

# **Mixed-stock analysis in the age of genomics: Rapture genotyping enables evaluation of stock-specific exploitation in a freshwater fish population with weak genetic structure**

Peter T. Euclide<sup>1</sup>, Tom MacDougall<sup>2</sup>, Jason M. Robinson<sup>3</sup>, Matthew D. Faust<sup>4</sup>, Chris C. Wilson<sup>5</sup>, Kuan-Yu Chen<sup>6</sup>, Elizabeth A. Marschall<sup>6</sup>, Wesley Larson<sup>7\*</sup>, Stuart A. Ludsin<sup>6</sup>

<sup>1</sup>Wisconsin Cooperative Fishery Research Unit, College of Natural Resources, University of Wisconsin-Stevens Point, 800 Reserve St., Stevens Point, WI 54481

<sup>2</sup>Lake Erie Management Unit, Ontario Ministry of Natural Resources and Forestry 49 Passmore Avenue, Port Dover, ON N0A 1N7, Canada

<sup>3</sup>New York State Department of Environmental Conservation, Lake Erie Fisheries Research Unit 178 Point Drive North, Dunkirk, NY 14048, USA

<sup>4</sup>Ohio Department of Natural Resources, Division of Wildlife, Sandusky Fisheries Research Station, 305 E. Shoreline Drive, Sandusky, Ohio 44870, USA

<sup>5</sup>Aquatic Research and Development Section, Ontario Ministry of Natural Resources and Forestry 2140 East Bank Drive, Peterborough, ON K9L 0G2, Canada

<sup>6</sup>Aquatic Ecology Laboratory, Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH 43210

<sup>7</sup>U.S. Geological Survey, Wisconsin Cooperative Fishery Research Unit, College of Natural Resources, University of Wisconsin-Stevens Point, 800 Reserve St., Stevens Point, WI 54481

\*Current address: National Oceanic and Atmospheric Administration, Ted Stevens Marine Research Institute, Juneau, Alaska, 99801

Keywords: Portfolio theory, stock discrimination, RAD-capture, RAD-seq, Great Lakes, genetic stock identification

Running title: Rapture enables walleye Great Lakes stock discrimination

## Abstract

Mixed-stock analyses using genetic markers have informed fisheries management in cases where strong genetic differentiation occurs among local spawning populations, yet many fisheries are supported by multiple spawning stocks that are weakly differentiated. Freshwater fisheries exemplify this problem, with many harvested populations supported by multiple stocks of young evolutionary age that are isolated across small spatial scales. As a result, attempts to conduct genetic mixed-stock analyses of inland fisheries have often been unsuccessful. Advances in genomic sequencing now offer the ability to discriminate among populations with weak population structure by providing the necessary resolution to conduct mixed-stock assignment among previously indistinguishable stocks. We demonstrate the use of genomic data to conduct a mixed-stock analysis of Lake Erie's commercial and recreational walleye (*Sander vitreus*) fisheries and estimate the relative harvest of weakly differentiated stocks (pairwise  $F_{ST} < 0.01$ ). We used RAD-capture (Rapture) to sequence and genotype individuals at 12,081 loci that had been previously determined to be capable of discriminating between western and eastern basin stocks with 95% reassignment accuracy, which was not possible in the past with microsatellite markers. Genetic assignment of 1,075 fish harvested from recreational and commercial fisheries in the eastern basin indicated that western basin stocks constituted the majority of individuals harvested during peak walleye fishing season (July – September). Composition of harvest changed seasonally, with eastern basin fish comprising much of the early season harvest (May – June). Clear spatial structure in harvest composition existed; more easterly sites contained more individuals of east basin origin than did westerly sites. Our study provides important stock contribution estimates for Lake Erie fishery management and demonstrates the power of

genomic data to facilitate mixed-stock analysis in exploited fish populations with weak population structure or limited existing genetic resources.

## Introduction

The sustainability of many populations depends on multiple, but often cryptic, breeding groups connected by shared habitat or reproductive behaviors (Hilborn et al. 2003; Cowen et al. 2007; Alves et al. 2010). Such complexity helps to improve population stability and resilience (Hilborn et al. 2003; Schindler et al. 2010; 2015), but it also complicates conservation and management by increasing the number of areas where conservation and regulatory actions could occur (Cooke et al. 2016). In some cases, populations have been sustainably managed by treating discrete local spawning stocks as parts of an overall portfolio of population diversity (Schindler et al. 2015; Waples et al. 2020). Although this idea has primarily been applied to marine fisheries, it has also been suggested to be relevant to freshwater fisheries (DuFour et al. 2015). However, fisheries management often presumes that assessment information represents a single stock as opposed to multiple stocks, which can lead to unintended overexploitation of local stocks and inappropriate management advice (Stephenson 1999; Li et al. 2015). For example, the collapse of Atlantic cod *Gadus morhua* populations is thought to have been caused in part by a failure to account for multiple stocks in the assessment and management of this fishery (Hutchinson 2008). An ability to accurately discriminate and identify population components (such as local spawning stocks) is integral to the conservation and management of populations that fit the portfolio theory model of ecology and evolution because it provides a means of accounting for variance in stock-specific productivity (Figge 2004; Sethi 2010). However, achieving accurate discrimination among population components can be difficult, especially for freshwater species in large lakes that often

experience high gene flow similar to marine populations, but at reduced spatial scales (Martinez et al. 2018) or are of young evolutionary age. Thus, the need exists for methods that can deal with the resultant weak population structure.

For decades, molecular markers have been used to resolve stock-specific contributions to mixed-stock fisheries in marine systems (Seeb and Crane 1999; Milner et al. 2008; Bernatchez et al. 2017; Waples et al. 2020). For example, mixed-stock assessments have become central to the management of Pacific salmonines (Shaklee et al. 1999) such that the contributions of hundreds of salmon spawning populations are evaluated annually using molecular markers (Dann et al. 2013; Beacham et al. 2020). Similar practices have become essential to Atlantic Cod *Gadus morhua* management by improving understanding of seasonal spawning dynamics (Dean et al. 2019) and limiting overharvest of less productive stocks (Dahle et al. 2018). Freshwater fisheries could benefit from similar portfolio-based approaches to management as they are often under high fishing pressure (DuFour et al. 2015; Fluet-Chouinard et al. 2018; Embke et al. *in press*) and experience more environmental stochasticity than marine systems (Strayer and Dudgeon 2010). However, application of mixed-stock assessments use of similar techniques has been limited, partially owing to difficulties in developing molecular resources capable of accurately discriminating among stocks (e.g., Chen et al. 2020, but see Andvik et al. 2016).

Advancing stock identification in freshwater systems is important because the same portfolio theory strategies used in marine systems have been suggested to improve population stability of exploited freshwater species that display complex portfolios of diversity (DuFour et al. 2015). Because the habitat of many freshwater species shifted southward and was reduced during the

last Ice Age, the recolonized populations remain relatively young (4,000 – 12,000 years old) and lack substantial genetic differences among spawning stocks (Bernatchez & Wilson, 1998). Additionally, freshwater populations often contain spawning stocks in relatively close proximity to each other and admixture among stocks can complicate stock reassignment (Waples and Gaggiotti 2006; Martinez et al. 2018). Therefore, traditional genetic approaches such as microsatellites often lack the statistical power to consistently discriminate among spawning stocks (Johnson et al. 2004). Instead, the development of high-resolution genomic marker panels that can discriminate among local spawning stocks could make mixed-stock management accessible for fisheries managers tasked with managing such species.

Developing high-resolution marker panels can require a large amount of time and analytical experience, which has limited the use of such panels to high-value marine fisheries, such as salmonids and cod. However, the cost of conducting genomic studies has decreased substantially (Meek and Larson 2019) and molecular and statistical advancements have made it possible to untangle complex patterns of population structure and conduct stock identification of mixed-stock harvests (Layton et al. 2020; Carrier et al. 2020). For example, advances in genotyping-by-sequencing technology, such as RAD-capture (Rapture), have made it possible to affordably and consistently genotype thousands of genetic markers to discriminate among finely structured populations (Ali et al. 2016; Sard et al. 2020). Rapture is a sequence capture based approach that uses biotinylated RNA capture baits to target loci identified using reduced representation approaches, such as restriction site associated DNA sequencing (RAD-Seq). Additionally, new bioinformatic approaches have led to increased discriminatory power among populations with weak genetic structure. For example, the identification and use of microhaplotypes, which are

polyallelic haplotypes of two or more SNPs that are in high linkage disequilibrium and can be sequenced together within the same 150-300 bp region (McKinney et al. 2017a; Baetscher et al. 2018). Advances in computing power and statistics have also resulted in improved reassignment accuracy through resampling cross-validation and machine-learning algorithms (Anderson et al. 2008; Chen et al. 2018). As a result, mixed-stock assignment using genomic data offers the potential to benefit management of fisheries supported by stocks that are difficult to distinguish and that are vulnerable to differentiated overexploitation by commercial, recreational, and subsistence fishing (Post et al. 2002; Fluet-Chouinard et al. 2018).

Fish populations supported by multiple stocks are common in the Laurentian Great Lakes, which have a long history of commercial, recreational, and subsistence fishing (Regier and Hartman 1973; Lynch et al. 2016). Walleye (*Sander vitreus*) are one of the most ecologically and economically important species in the Great Lakes (Hatch et al. 1987; Ludsin et al. 2014) and has been the focus of many previous stock discrimination research efforts (Johnson et al. 2004; Stepien et al. 2015; Chen et al. 2017). Lake Erie's walleye population is the largest of the five Laurentian Great Lakes and is supported by multiple discrete stocks that spawn throughout the lake (Stepien and Faber 1998; Zhao et al. 2011; Stepien et al. 2012). Most of the lake's walleye production occurs in the western basin where tributary and open-lake spawning stocks of varying productivity exist (DuFour et al. 2015). Individuals from these stocks move throughout Lake Erie during non-spawning periods in search of preferred habitat (Kersher et al. 1999; Raby et al. 2018), where they intermix with the smaller spawning stocks from the central and eastern basins (Zhao et al. 2011; Vandergoot and Brenden 2014; Matley et al. 2020). Migration of western basin stocks into Lake Erie's eastern basin is likely to have a disproportionate influence on local

commercial and recreational fisheries because of presumed differences in population productivity and abundance (Zhao et al. 2011). However, the degree to which western versus eastern stocks are harvested by recreational and commercial fisheries in a given year or season is unknown. Further, efforts to discriminate among western and eastern stocks to facilitate mixed-stock assessments using biological markers have been largely unsuccessful (Riley and Carline 1982; Hedges 2002; Johnson et al. 2004; Chen et al. 2017). These needs, in turn, limit management in Lake Erie (MF, JR, TM *personal communication*).

