	5.6 m ² /s	224600	50 days
2016	June 26 (178)	Aug 15 (228)	0.0003 m/s	5.6 m ² /s	224600	50 days
2016	July 03 (185)	Aug 22 (235)	0.0003 m/s	5.6 m ² /s	224600	50 days
2017	May 28 (148)	July 17 (198)	0.0003 m/s	5.6 m ² /s	224600	50 days
2017	June 04 (155)	July 24 (205)	0.0003 m/s	5.6 m ² /s	224600	50 days
2017	June 11 (162)	July 31 (212)	0.0003 m/s	5.6 m ² /s	224600	50 days
2017	June 18 (169)	Aug 07 (219)	0.0003 m/s	5.6 m ² /s	224600	50 days
2017	June 25 (176)	Aug 14 (226)	0.0003 m/s	5.6 m ² /s	224600	50 days
2017	July 02 (183)	Aug 21 (233)	0.0003 m/s	5.6 m ² /s	224600	50 days
2018	May 27 (147)	July 16 (197)	0.0003 m/s	5.6 m ² /s	224600	50 days
2018	June 03 (154)	July 23 (204)	0.0003 m/s	5.6 m ² /s	224600	50 days
2018	June 10 (161)	July 30 (211)	0.0003 m/s	5.6 m ² /s	224600	50 days
2018	June 17 (168)	Aug 06 (218)	0.0003 m/s	5.6 m ² /s	224600	50 days
2018	June 24 (175)	Aug 13 (225)	0.0003 m/s	5.6 m ² /s	224600	50 days
2018	July 01 (182)	Aug 20 (232)	0.0003 m/s	5.6 m ² /s	224600	50 days
2019	May 26 (146)	July 15 (196)	0.0003 m/s	5.6 m ² /s	224600	50 days
2019	June 02 (153)	July 22 (203)	0.0003 m/s	5.6 m ² /s	224600	50 days
2019	June 09 (160)	July 29 (210)	0.0003 m/s	5.6 m ² /s	224600	50 days
2019	June 16 (167)	Aug 05 (217)	0.0003 m/s	5.6 m ² /s	224600	50 days
2019	June 23 (174)	Aug 12 (224)	0.0003 m/s	5.6 m ² /s	224600	50 days
2019	June 30 (181)	Aug 19 (231)	0.0003 m/s	5.6 m ² /s	224600	50 days

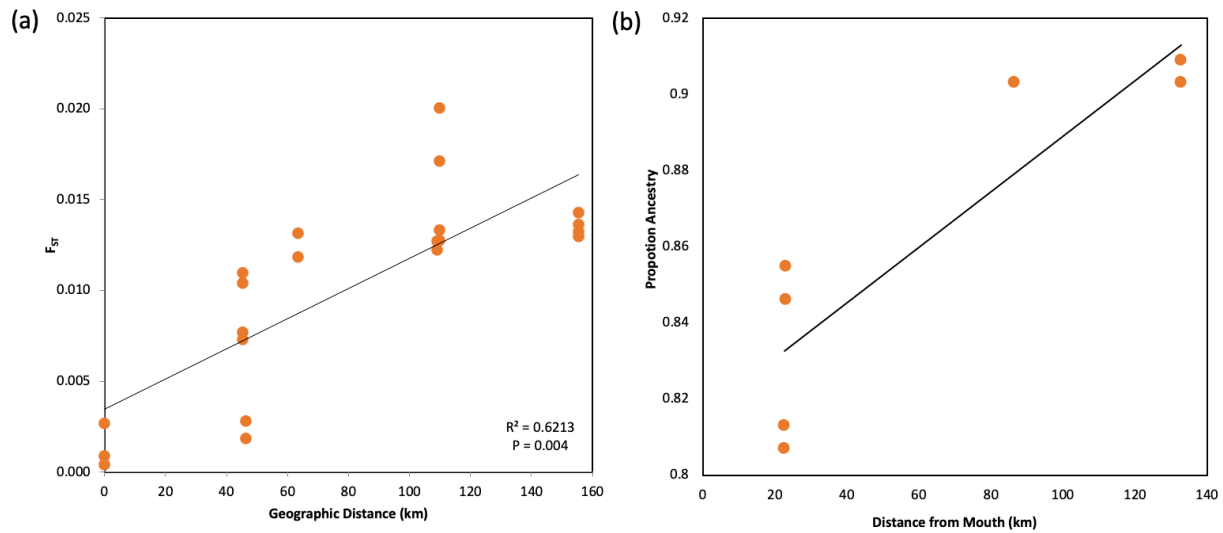


Figure 1.2. (a) Isolation by distance (IBD) plot of F_{ST} values against geographic (Euclidean) distance among Green Bay populations. (b) Regression analysis of proportion Green Bay ancestry and distance from the mouth of Green Bay.

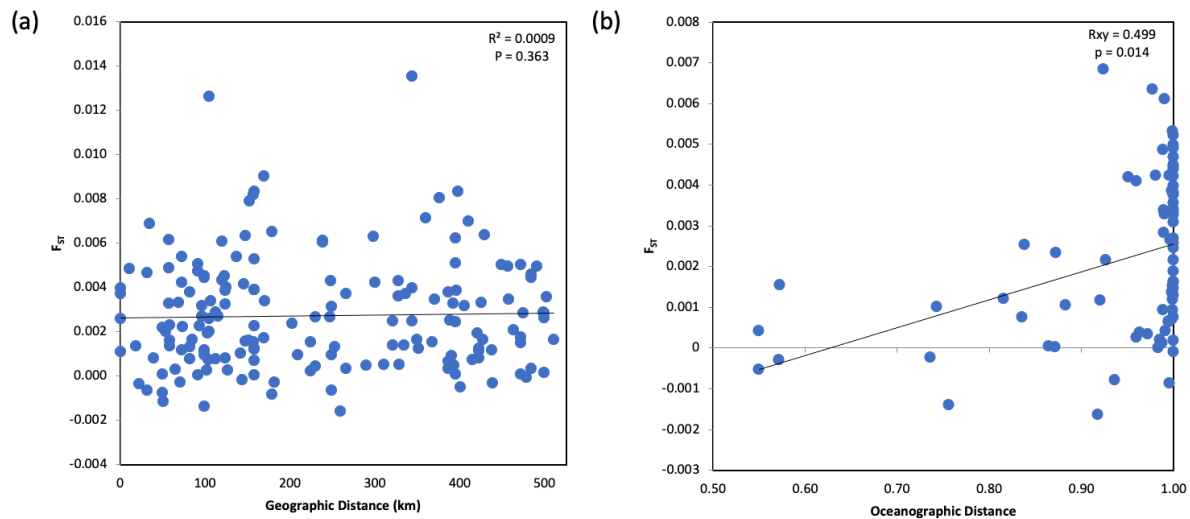


Figure 1.3. (a) Isolation by distance (IBD) plot of F_{ST} values against geographic (Euclidean) distance among main basin populations. (b) Regression analysis of genetic distance (F_{ST}) and oceanic distance (1 - dispersal probability), showing increased genetic distance with increased oceanic distance.