Genomic tools represent an affordable solution for understanding stock-specific dynamics with a recent study using thousands of single nucleotide polymorphisms (SNPs) genotyped with RAD-seq suggesting that high-accuracy reassignment of walleye to basin of origin is possible using genomic tools (Chen et al. 2020). Here, we developed a Rapture panel containing thousands of genetic markers to conduct a mixed-stock analysis of walleye harvest in Lake Erie's eastern basin. Specifically, our objectives were to 1) quantify the relative contributions of western and eastern basin walleye populations to commercial and recreational harvest within the eastern basin, and 2) determine if the contributions of western and eastern basin populations varied spatio-temporally within the eastern basin. We hypothesized that walleye from the more productive western basin spawning stocks would comprise the majority of harvest for both types of fisheries, but that harvest composition would vary spatially and temporally based on annual seasonal migrations (Kershner et al. 1999; Raby et al. 2018). In testing this hypothesis, we ultimately demonstrate the ability of genomic data to resolve genetic stock structure and facilitate mixed-stock analysis in systems with weak population structure, to the benefit of fisheries management.

## Materials and Methods

### *Study system*

Lake Erie, which has a surface area of ~ 26,000 km<sup>2</sup>, is composed of three distinct basins. The western basin is shallowest with most areas ranging 3 – 7 m in depth. The central basin is deeper (15 – 18 m deep), with the eastern basin being the deepest (15 – 25 m) and also containing the deepest point in the lake (65 m; Holcombe et al. 2003). Lake Erie supports the largest commercial and recreational walleye fisheries in the world, which are collectively valued at more than 1 billion dollars annually (Roseman et al. 2014). These fisheries are managed via inter-jurisdictional fisheries management with an annual total allowable catch that is distributed across the lake without consideration of relative stock contributions (Kayle et al. 2015; Fig. 1).

In the western basin, four large stocks exist. The largest stocks spawn on open-lake reefs, whereas the others spawn in three large tributaries (Maumee, Sandusky, and Detroit Rivers) that drain into the western basin (DuFour et al. 2015; Fig. 1). Spawning is known to occur in the central basin on nearshore reefs and in the Grand River, Ohio (Stepien et al. 2018). Although the contributions of the resulting offspring to the lake-wide fishery remains unknown for these stocks, it is suspected to be much less than that from the western basin stocks. More spawning stocks exist in the eastern basin, where small spawning aggregations have been reported in small New York tributaries and nearshore reefs dispersed along the south-eastern shore of the east basin, as well as in Ontario's Grand River on the northern shore of the eastern basin (Zhao et al. 2011; Fig. 1).



## Rapture panel development

We developed a Rapture bait panel (Ali et al. 2016) to reduce the cost and increase the power and consistency of genetic stock identification. Loci for our Rapture panel were identified by conducting a preliminary RAD-sequencing on 96 walleye (Table 1). These individuals were sampled on known spawning grounds during the peak spawning season and were and therefore presumed to have originated from that site (Chen et al. 2020). All of these fish were collected by Lake Erie management agencies, including: Ohio Department of Natural Resources, New York Department of Environmental Conservation, and Ontario Ministry of Natural Resources and Forestry. Fin clips were preserved in 95% ethanol until DNA extraction using Qiagen DNEasy 96 kits.

A single RAD-seq library was prepared using *SbfI* enzyme and the standard library preparation approach outlined in Ali et al. (2016) and detailed in Ackiss et al. (2020). Samples were sequenced using paired-end 150 BP sequencing on an Illumina HiSeq4000 at NovoGene (Sacramento, CA). Loci were then assembled *de novo* in STACKS v. 2.0 (Catchen et al. 2011; 2013) and a catalog of all putative loci was created in *cstacks* from data of all 96 individuals. Samples were demultiplexed using *process\_radtags* (-e sbfI -i gzfastq -c -q -r --filter\_illumina - bestrad), assembled *de novo* in *ustacks* (--disable-gapped, -m3, -M 3, -H, --max\_locus\_stacks 4, -model\_type bounded, --bound\_high 0.05), matched in *sstacks* (--disable gapped), converted to bam files using *tsv2bam*, and genotyped in *gstacks*. Finally, genotypes were called for all SNPs with a minor allele count greater than three (--mac 3). We then removed putative paralogs identified in HDPlot (McKinney et al. 2017b), as well as loci with minor allele frequencies <0.01, heterozygosity <0.05, and genotype rate < 50%. Sequence data for the remaining loci that

met our quality standards were sent to Arbor Biosciences (Ann Arbor, MI) for bait development. Arbor Biosciences conducted additional quality filters including a blast alignment to the yellow perch (*Perca flavescens*) genome (Feron et al. 2019) and then designed and synthesized two baits per-locus resulting in a final panel of 12,081 loci and 24,162 unique baits. STACKS 2 catalog files and a fasta file for all 12,081 baited loci can be found on Dryad (*data will be available upon acceptance*).

### *Baseline sequencing*

The baseline population structure among six east basin, four west basin, and one central basin walleye spawning stocks (Fig. 1) was evaluated by sequencing 397 individuals collected during the 2012-2018 spawning seasons using our Rapture panel described above (Table 1). All samples were collected and extracted under the same protocols as Rapture panel development. Whenever possible samples from at least two different years collected from the same site were used to account for possible temporal variation (Table 1). Individuals sampled on known spawning grounds during the peak spawning season were presumed to have spawned in those areas. RAD-seq libraries were prepared identically as panel development individuals and then bait capture was conducted for each library following the myBaits protocol v.4.01(<https://arborbiosci.com/mybaits-manual/>) with minor modifications. In short, RAD-seq libraries were hybridized with the bait mixture for 16 hours at 65°C and then amplified using 10 PCR cycles, universal primers, and an annealing temperature of 56°C. Final Rapture libraries were then purified with a 1X Ampure bead solution and submitted for sequencing on 1/2 of a S4 NovaSeq lane at NovoGene (Sacramento, CA). Data were processed using STACKS v.2.3 (Rochette et al. 2019) with identical parameters and catalog as panel development. Finally, loci

were filtered using a locus-specific whitelist that included only the 12,081 loci included in the final Rapture panel. Microhaplotypes at each locus were identified in the *populations* step of STACKS and used for all downstream analysis. To ensure that SNPs used in microhaplotype genotypes were not the result of sequencing error, loci had to be genotyped in at least 80% of individuals (both baseline and mixed-stock) and SNPs had to have a minor allele count of three or more to be processed.

#### *Baseline population structure and identification of reporting groups (natal sources)*

*F*-statistics and discriminant analysis of principal components (DAPC; Jombart et al. 2010) were used to describe overall patterns of genetic structure in Lake Erie and to identify putative reporting groups for reassignment tests to determine classification accuracies. Observed and expected heterozygosity, inbreeding coefficient, and pairwise Weir and Cockerham  $F_{ST}$  were estimated in the DiveRsity v.1.1.9 R package (Keenan et al. 2013; R Core Team 2019). We used DAPC to identify putative clusters of spawning sites that could be combined into reporting groups for assignment of individuals of unknown origin. DAPC was run using the ADEGENET v.2.1.2 R package first with individuals grouped by spawning stock and next with individuals grouped into four clusters identified with the `find.clusters()` function (Jombart 2008). To avoid model over-fitting, the optimal number of principal components necessary to explain the variance among groups was identified in ADEGENET with the *optim.a.score()* function (Jombart 2008). Finally, we used AMOVA and permutation tests of significance implemented in the poppr v.2.8.3 R package to determine how much variance was explained when sites were grouped by putative reporting groups (Kamvar et al. 2014).

We tested reassignment accuracy of putative reporting groups identified through baseline population genetic analysis in the assignPOP v.1.1.8 R package (Chen et al. 2018). The assignment tests were performed using Monte-Carlo cross-validation in which individuals were randomly resampled as a training set, using remaining individuals as a test (holdout) set to avoid upwardly biased test results (Anderson 2010, Waples 2010). Specifically, we chose three proportions (0.5, 0.7, and 0.9) of individuals from each reporting group and used either random half or all loci as training data (combination  $N = 6$ ) to perform the assignment test. Each combination of training data and the test were iterated 30 times for a total of 180 assignment tests. This procedure allowed us to evaluate variation in assignment accuracy and how different proportions of training individuals influenced the assignment results.

The assignPOP predictive models were built using a support vector machine (with linear kernel and parameter cost = 1), which has been shown to generate higher assignment accuracies than other models (i.e., LDA, naiveBayes, decision tree, and random forest; Chen et al. 2018) and had the highest accuracy in preliminary trial runs. We ran the reassignment test twice: once with every spawning stock kept separate and once with stocks clustered into our reporting groups identified with DAPC and pairwise  $F_{ST}$ . Because individual assignments can sometimes bias mixture results, we compared assignment accuracy identified in assignPOP with 100% mixture assignments tests conducted in the rubias R package v.0.3.0 (Moran and Anderson 2019). Assignment accuracy of 100% mixtures for all reporting groups were estimated using a leave-one-out approach, 25 replicates, and a mixture size of 100. All presented graphics were constructed in R primarily using the ggplot 3.3.0 (Wickham 2009), ggpubr v.0.2.5 (Kassambara

2020), ggsci v.2.9 (Xiao 2018), and the sf v.0.7-7 (Pebesma 2018), and scatterpie v.0.1.4 (Yu 2019) packages.

#### *Mixed-stock assignment of harvested walleye*

To determine the relative contribution of Lake Erie's walleye stock reporting groups to the east basin commercial and recreational harvest, 1,274 individuals of unknown origin were sampled during the spring, summer, and fall of the 2016-2018 fishing seasons (Table 2). For both commercial and recreational fishery samples, anglers and fishers reported the grid location of their harvest on a map provided by creel agents at boat ramps/docks. The total length (nearest 1 mm) of all sampled individuals was also recorded. To maximize our ability to estimate seasonal variance in stock structure, the majority of our mixed-stock samples (N=1,021, 80%) were taken from the 2017 harvest that occurred between May and December (Table 2). However, 130 samples were selected from the July 2016 commercial harvest and 122 from the July 2018 recreational harvest to investigate interannual variation. Fin clips were preserved in 95% ethanol until DNA extraction using Qiagen DNEasy 96 kits.

Rapture libraries for the mixed-stock analysis were constructed using identical procedures as baseline samples with one exception; to reduce the number of bait capture reactions necessary to genotype mixed-stock samples, DNA from two RAD-seq libraries were included in each bait capture reaction. All mixed-stock Rapture libraries were pooled and sequenced on four Illumina HighSeq4000 lanes at NovoGene (Sacramento, CA). Sequence data were processed in STACKS v.2.3 using identical procedures and filters as baseline samples with the exception of individual genotype rate, which was reduced to 50% because assignPOP indicated similar reassignment

accuracy between 100% and 50% retained loci tests. Once all individuals were genotyped, each was assigned to one of the three reporting groups identified during baseline analysis using the assign.x function in assignPOP using microhaplotype genotypes and the SVM parameter. As an additional estimate of mixed-stock proportion in 2017, the total stock mixture was estimated in rubias with identical reporting groups and reference as assignPOP and bootstrapped proportions (100 bootstraps). As confirmation of individual assignments, we also used rubias to estimate the stock mixture of all individuals assigned to each reporting group. Individual assignments were presumed to be correct as long as the genetic mixture of individuals in each reporting group exceeded 90%.

#### *Stock-specific harvest dynamics*

To determine how contributions to the eastern basin fisheries varied through space and time, all walleye that could be successfully assigned to one of our three reporting groups were summarized by year, month, location (commercial fishery or creel-grid centroid), and fishery type (recreational or commercial). For walleye sampled during 2017, the proportion of individuals assigned to each reporting group was calculated by dividing the number of walleye assigned to a particular reporting group by the total walleye assigned to any reporting group for a particular month, location, or fishery type. Because some locations contained a small number of assigned individuals in certain months, only month/location samples with greater than six successfully assigned walleye were included in these analyses ( $N = 48$  grid by month samples). Using the proportion of walleye assigned to the west/central basin reporting group as our response variable, we constructed linear models using only 2017 data with fishery type, month, and longitude as predictor variables. To better identify patterns in harvest, longitude was used in

place of grid identification numbers to orient positions along the east-west axis of Lake Erie. Estimates of stock proportion were also extrapolated to actual estimates of harvest (number of fish) for each fishery type, location, and month for 2017 (Walleye Task Group 2018). To describe inter-annual variation from 2016 to 2018, the proportional contributions of each reporting group during July was compared. Only samples collected during July 2016-2018 were used because July is consistently one of the peak months of walleye harvest in Lake Erie and we had a high number of available samples for all three years. These data were compared by calculating the proportion of individuals assigned to each reporting group relative to the total fish assigned to any reporting group that year. Because only a single month of samples was genotyped during 2016 and 2018, comparisons should be only considered as point estimates of proportional harvest.

Because the relative proportions of fish assigned to a particular reporting group throughout the season is most likely predicted by more than a single variable, we evaluated a series of seven generalized linear models (GLM) explaining the proportional harvest of walleye reporting groups. Information theoretic criteria (AIC<sub>c</sub>; Burnham and Anderson 2004) was used to identify a best model using available metadata for each fish: the type of fishery (commercial or recreational), average length of individuals sampled, latitude of capture, longitude of capture, and month of capture. All explanatory variables were first run independently against the response variable (proportion of walleye assigned to the west/central basin reporting group) and any variable not found to be significant ( $\alpha = 0.05$ ) was removed from future models. Once a subset of variables that were all independently significant were identified, all possible combinations of these variables were fit, and the best model was chosen based on the approach

outlined in Symonds and Moussalli (2011). In short, each model was fit in R using the `glm()` function using default settings (Gaussian error structure) and then ranked based on  $AIC_c$  using the `AICcmodavg` v.2.3 R package whereby the model with the lowest  $AIC_c$  was considered to be the best fit. Finally, the Evidence Ratio (ER) was calculated in the `qpcR` v.1.4-1 R package as a measure of how much more likely the best model was than other models.

## Results

### *Panel development*

Panel development sequencing produced a total of 61,350,730 retained reads and an average effective per-sample coverage of 11.4 (standard deviation = 4.0). Of the 263,723 putative SNPs and 43,884 loci identified in the 96 individuals sequenced for panel development, 14,418 loci passed our quality control filters and were sent to Arbor Biosciences for bait development. Of the 14,418 sequences sent, baits were successfully designed for 12,081.

### *Baseline stock structure*

Using 395 walleye collected from spawning sites during the spawning season, we described the population structure of walleye spawning stocks in Lake Erie and identified three putative reporting groups. Baseline Rapture sequencing used to identify reporting groups produced a total of 1,953,890,992 reads and an average effective per-sample coverage of 124.2 (standard deviation = 44.2). Of the 12,091 baited loci, 8,482 loci passed the genotyping rate (80%) and minor allele count filters (minor allele count > 3). Microhaplotypes at each locus had an average of 3.8 alleles, ranging from 1 to 21 alleles. Two samples contained genotypes at fewer than 70%



of loci and were removed prior to panel testing for a final total number of samples of 395 individuals and overall genotyping rate of 92%.

Three to four clusters of walleye spawning stocks that had similar allele frequencies were identified based on both pairwise  $F_{ST}$  (Table 3) and DAPC analyses (Fig. 1; Fig. 2). DAPC identified four distinct genetic clusters (Fig. 1, Fig. 2B). The first cluster identified by DAPC contained 142 out of 143 of the western basin and 13 out of 13 central basin walleye. The second cluster identified by DAPC contained 139 out of 142 individuals collected from 5 of the 6 eastern basin stocks sampled. Finally, individuals from Ontario's Grand River in the eastern basin formed the remaining two clusters with 24 individuals assigned to cluster 3, 56 individuals assigned to cluster 4. Additionally, 16 individuals were genetically similar to west/central basin individuals and assigned to cluster 1 (Fig. 2). This same pattern was apparent with pairwise  $F_{ST}$  values (Table 3) whereby pairwise  $F_{ST}$  was highest between Ontario's Grand River and all other sites in the eastern basin and all of the western basin sites (average  $F_{ST} = 0.032$ ). Pairwise  $F_{ST}$  was lower for all non-Ontario Grand River comparisons but higher between basins (mean pairwise  $F_{ST} = 0.007$ ) than within them (mean pairwise  $F_{ST} = 0.003$ ).

Reassignment accuracy to individual spawning stocks was low (< 50 – 75% for most stocks; Suppl. Fig 1). Thus, based on DAPC and pairwise  $F_{ST}$ , we decided to group spawning stocks into three reporting groups hereafter referred to as the west/central basin (containing all individuals from western basin and central basin stocks), east basin (containing individuals from 5 out of 6 eastern basin stocks), and Ontario's Grand River (containing individuals from the Ontario Grand River stock). When spawning stocks were classified into these three reporting groups, a

significant amount of variance was explained (AMOVA  $p = 0.01$ ; variance between reporting groups = 1.9%, variance between populations within reporting group = 0.2%). The average overall reassignment accuracy to the west/central basin, east basin, and Ontario Grand River reporting groups was estimated to be 93% (Fig. 3). Reassignment to each reporting group varied and was lowest for Ontario Grand River (mean = 85%), and more similar between the east basin and west/central basin reporting group (mean = 96% and mean = 99% respectively). Reassignment accuracy was similar for all sets of training individuals and when either 50% or 100% of loci were used (Fig. 3). Baseline assessment with rubias found similar reporting group accuracy as assignPOP based on posterior mean reporting group scores (Suppl. Fig. 2). Thus, we have high confidence in using our three reporting groups to identify the source origins of walleye harvested in eastern Lake Erie's fisheries.

#### *Assignment of mixed-stock individuals of unknown origin*

Rapture sequencing of the mixed-stock walleye samples of unknown origin produced a total of 3,331,974,311 retained reads and an average effective per-sample coverage of 41.8 (standard deviation = 24.3). Of the 12,081 baited loci, 8,482 loci passed genotyping rate and minor allele count filters and overlapped with loci used for baseline analysis. Of the 1,274 individuals analyzed, 199 (15%) failed to genotype in at least 50% of loci and were removed from analysis (Table 2). However, because the removed samples were spread across sampling dates and locations, we do not believe that their removal led to important sampling bias or was the result of consistent laboratory error. All individuals assigned to either west/central basin or east basin reporting groups and no individuals assigned to the Ontario Grand River reporting group. During 2017, the percentage of individuals of west/central basin origin was identical between individual

assignments determined in assignPOP and stock mixtures estimated in rubias (west/central basin = 54%; east basin = 46%; Ontario Grand River = 0%). Owing to the variable sample sizes among collections and similarity between assignPOP and rubias, we limited our subsequent analysis to individual assignments. The percentage of west/central basin origin walleye was almost identical between recreational and commercial fisheries (51% and 49%, respectively) indicating that both fisheries were equally supported by fish originating in west/central and east basin reporting groups ( $F = 0.1$ ;  $P = 0.71$ ; Fig. 4A).

Although about half of all fish harvested in the east basin were found to be of west/central basin origin, the percent of west/central basin origin walleye harvested varied spatially and temporally (Fig. 4B; Fig. 5). During the spring, the proportion of west/central basin origin walleye in the harvest was low with increasing contributions from this reporting group throughout the summer ( $F = 25.1$ ;  $P < 0.01$ ). For example, 6% of walleye were of west/central basin origin in May 2017, which increased to 31% by the end of June 2017 (Fig. 4B). The presence of walleye of west/central basin origin in the commercial fishery remained high throughout the summer and into the fall (July – October average percent of west/central basin origin = 57%) Harvest composition also varied spatially, with more easterly sites having fewer individuals of west/central basin origin than more westwardly ones ( $F = 24.0$ ;  $P < 0.01$ ; Fig. 4C).

Based on examination of samples collected during July 2016-2018 harvest composition varied among years (Table 2). The percentage of west/central basin fish in the July 2016 recreational fishery was 20%, which was smaller than the percentage of west/central basin origin fish harvested in the recreational (54%) or commercial (38%) fisheries during 2017. However, during

July 2018 the percentage of walleye of west/central basin origin in the commercial harvest was 90%, which was higher than either the commercial or recreational harvest during July the previous years.

#### *Predicting contributions from the west/central basins*

Modeling showed that the proportional contribution of the western/central basin spawning stocks to the eastern basin fisheries in 2017 varied both seasonally and spatially. The candidate model with the highest level of support in our AIC<sub>c</sub> contained three of the five potential predictors (longitude, latitude, and month), and had an evidence ratio that indicated it was 10 times more likely than the next best model (Table 4; AIC<sub>c</sub> = -5.5; degrees freedom = 47; Intercept = 2.8; latitude = -0.74, longitude = -0.37, date = 0.05). Fishery type and mean total length were not significant predictors of proportion of west/central basin fish, and so were removed prior to AIC<sub>c</sub> analysis. All other variables appeared to explain a significant amount of the variance and were included in our AIC<sub>c</sub> assessment. Longitude was found to be the most important predictor of proportion of west/central basin walleye with proportional contributions of west/central basin walleye in harvest declining from west to east (Fig. 4C), however date and latitude were also important based on ER and delta AIC<sub>c</sub> with proportional contributions of west/central basin walleye increasing from spring to fall and decreasing from North to South.

#### **Discussion**

Our study provides the first successful attempt to use a sequencing-based genetic panel to determine the relative contributions of local spawning stocks to harvest in the Great Lakes.

Walleye of unknown origin harvested by commercial and recreational fisheries in eastern Lake Erie were assigned to one of three reporting groups with high accuracy using a Rapture panel and genotype data from 8,482 microhaplotype markers. Results supported the expectation that walleye from the productive west/central basin spawning stocks contributed substantially to the eastern basin fisheries for much of the year. Even so, seasonal and spatial variance in the proportion of walleye originating from the west/central basin indicated that the smaller eastern basin stocks (with the exception of the Ontario Grand River) comprised a large portion of harvest during certain times, such as the spring, or in certain parts of the basin, such as more easterly parts of the basin. In addition to providing critical information on stock-specific harvest of walleye in the eastern basin of Lake Erie that can benefit fishery management, our study represents one of the first uses of Rapture data for mixed-stock assignment in fisheries (but see Carrier et al. 2020) and the first in a freshwater ecosystem.

#### *Population structure and reassignment accuracy*

By using microhaplotype genotypes at 8,482 loci, we were able to accurately (> 95%) assign fish of unknown origin to one of three reporting groups of walleye spawning stocks in Lake Erie. Our high assignment accuracy is an improvement relative to previous attempts at differentiating Lake Erie stocks of walleye using microsatellite loci (Stepien et al. 2012; Brenden et al. 2015) or mitochondrial haplotypes (Stepien and Faber 1998; Gatt et al. 2004; Haponski et al. 2014). Unlike our approach, which aggregated spawning stocks with low genetic differentiation, these earlier studies treated each spawning stock independently (Strange and Stepien 2007; Stepien et al. 2012), and often included only a subset of known stocks within a given basin (Brenden et al. 2015; Chen et al. 2019). In turn, while microsatellite loci and mitochondrial haplotypes identified

significant genetic differences between the west and east basins, the reassignment accuracy to individual stocks in these studies was generally low (<25%) and inconsistent between western basin versus eastern basin reporting groups (20 – 87%). When taken in context with these previous studies, our results indicate that walleye of unknown spawning stock origin can accurately be assigned to one of three reporting groups (i.e., western/central basin, eastern basin, and Ontario Grand River) but not to a single spawning stock within a reporting group. The high reassignment accuracy in our study can be attributed to the increased statistical power attained by genotyping thousands of genomic markers and by conducting a comprehensive baseline assessment of lake-wide spawning stocks. Both of these factors made it possible to distinguish among weakly structured spawning stocks and provide a realistic representation of the interannual movement and mixture of walleye stocks in the eastern basin of Lake Erie during our study.

#### *Stock contributions to eastern basin fisheries*

Walleye from the west/central basin substantially contributed to the recreational and commercial harvest in the eastern basin, although their contribution varied among years, seasons, and locations. Walleye from the western basin are known to migrate seasonally towards the eastern basin after spawning (Zhao et al. 2011; Vandergoot and Brenden 2014; Matley et al. 2020), but the extent to which these migrations influenced the actual harvest of walleye was previously unknown. Because western basin walleye spawning stocks are highly productive (DuFour et al. 2015) and have been tracked migrating to the eastern basin (Zhao et al. 2011; Matley et al. 2020), we were not surprised that western basin walleye contribute to eastern basin harvests. Importantly, our first quantification of seasonal, annual, and spatial variation in the contributions

of western basin migrants to eastern basin fisheries supports previous evidence that walleye stocks can be treated as a portfolio whereby both eastern and western basin spawning stocks contribute to lake-wide harvest opportunities (DuFour et al. 2015). Further, these results provide an important first step toward a lake-wide approach to assessment and management of Lake Erie walleye, which was identified as a key priority in the current walleye management plan (Kayle et al. 2015).

The proportion of walleye of west/central basin origin caught in the eastern basin was low during the spring fishing season and increased rapidly following spring spawning period in April. By early summer (June – July) west/central basin walleye already made up nearly 50% of the harvest in our 2017 samples, and by August, the majority of genotyped fish in both the recreational and commercial harvest were of west/central basin origin. This finding is consistent with recent findings for large-scale acoustic telemetry studies examining annual migration patterns of Lake Erie walleye (Matley et al. 2020). Specifically, walleye from the western basin began being detected at central and eastern basin sites in higher proportions in June and continued to be detected throughout the summer (Matley et al. 2020). These annual movements have been hypothesized to be associated with such as thermoregulation (Raby et al. 2018) and prey availability (Kershner et al. 1999). Therefore, interannual variation in water temperature could influence the timing of spawning and subsequent eastward migration of west basin walleye which in turn could drive harvest dynamics (Dippold et al 2020). Similarly, changes in abundance of preferred forage fishes (e.g., emerald shiner *Notropis atherinoides*, rainbow smelt *Osermerus mordax*) may also influence migratory tendencies of western basin walleye.

During our study, the proportion of west/central basin walleye harvested in the eastern basin during July varied from 20% in 2016 to 90% in 2018 across both fisheries (a 4.5-fold change), indicating interannual variation in stock composition can be high. This variation may be in part driven by stock-specific behavior and productivity. For instance, western basin walleye stocks more strongly selected depths > 13 m than did their eastern basin counterparts, suggesting that there could be potential for differential exposure to commercial and recreational fisheries (Matley et al. 2020). Similarly, in years when eastward migration is either delayed or limited, eastern basin stocks may be exploited to a higher degree (Dippold et al. 2020). Exploitation of individual stocks may also vary among years owing to differences in stock productivity and recruitment success for eastern and western basin stocks. For example, no fish in our mixed-stock analysis were assigned to the Ontario Grand River. However, this stock has been suggested to substantially contribute to eastern basin harvest in other years (C. Wilson, *unpublished data*). Therefore, periodic (e.g., every 3-5 years) reassessments of stock contributions to the harvest are likely necessary to allow management agencies to characterize spatio-temporal variation in relative stock contributions.

Even when walleye of west/central basin origin were present in the eastern basin, their contribution to harvest was spatially variable, with higher proportions of these individuals being recovered in more westerly areas of the eastern basin relative to more easterly sites. This trend was especially clear during June and July, where west/central origin fish made up 50 – 75% of assigned individuals in our westerly samples, whereas this percentage was closer to 25% in easterly samples (Fig. 5). This finding is also in accordance with previous studies of walleye movement, where the number of western basin walleye detected decreased with distance from



the western basin (Raby et al. 2018; Matley et al. 2020). Month and longitude of harvest alone described over 60% of the variance in stock proportions during 2017, suggesting that the contribution of western basin walleye to the eastern basin harvest is non-random and that these movements should be considered in management of walleye. Understanding the migration patterns and environmental predictors of stock structure have proven crucial to defining harvest quotas in both Pacific and Atlantic salmon populations (Shaklee et al. 1999; Vähä et al. 2011; Bradbury et al. 2016). Once a baseline of genetic connectivity and movement is established, assessments of mixed-stock composition can help identify interannual shifts in abundance and help managers anticipate changes in stock abundance. Our findings support previous work documenting movement of western basin walleye into the eastern basin (e.g., Zhao et al. 2011; Matley et al. 2020) and advanced understanding by providing a baseline estimate of harvest composition in the eastern basin.

Eastern basin walleye stocks clearly do have an influence on eastern basin fishing opportunities and contribute to the stability of the Lake Erie walleye population year-to-year. For example, because the presence of west/central basin walleye decreases from west to east (see Fig. 4), recreational walleye fishing in the eastern most portion of the eastern basin is likely supported primarily by walleye of east basin origin for much of the year. Additionally, an estimated 11,694 walleye were harvested by commercial fishers in the east basin during May 2017 at a time when greater than 90% of walleye that we sampled originated from the east basin. During June, the estimated catch rose to 76,579, 40% of which may have originated from the smaller, less productive east basin spawning stocks (Suppl. Fig. 3). These simple extrapolations illustrate the utility of mixed-stock analysis for providing important information that can be used to protect

small and vulnerable stocks such as those found in the eastern basin of Lake Erie. Further, understanding how and when smaller stocks are exploited in high numbers makes targeted action to limit overharvest possible (Dann et al. 2013), thus offering a means to help maintain a diverse portfolio of stocks that we expect to be critical for sustaining fisheries production in the face of continued anthropogenic change.

No walleye originating from the Ontario Grand River reporting group were identified in either fishery. While this finding could be real, the possibility also exists that either we failed to collect Ontario Grand River walleye in our samples or that Ontario Grand River individuals were misassigned in our analysis. The estimated reassignment accuracy of Ontario Grand River walleye was the lowest of our three reporting groups but was still high (87%) relative to other genetic studies of walleye in Lake Erie (Johnson et al. 2004). The variability in our assignment accuracy of the Ontario Grand River reporting group can be attributed to the 16 baseline individuals sampled from the Ontario Grand River that were genetically similar to the west/central basin reporting group. Therefore, any fish sampled that was genetically similar to the majority of Ontario Grand River walleye used in the baseline would most likely have been successfully assigned to this source stock, suggesting that misassignment is not the cause of failed detection. However, some behavioral evidence exists that suggest that walleye from Ontario's Grand River remain close to the northern shore of Lake Erie throughout the summer, and do not mix evenly with walleye from the west/central and east basin reporting groups (TM, *unpublished data*; Matley et al. 2020). Therefore, our sample collections from commercial fisheries may not have included Ontario Grand River walleye. Additional assignment studies focused on sampling the Ontario Grand River reporting group could help to determine why this particular stock is unique

from the rest of the Lake Erie walleye stocks, and help to verify if indeed the Ontario Grand River spawning stock does not significantly support the east basin recreational and commercial walleye harvest.

# *Implications for inland fisheries conservation and management*

By developing a Rapture panel, we were able to genotype a large number of individuals (1,671) at a consistent set of markers, which provided a clearer description of walleye genetic stock structure than was previously possible. The increased diagnostic power of our Rapture panel also allowed us to achieve higher confidence in individual assignments of natal origin, which is crucial for obtaining unbiased estimates of harvest proportions. When assignment accuracy is low, using individual assignments rather than mixture analysis for mixed-stock analysis can bias mixture estimates (Manel et al. 2005). However, the mixture proportions estimated based on our individual assignments were very similar to those estimated with traditional mixture analysis. Therefore, the ability to use thousands of loci for population assignment allowed us to obtain a high resolution and specificity of individual assignment without losing the precision of mixture analysis. Previous studies have shown that individual assignments can provide similar stock composition estimates as mixture analysis, but genetic differentiation among populations in these studies was larger than our study (Potvin and Bernatchez 2001). Indeed, when genetic differentiation is low, having a high number of genetic markers becomes more important for assignment accuracy (Waples and Gaggiotti 2006; Larson et al. 2014; Benestan et al. 2015). Genotyping-by-sequencing methods, like Rapture, are a feasible way to obtain this level of diagnostic power in a non-model species and can provide reliable mixture estimates. Once stock differences are identified, markers from high-density marker panels can be filtered to identify

smaller subsets of highly-informative markers that can be incorporated into even higher throughput genotyping-in-thousands (GT-seq) panels for long-term monitoring and annual mixed-stock assessments (Bootsma et al. 2020). For this reason, we are optimistic that our approach will offer a new means for agencies managing fisheries with weak population structure to discern local spawning stocks such that their relative contributions can be quantified.

Our study provides the first estimate of proportional harvest of eastern and western basin walleye stocks, and therefore has direct implications for the sustainable management of this ecologically and economically important population (Hatch et al. 1987; Kayle et al. 2015). The baseline assessment of spawning stocks that we used to identify reporting groups (putative source stocks) refined previous assessments of walleye connectivity among spawning sites and helped to describe a third reporting group, the Ontario Grand River population, which had previously been suggested to be diverged from other eastern basin stocks (MacDougall et al. 2007). Subsequent mixed-stock assessments in the eastern basin identified clear patterns in harvest, helping to illuminate where and when different stocks were being exploited that would have otherwise been unattainable. Our results indicate that similar approaches could be used successfully in other ecosystems to obtain a snapshot of stock structure and exploitation without intensive up-front research. Harvest from rivers and lakes has been increasing worldwide (Fluet-Chouinard et al. 2018) but a lack of basic knowledge exists about many of these fisheries precludes the determination of necessary regulations and development of sustainable management plans (Allan et al. 2005; Bower et al. 2020). Determinations of relative stock contributions, like the one we present here, have the potential to inform sustainable harvest in many of these fisheries by identifying over- versus under-exploited stock components (Beard et al. 2011). This knowledge

could open up the opportunity to identify stock components of other populations being stabilized by the portfolio effect and predict the sustainability of these populations long-term.

Genomics is emerging as an influential tool for managing and conserving freshwater fishes globally (Tibihika et al. 2020; Biesack et al. 2020). In Africa and Asia, genomics has become an important tool to quantify diversity and identify critically endangered populations (Malinsky et al. 2018; Ackiss et al. 2019). In Europe, genomics has become central to domestication of native species for aquaculture (Toomey et al. 2020). In the Laurentian Great Lakes, researchers have made substantial advances in the last decade to incorporate molecular techniques into the management and conservation native communities including lake trout *Salvelinus namaycush* (Bernatchez et al. 2016), brook trout *Salvelinus fontinalis* (Elias et al. 2018), coregonines (Ackiss et al. 2020), and walleye (Chen et al. 2019). Much of this success can be attributed to increased access to molecular resources for non-model organisms and reduced sequencing costs that have made large-scale genomic assessment feasible for freshwater fisheries that often lack resources (Lynch et al. 2016). With the promise that genomic tools hold, we are hopeful that molecular studies of exploited freshwater populations can move beyond describing the extent of genetic diversity present and population structure in the system and begin being used to consistently monitor contemporary changes in population structure and microevolution in response to anthropogenic change.

Our study highlights the ability for genomic techniques to be applied directly to active fisheries and to quantify contributions of genetically distinct units to harvest, even those that are weakly differentiated. Without substantial prior genomic resources, we successfully completed all steps

of a mixed-stock analysis that facilitated the quantification of stock contributions to fisheries harvest and identified predictors of stock composition of a large fishery in a single study. Genotyping-by-sequencing approaches allowed us to identify informative markers in a non-model organism while the construction of a Rapture panel enabled the consistent genotyping of thousands of SNPs across a large number of individuals. Our approach offers a pathway to obtain a snapshot of mixed-stock composition that has the potential to inform stock connectivity and temporal and spatial changes in harvest for data-poor fisheries with weak population structure where an understanding of stock structure and relative contributions to the fishery has not been feasible but is needed (Allan et al. 2005; Irvine et al. 2019). In this way, we are confident that the continued use of genomic approaches, like the one we demonstrated with Lake Erie walleye, will help untangle the spatio-temporal contributions of spawning stock portfolios to fishery harvest.

## Acknowledgements

We thank Michael Sovic for his initial work writing the grant and developing pilot data, Kristen Gruenthal, Tina Werner, and Kevin Smith for help processing tissue samples and conducting lab work, and Zac Driscoll and Zachary Feiner for help with statistical analysis. This work was funded by the Ohio Sea Grant College Program (#NA18OAR417100 to SAL) and research was supported by the Research Computing clusters at Old Dominion University. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. There is no conflict of interest declared in this article.

## Data Archiving Statement

The data that support the findings of this study are openly available in Geome at [https://geome-](https://geome-db.org/workbench/project-overview?projectId=190)  
[db.org/workbench/project-overview?projectId=190](https://doi.org/10.5061/dryad.4b8gthtb2) and Dryad at  
<https://doi.org/10.5061/dryad.4b8gthtb2>. All raw sequence files will be uploaded and stored on  
the NCBI Short Read Archive upon manuscript acceptance.

## Literature Cited

- Ackiss, A. S., B. T. Dang, C. E. Bird, E. E. Biesack, P. Chheng, L. Phounvisouk, Q. H. D. Vu, S. Uy, and K. E. Carpenter. (2019). Cryptic lineages and a population dammed to incipient extinction? Insights into the genetic structure of a Mekong River catfish. *Journal of Heredity*, 110(5), 535–547.
- Ackiss, A. S., W. A. Larson, and W. Stott. (2020). Genotyping-by-sequencing illuminates high levels of divergence among sympatric forms of coregonines in the Laurentian Great Lakes. *Evolutionary Applications*, 13(5), 1037–1054.
- Ali, O. A., S. M. O’Rourke, S. J. Amish, M.H. Meek, G. Luikart, C. Jeffres, and M. R. Miller. (2016). Rad capture (Rapture): Flexible and efficient sequence-based genotyping. *Genetics*, 202(2), 389–400.
- Allan, J. D., R. Abell, Z. Hogan, C. Revenga, B. W. Taylor, R. L. Welcomme, and K. Winemiller. (2005). Overfishing of inland waters. *BioScience*, 55(12), 1041.
- Alves, J. A., P. M. Lourenço, T. Piersma, W. J. Sutherland, and J. A. Gill. (2010). Population overlap and habitat segregation in wintering Black-tailed Godwits *Limosa limosa*. *Bird Study*, 57(3), 381–391.

- Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. *Molecular Ecology Resources*, 10(4), 701–710.
- Anderson, E. C., R. S. Waples, and S. T. Kalinowski. (2008). An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(7), 1475–1486.
- Andvik, R. T., B. L. Sloss, J. A. VanDeHey, R. M. Claramunt, S. P. Hansen, and D. A. Isermann. (2016). Mixed stock analysis of Lake Michigan's lake whitefish *Coregonus clupeaformis* commercial fishery. *Journal of Great Lakes Research*, 42, 660–667.
- Baetscher, D. S., A. J. Clemento, T. C. Ng, E. C. Anderson, and J. C. Garza. (2018). Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, 18(2), 296–305.
- Beacham, T. D., C. Wallace, K. Jonsen, B. McIntosh, J. R. Candy, E. B. Rondeau, J.-S. Moore, L. Bernatchez, and R. E. Withler. (2020). Accurate estimation of conservation unit contribution to coho salmon mixed-stock fisheries in British Columbia, Canada, using direct DNA sequencing for single nucleotide polymorphisms. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(8), 1–14.
- Beard, T. D., R. Arlinghaus, S. J. Cooke, P. B. McIntyre, S. De Silva, D. Bartley, and I. G. Cowx. (2011). Ecosystem approach to inland fisheries: research needs and implementation strategies. *Biology Letters*, 7(4), 481–483.
- Benestan, L., T. Gosselin, C. Perrier, B. Sainte-Marie, R. Rochette, and L. Bernatchez. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population



assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Molecular Ecology*, 24(13), 3299–3315.

Bernatchez, S., M. Laporte, C. Perrier, P. Sirois, and L. Bernatchez. (2016). Investigating genomic and phenotypic parallelism between piscivorous and planktivorous lake trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses. *Molecular Ecology*, 25(19), 4773–4792.

Bernatchez, L., M. Wellenreuther, C. Araneda, D. T. Ashton, J. M. I. Barth, T. D. Beacham, G. E. Maes, J. T. Martinsohn, K. M. Miller, K. A. Naish, J. R. Ovenden, C. R. Primmer, H. Young Suk, N. O. Therkildsen, and R. E. Withler. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology and Evolution*, 32(9), 665–680.

Bernatchez, L., & Wilson, C. C. (1998). Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, 7(4), 431–452.

Biesack, E. E., B. T. Dang, A. S. Ackiss, C. E. Bird, P. Chheng, L. Phounvisouk, O. T. Truong, and K. E. Carpenter. (2020). Evidence for population genetic structure in two exploited Mekong River fishes across a natural riverine barrier. *Journal of Fish Biology*, 97(3), 696–707.

Bootsma, M. L., K. M. Gruenthal, G. J. McKinney, L. Simmons, L. Miller, G. G. Sass, and W. A. Larson. (2020). A GT-seq panel for walleye (*Sander vitreus*) provides important insights for efficient development and implementation of amplicon panels in non-model organisms. *Molecular Ecology Resources*, (in press). <https://doi.org/10.1111/1755-0998.13226>.

Bower, S. D., Ø. Aas, R. Arlinghaus, T. Douglas Beard, I. G. Cowx, A. J. Danylchuk, K. M. F. Freire, W. M. Potts, S. G. Sutton, and S. J. Cooke. (2020). Knowledge gaps and

management priorities for recreational fisheries in the developing world. *Reviews in Fisheries Science & Aquaculture*, 28(4), 518-535.

<https://doi.org/10.1080/23308249.2020.1770689>

Bradbury, I. R., L. C. Hamilton, G. Chaput, M. J. Robertson, H. Goraguer, A. Walsh, V. Morris, D. Reddin, J. B. Dempson, T. F. Sheehan, T. King, and L. Bernatchez. (2016). Genetic mixed stock analysis of an interceptory Atlantic salmon fishery in the Northwest Atlantic. *Fisheries Research*, 174, 234–244.

Brenden, T. O., K. T. Scribner, J. R. Bence, I. Tsehay, J. Kanefsky, C. S. Vandergoot, and D. G. Fielder. (2015). Contributions of Lake Erie and Lake St. Clair Walleye populations to the Saginaw Bay, Lake Huron, recreational fishery: evidence from genetic stock identification. *North American Journal of Fisheries Management*, 35(3), 567–577.

Carrier, E., A. A. Ferchaud, E. Normandeau, P. Sirois, and L. Bernatchez. (2020). Estimating the contribution of Greenland Halibut (*Reinhardtius hippoglossoides*) stocks to nurseries by means of genotyping-by-sequencing: Sex and time matter. *Evolutionary Applications*, 13(9), 2155-2167. <https://doi.org/10.1111/eva.12979>

Chen, K.-Y., S. A. Ludsin, M. M. Corey, P. D. Collingsworth, M. K. Nims, J. W. Olesik, K. Dabrowski, J. J. van Tassell, and E. A. Marschall. (2017). Experimental and field evaluation of otolith strontium as a marker to discriminate between river-spawning populations of walleye in Lake Erie. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(5), 693–701.

Chen, K. Y., P. T. Euclide, S. A. Ludsin, W. A. Larson, M. G. Sovic, H. L. Gibbs, and E. A. Marschall. (2020). RAD-Seq refines previous estimates of genetic structure in Lake Erie Walleye. *Transactions of the American Fisheries Society*, 149(2), 159-173.

- Chen, K. Y., E. A. Marschall, M. G. Sovic, A. C. Fries, H. L. Gibbs, and S. A. Ludsin. (2018). assignPOP: An r package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution*, 9(2), 439–446.
- Chen, K-Y., S.A. Ludsin, B.J. Marcek, J.W. Olesik, and E.A. Marschall. (2020). Otolith microchemistry shows natal philopatry of walleye in western Lake Erie. *Journal of Great Lakes Research*, 46(5), 1349-1357.
- Cowen, R. K., G. Gawarkiewicz, J. Pineda, S. R. Thorrold, and F. E. Werner. (2007). Population connectivity in marine systems: An overview. *Oceanography*, 20(SPL.ISS. 3):14–21.
- Dahle, G., T. Johansen, J. I. Westgaard, A. Aglen, and K. A. Glover. (2018). Genetic management of mixed-stock fisheries “real-time”: the case of the largest remaining cod fishery operating in the Atlantic in 2007–2017. *Fisheries Research*, 205, 77–85.
- Dann, T. H., C. Habicht, T. T. Baker, J. E. Seeb, and D. Fraser. (2013). Exploiting genetic diversity to balance conservation and harvest of migratory salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(5), 785–793.
- Dean, M. J., S. P. Elzey, W. S. Hoffman, N. C. Buchan, J. H. Grabowski, and A. Kuparinen. (2019). The relative importance of sub-populations to the Gulf of Maine stock of Atlantic cod. *ICES Journal of Marine Science*, 76(6), 1626–1640.
- Dippold, D. A., G. D. Adams, and S. A. Ludsin. (2020). Spatial patterning of walleye recreational harvest in Lake Erie: role of demographic and environmental factors. *Fisheries Research*, 230 (in press).<https://doi.org/10.1016/j.fishres.2020.105676>
- DuFour, M. R., C. J. May, E. F. Roseman, S. A. Ludsin, C. S. Vandergoot, J. J. Pritt, M. E. Fraker, J. J. Davis, J. T. Tyson, J. G. Miner, E. A. Marschall, and C. M. Mayer. (2015).

Portfolio theory as a management tool to guide conservation and restoration of multi-stock fish populations. *Ecosphere*, 6(12), art296–art296.

Elias, A., R. McLaughlin, R. Mackereth, C. Wilson, and K. M. Nichols. (2018). Population structure and genomic variation of ecological life history diversity in wild-caught Lake Superior brook trout, *Salvelinus fontinalis*. *Journal of Great Lakes Research*, 44(6), 1373–1382.

Embke, H.S., T.D. Beard Jr., A.J. Lynch, and M.J. Vander Zanden. (2020). Fishing for food: quantifying recreational fisheries harvest in Wisconsin lakes. *Fisheries*, (in press).

Feron, R., M. Zahm, C. Cabau, C. Klopp, C. Roques, O. Bouchex, C. Eche, S. Valiere, C. Donnadieu, P. Haffray, A. Bestin, R. Morvezen, H. Acloque, P. T. Euclide, M. Wen, E. Jouano, M. Scharl, J. H. Postlethwait, C. Schraidt, M. R. Christie, W. A. Larson, A. Herpin, and Y. Guiguen. 2019. 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.

Figge, F. (2004). Bio-folio: applying portfolio theory to biodiversity. *Biodiversity and Conservation*, 13, 827-849.

Fluet-Chouinard, E., S. Funge-Smith, and P. B. McIntyre. (2018). Global hidden harvest of freshwater fish revealed by household surveys. *Proceedings of the National Academy of Sciences of the United States of America*, 115(29), 7623–7628.

Gatt, M. H., T. L. McParland, L. C. Halyk, and M. M. Ferguson. (2004). Mitochondrial DNA variation and mixed-stock analysis of recreational and commercial walleye fisheries in eastern Lake Erie. *North American Journal of Fisheries Management*, 23(2), 431–440.

825 Haponski, A. E., H. Dean, B. E. Blake, and C. A. Stepien. (2014) Genetic history of Walleyes  
826 spawning in Lake Erie's Cattaraugus Creek: a comparison of pre- and post- stocking.  
827 *Transactions of the American Fisheries Society*, 143(5), 1295-1307.

828 Hatch, R. W., Nepszy, S. J., Muth, K. M., & Baker, C. T. (1987). Dynamics of the recovery of  
829 the western Lake Erie walleye (*Stizostedion vitreum vitreum*) stock. *Canadian Journal of*  
830 *Fisheries and Aquatic Sciences*, 44(Suppl.2), 15–22. <https://doi.org/10.1139/f87-305>

831 Hedges, K. J. (2002). Use of calcified structures for stock discrimination in Great Lakes walleye  
832 (*Stizostedion vitreum*). Master's thesis. University of Windsor, Ontario, Canada.

833 Hilborn, R. T.P. Quinn, D.E. Schindler, and D.E. Rogers. (2003). Biocomplexity and fisheries  
834 sustainability. *Proceedings of the National Academy of Sciences*, 100, 6564-6568.

835 Holcombe, T. L., L. A. Taylor, D. F. Reid, J. S. Warren, P. A. Vincent, and C. E. Herdendorf.  
836 (2003). Revised Lake Erie postglacial lake level history based on new detailed bathymetry.  
837 *Journal of Great Lakes Research*, 29(4), 681–704.

838 Hutchinson, W. F. (2008). The dangers of ignoring stock complexity in fishery management: the  
839 case of the North Sea cod. *Biology Letters*, 4(6), 693–695.

840 Irvine, K., C. A. Etiegni, and O. L. F. Weyl. (2019). Prognosis for long-term sustainable fisheries  
841 in the African Great Lakes. *Fisheries Management and Ecology*, 26(5), 413–425.

842 Johnson, T. B., B. Dixon, C. A. Stepien, and C. C. Wilson. (2004). Stock discrimination of Lake  
843 Erie walleye: a mixed stock analysis contrasting genetic techniques. *Great Lake Fishery*  
844 *Commission completion report*. Available from  
845 [http://www.glfc.org/lakecom/lec/WTG\\_docs/annual\\_reports/WTG\\_report\\_2006.pdf](http://www.glfc.org/lakecom/lec/WTG_docs/annual_reports/WTG_report_2006.pdf).

846 Jombart, T. (2008). Adegnet: A R package for the multivariate analysis of genetic markers.  
847 *Bioinformatics*, 24(11), 1403–1405.

- Jorde, P. E., A. R. Kleiven, M. Sodeland, E. M. Olsen, K. Ferter, S. Jentoft, H. Knutsen, and W. S. Grant. (2018). Who is fishing on what stock: population-of-origin of individual cod (*Gadus morhua*) in commercial and recreational fisheries. *ICES Journal of Marine Science* 75(6):2153–2162.
- Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281.
- Kassambara, A. 2020. ggpubr: “ggplot2” Based Publication Ready Plots.
- Kayle, K., Francis, J., Murray, C., & Markham, J. (2015). Lake Erie Walleye management plan (2015–2019). Great Lakes Fishery Commission, Ann Arbor, Michigan.
- Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, and P. A. Prodöhl. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4(8), 782–788.
- Larson, W. A., L. W. Seeb, M. V. Everett, R. K. Waples, W. D. Templin, and J. E. Seeb. (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, 7(3), 355–369.
- Layton, K. K. S., B. Dempson, P. V. R. Snelgrove, S. J. Duffy, A. M. Messmer, I. G. Paterson, N. W. Jeffery, T. Kess, J. B. Horne, S. J. Salisbury, D. E. Ruzzante, P. Bentzen, D. Côté, C. M. Nugent, M. M. Ferguson, J. S. Leong, B. F. Koop, and I. R. Bradbury. (2020). Resolving fine-scale population structure and fishery exploitation using sequenced microsatellites in a northern fish. *Evolutionary Applications*, 13(5), 1055-1068.

869 Li, Y., J. R. Bence, and T. O. Brenden. (2015). An evaluation of alternative assessment  
870 approaches for intermixing fish populations: a case study with Great Lakes lake whitefish.  
871 *ICES Journal of Marine Science: Journal du Conseil* 72, 70–81.

872 Ludsin, S. A., De Vanna, K. M., & Smith, R. E. H. (2014). Physical-biological coupling and the  
873 challenge of understanding fish recruitment in freshwater lakes. *Canadian Journal of*  
874 *Fisheries and Aquatic Sciences*, 71(5), 775–794. <https://doi.org/10.1139/cjfas-2013-0512>

875 Lynch, A. J., S. J. Cooke, A. M. Deines, S. D. Bower, D. B. Bunnell, I. G. Cowx, V. M. Nguyen,  
876 J. Nohner, K. Phouthavong, B. Riley, M. W. Rogers, W. W. Taylor, W. Woelmer, S. J.  
877 Youn, and T. D. Beard. (2016). The social, economic, and environmental importance of  
878 inland fish and fisheries. *Environmental Reviews*, 24(2), 115–121.

879 MacDougall, T. M., Wilson, C. C., Richardson, L. M., Lavender, M., & Ryan, P. A. (2007).  
880 Walleye in the Grand River, Ontario: an overview of rehabilitation efforts, their  
881 effectiveness, and implications for eastern Lake Erie fisheries. *Journal of Great Lakes*  
882 *Research*, 33(1), 103–117.

883 Malinsky, M., H. Svardal, A. M. Tyers, E. A. Miska, M. J. Genner, G. F. Turner, and R. Durbin.  
884 (2018). Whole-genome sequences of Malawi cichlids reveal multiple radiations  
885 interconnected by gene flow. *Nature Ecology and Evolution*, 2(12), 1940–1955.

886 Manel, S., O. E. Gaggiotti, and R. S. Waples. (2005). Assignment methods: Matching biological  
887 questions with appropriate techniques. *Trends in Ecology & Evolution*, 20(3), 136–142.

888 Martinez, A. S., J. R. Willoughby, and M. R. Christie. (2018). Genetic diversity in fishes is  
889 influenced by habitat type and life-history variation. *Ecology and Evolution*, 8(23), 12022–  
890 12031.

Matley, J. K., M. D. Faust, G. D. Raby, Y. Zhao, J. Robinson, T. MacDougall, T. A. Hayden, A. T. Fisk, C. S. Vandergoot, and C. C. Krueger. (2020). Seasonal habitat-use differences among Lake Erie's walleye stocks. *Journal of Great Lakes Research*, 46(3), 609-621.

McKinney, G. J., J. E. Seeb, and L. W. Seeb. (2017a). Managing mixed-stock fisheries: genotyping multi-SNP haplotypes increases power for genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(4), 429-434.

McKinney, G. J., R. K. Waples, L. W. Seeb, and J. E. Seeb. (2017b). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources*, 17(4), 656-669.

Meek, M. H., and W. A. Larson. (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources*, 19(4), 795-803.

Milner, G. B., D. J. Teel, F. M. Utter, and G. A. Winans. (2008). A genetic method of stock identification in mixed populations of Pacific salmon, *Oncorhynchus* spp. *Marine Fisheries Review*, 47(1), 1-8.

Moran, B. M., and E. C. Anderson. (2019). Bayesian inference from the conditional genetic stock identification model. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(4), 551-560.

Pebesma, E. (2018). Simple features for R: standardized support for spatial vector data. *The R Journal* 10(1):439.

Post, J. R., M. Sullivan, S. Cox, N. P. Lester, C. J. Walters, E. A. Parkinson, A. J. Paul, L. Jackson, and B. J. Shuter. (2002). Canada's recreational fisheries: the invisible collapse? *Fisheries*, 27(1), 6-17.



913 Potvin, C., and L. Bernatchez. (2001). Lacustrine spatial distribution of landlocked Atlantic  
 914 salmon populations assessed across generations by multilocus individual assignment and  
 915 mixed-stock analyses. *Molecular Ecology*, 10(10), 2375–2388.

916 R Core Team. (2019). R: A language and environment for statistical computing. R Foundation  
 917 for Statistical Computing, Vienna, Austria.

918 Raby, G. D., C. S. Vandergoot, T. A. Hayden, M. D. Faust, R. T. Kraus, J. M. Dettmers, S. J.  
 919 Cooke, Y. Zhao, A. T. Fisk, and C. C. Krueger. (2018). Does behavioral thermoregulation  
 920 underlie seasonal movements in Lake Erie walleye? *Canadian Journal of Fisheries and*  
 921 *Aquatic Sciences*, 75(3), 1–9.

922 Regier, H.A., and W.L. Hartman. (1973). Lake Erie's fish community: 150 years of cultural  
 923 stresses. *Science*, 180, 1248-1255.

924 Riley, L. M., and R. F. Carline. (1982). Evaluation of scale shape for the identification of walleye  
 925 stocks from western Lake Erie. *Transactions of the American Fisheries Society*, 111, 736–  
 926 741.

927 Rochette, N. C., A. G. Rivera-Colón, and J. M. Catchen. (2019). Stacks 2: Analytical methods for  
 928 paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*,  
 929 28(21), 4737–4754.

930 Roseman, E. F., Thompson, P. A., Farrell, J. M., Mandrak, N. E., & Stepien, C. A. (2014).  
 931 Conservation and management of fisheries and aquatic communities in Great Lakes  
 932 connecting channels. *Journal of Great Lakes Research*, 40, 1–6.

933 Sard, N. M., S. R. Smith, J. J. Homola, J. Kanefsky, G. Bravener, J. V. Adams, C. M. Holbrook,  
 934 P. J. Hrodey, K. Tallon, and K. T. Scribner. (2020). RAPTURE (RAD capture) panel

facilitates analyses characterizing sea lamprey reproductive ecology and movement dynamics. *Ecology and Evolution*, 10(3), 1469–1488.

Schindler, D. E., J. B. Armstrong, and T. E. Reed. (2015). The portfolio concept in ecology and evolution. *Frontiers in Ecology and the Environment*, 13(5), 257–263.

Seeb, L.W. and P.A. Crane. (1999). Allozymes and mitochondrial DNA discriminate Asian and North American populations of chum salmon in mixed-stock fisheries along the south coast of the Alaskan Peninsula. *Transactions of the American Fisheries Society*, 128, 88-103.

Sethi, S. A. (2010). Risk management for fisheries. *Fish and Fisheries*, 11(4), 341–365.

Shaklee, J. B., T. D. Beacham, L. Seeb, and B. A. White. (1999). Managing fisheries using genetic data: Case studies from four species of Pacific salmon. *Fisheries Research*, 43(1–3), 45–78.

Stephenson, R. L. (1999). Stock complexity in fisheries management: A perspective of emerging issues related to population sub-units. *Fisheries Research*, 43(1–3), 247–249.

Stepien, C. A., J. A. Banda, D. M. Murphy, and A. E. Haponski. (2012). Temporal and spatial genetic consistency of Walleye spawning groups. *Transactions of the American Fisheries Society*, 141(3), 660–672.

Stepien, C. A., and J. E. Faber. (1998). Population genetic structure, phylogeography and spawning philopatry in walleye (*Stizostedion vitreum*) from mitochondrial DNA control region sequences. *Molecular Ecology*, 7(12), 1757–1769.

Stepien, C. A., O. J. Sepulveda-Villet, and A. E. Haponski. (2015). Comparative genetic diversity, population structure, and adaptations of walleye and yellow perch across North America. Pages 643–689 in P. Kestemont, K. Dabrowski, and R. C. Summerfelt, editors. *Biology and Culture of Percid Fishes: Principles and Practices*.

958 Stepien, C. A., Snyder, M. R., & Knight, C. T. (2018). Genetic divergence of nearby walleye  
 959 spawning groups in central Lake Erie: Implications for management. *North American*  
 960 *Journal of Fisheries Management*, 38(4), 783–793.

961 Strange, R. M., and C. A. Stepien. (2007). Genetic divergence and connectivity among river and  
 962 reef spawning groups of walleye (*Sander vitreus vitreus*) in Lake Erie. *Canadian Journal of*  
 963 *Fisheries and Aquatic Sciences*, 64(3), 437–448.

964 Strayer, D. L., and D. Dudgeon. (2010). Freshwater biodiversity conservation: Recent progress  
 965 and future challenges. *Journal of the North American Benthological Society*, 29(1), 344–  
 966 358.

967 Tibihika, P. D., M. Curto, E. Alemayehu, H. Waidbacher, C. Masembe, P. Akoll, and H.  
 968 Meimberg. (2020). Molecular genetic diversity and differentiation of Nile tilapia  
 969 (*Oreochromis niloticus*, L. 1758) in East African natural and stocked populations. *BMC*  
 970 *Evolutionary Biology* 20(1), 16.

971 Toomey, L., S. Dellicour, T. Vanina, J. Pegg, Z. Kaczkowski, J. Kouřil, F. Teletchea, M. Bláha,  
 972 P. Fontaine, and T. Lecocq. (2020). Getting off on the right foot: integration of spatial  
 973 distribution of genetic variability for aquaculture development and regulations, the  
 974 European perch case. *Aquaculture*, 521, 734981.

975 Vähä, J. P., J. Erkinaro, E. Niemelä, C. R. Primmer, I. Saloniemi, M. Johansen, M. Svenning, and  
 976 S. Brørs. (2011). Temporally stable population-specific differences in run timing of one-sea-  
 977 winter Atlantic salmon returning to a large river system. *Evolutionary Applications*, 4(1),  
 978 39–53.

979 Vandergoot, C. S., and T. O. Brenden. (2014). Spatially varying population demographics and  
 980 fishery characteristics of Lake Erie walleyes inferred from a long-term tag recovery study.  
 981 *Transactions of the American Fisheries Society*, 143(1), 188–204.

982 Vinther, M., S. A. Reeves, and K. R. Patterson. (2004). From single-species advice to mixed-  
 983 species management: taking the next step. *ICES Journal of Marine Science*, 61(8), 1398–  
 984 1409.

985 Walleye Task Group, (WTG). (2018). Report of the Lake Erie Walleye Task Group to the  
 986 Standing Technical Committee. Great Lakes Fishery Commission, Ann Arbor MI

987 Waples, R. S. 2010. High-grading bias: subtle problems with assessing power of selected subsets  
 988 of loci for population assignment. *Molecular Ecology*, 19(13), 2599–2601.

989 Waples, R. S., and O. Gaggiotti. (2006). What is a population? An empirical evaluation of some  
 990 genetic methods for identifying the number of gene pools and their degree of connectivity.  
 991 *Molecular Ecology*, 15(6), 1419-1439.

992 Waples, R. S., K. A. Naish, and C. R. Primmer. (2020). Conservation and management of salmon  
 993 in the Age of Genomics. *Annual Review of Animal Biosciences*, 8(1), 117–143.

994 Wickham, H. (2009). ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York,  
 995 NY.

996 Wills, T., J. Markham, K. Kayle, M. Turner, A. M. Gorman, C. Vandergoot, M. Belore, A. Cook,  
 997 R. Drouin, T. MacDougall, Y. Zhao, C. Murray, and M. Hosack. (2018). Report for 2017 by  
 998 the Lake Erie Walleye Task Group. Great Lakes Fishery Commission, Ann Arbor MI

999 Xiao, N. (2018). ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for “ggplot2.”

1000 Yu, G. (2019). scatterpie: Scatter Pie Plot.

1001 Zhao, Y., D. W. Einhouse, and T. M. MacDougall. (2011). Resolving some of the complexity of  
 1002 a mixed-origin walleye population in the east basin of Lake Erie using a mark–recapture  
 1003 study. *North American Journal of Fisheries Management*, 31(2), 379–389.  
 1004  
 1005

## 1006 **Tables**

1007 Table 1: Walleye samples collected and genotyped for Rapture panel development and  
 1008 identification of reporting groups in mixed-stock analysis of recreational and commercial harvest  
 1009 in the East Basin of Lake Erie during 2016-2018. Columns show total sample size (N) and  
 1010 genetic diversity estimates including observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity, allelic  
 1011 richness (AR), and inbreeding coefficient ( $F_{IS}$ ) for final Rapture panel single nucleotide  
 1012 polymorphism (SNP) genotypes. These SNPs were identified during initial RAD-sequencing for  
 1013 panel development and microhaplotype genotypes identified through Rapture sequencing of 397  
 1014 baseline individuals used for reporting group identification (note the higher AR of  
 1015 microhaplotype loci than SNP loci in Baseline individuals). *Panel Development* samples  
 1016 generated the initial RAD-sequencing data that were subsequently used to develop the Rapture  
 1017 panel. This panel, in turn, was used to describe the *Baseline* population structure in Lake Erie.

Location								
Spawning Stock	Year(s)	of Stock	Reporting Group	N	$H_O$	$H_E$	AR	$F_{IS}$
<i>Panel Development</i>								
Maumee River	2012-2014	West	West/central basin	13	0.31	0.29	1.67	-0.08
Sandusky River	2012-2014	West	West/central basin	18	0.30	0.30	1.67	-0.03
Detroit River	2012-2014	West	West/central basin	13	0.29	0.31	1.66	-0.08
West Basin Mixed	2012-2014	West	West/central basin	26	0.29	0.29	1.67	-0.01
Ohio Grand River	2012-2014	Central	West/central basin	13	0.29	0.28	1.67	-0.03
Van Buren Bay	2012-2014	East	East basin	13	0.28	0.29	1.66	-0.00
<i>Baseline</i>								
Maumee River	2017, 2018	West	West/central basin	36	0.32	0.28	2.28	-0.09
Sandusky River	2017, 2018	West	West/central basin	36	0.31	0.29	2.33	-0.04
Detroit River	2017, 2018	West	West/central basin	36	0.30	0.29	2.32	-0.02

Ohio Reef Complex	2017	West	West/central basin	36	0.31	0.29	2.30	-0.06
Ohio Grand River	2012-2014	Central	West/central basin	13	0.30	0.29	2.38	-0.03
Van Buren Bay	2017, 2018	East	East basin	38	0.29	0.29	2.32	-0.01
Cattaraugus Creek	2017, 2018	East	East basin	18	0.29	0.28	2.32	-0.02
Lackawanna Shore	2017	East	East basin	24	0.30	0.28	2.29	-0.03
Bournes Beach	2018	East	East basin	14	0.33	0.28	2.25	-0.15
Shorehaven	2017	East	East basin	50	0.29	0.29	2.32	0.00
Ontario Grand								
Ontario Grand River	2014, 2016	East	River	96	0.29	0.29	2.29	0.03

---

1018

1019

1020

1021 Table 2: Number of walleye of unknown origin collected in the east basin of Lake Erie's

1022 commercial and recreational fisheries that were genotyped (N total), successfully assigned to any

1023 one of three reporting groups: east basin, west/central basin, and Ontario Grand River (N

1024 reassigned), or assigned to West Basin (N assigned to west/central basin). Assignments to

1025 reporting groups were grouped by year collected, sample source (either New York Department of

1026 Environment and Conservation [NY] or the Ontario Ministry of Natural Resources [ON]), month

1027 of harvest and fishery type. Individuals that could not be assigned to a natal source had > 50%

1028 missing genotypes and were removed from analysis. For easier interpretation, the number of

1029 walleye reassigned to west/central basin is also reported as a percent of total fish assigned.

Year	Data Source	Month	Harvest	N total	N assigned	N assigned to	% assigned to
						west/central basin	west/central basin
2016	NY	June	Recreational	7	5	1	20
2016	NY	July	Recreational	115	98	20	20
2016	NY	August	Recreational	7	6	6	100
<b>2016 Total:</b>				<b>129</b>	<b>109</b>	<b>27</b>	<b>25</b>
2017	NY	May	Recreational	28	26	1	4
2017	NY	June	Recreational	118	101	23	23
2017	NY	July	Recreational	138	120	65	54
2017	NY	August	Recreational	70	64	50	78
2017	NY	September	Recreational	126	114	91	80
2017	NY	October	Recreational	35	34	22	65
2017	ON	May	Commercial	9	9	1	11
2017	ON	June	Commercial	117	98	39	40
2017	ON	July	Commercial	124	106	41	39



2017	ON	August	Commercial	40	34	16	47
2017	ON	September	Commercial	125	105	74	70
2017	ON	October	Commercial	30	30	7	23
2017	ON	November	Commercial	30	28	26	93
2017	ON	December	Commercial	30	27	27	100
<b>2017 Total:</b>				<i>1020</i>	<i>896</i>	<i>483</i>	<i>54</i>
2018	ON	July	Commercial	110	62	56	90
2018	ON	August	Commercial	12	5	4	80
<b>2018 Total:</b>				<i>122</i>	<i>67</i>	<i>60</i>	<i>90</i>

---

1030

Table 3: Pairwise genetic differences of walleye collected at different Lake Erie spawning locations during 2014 - 2017 ( $F_{ST}$ ; Weir and Cockerham, 1984). Individuals from all spawning stocks were used in the baseline assessment of population structure of Lake Erie and used to define reporting groups for mixed-stock analysis.

Spawning Stock	Maumee River	Sandusky River	Detroit River	Ohio Reef Complex	Ohio Grand River	Van Buren Bay	Cattaraugus Creek	Lackawanna Shore	Bournes Beach	Shorehaven
Maumee River	0									
Sandusky River	0.0030	0								
Detroit River	0.0039	0.0018	0							
Ohio Reef Complex	0.0034	0.0022	0.0017	0						
Ohio Grand River	0.0060	0.0031	0.0019	0.0044	0					
Van Buren Bay	0.0069	0.0056	0.0054	0.0055	0.0054	0				
Cattaraugus Creek	0.0081	0.0066	0.0061	0.0070	0.0054	0.0012	0			
Lackawanna Shore	0.0069	0.0052	0.0055	0.0052	0.0058	4.00E-04	0.0015	0		

Bournes Beach	0.0118	0.0107	0.0098	0.0107	0.0126	0.0045	0.0068	0.0055	0	
Shorehaven	0.0070	0.0057	0.0056	0.0055	0.0058	6.00E-04	0.0011	5.00E-04	0.005	0
Ontario Grand										
River	0.0279	0.0260	0.0274	0.0277	0.0246	0.0312	0.0309	0.0315	0.0342	0.0306

---

Table 4: Statistics from comparisons of general linear models used to describe the proportion of west basin walleye (PropWB) harvested in the east basin of Lake Erie's recreational and commercial fisheries during 2017. The number of estimated parameters for each model (K), Akaike's information criterion for small sample sizes ( $AIC_c$ ), change in information criterion between sequential models (Delta  $AIC_c$ ),  $AIC_c$  weight ( $AIC_c Wt$ ), and evidence ratio (ER) are provided. The most parsimonious model is in bold-face text

Model	K	$AIC_c$	Delta $AIC_c$	$AIC_c Wt$	ER
<b>PropWB~lon+lat+month</b>	<b>5</b>	<b>-5.54</b>	<b>-</b>	<b>0.869</b>	<b>-</b>
PropWB~lon+lat	4	-0.99	4.52	0.091	10.14
PropWB~lon+month	4	1.28	6.79	0.029	31.52
PropWB~lat+month	4	3.64	9.15	0.009	102.57
PropWB~lon	3	7.61	13.12	0.001	788.41
PropWB~month	3	8.33	13.84	0.001	1130.40
PropWB~lat	3	19.11	24.62	3.92E-06	247246.43

## Figures

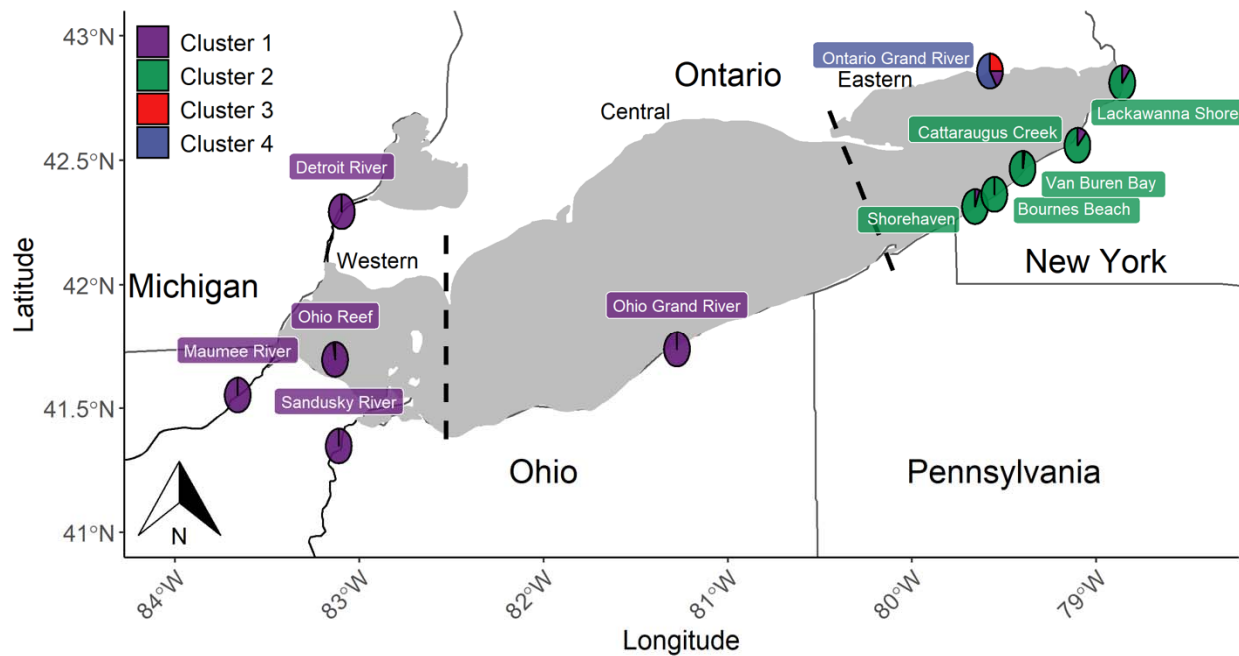


Fig. 1: Baseline sample locations and discriminant analysis of principal components (DAPC) results for walleye clustered into the four most parsimonious groups based on Bayesian Information Criterion. Pie charts display the posterior probability of each DAPC cluster proportion averaged across all individuals at the site. The background color of the site label indicates the final reporting group into which the stock was grouped for assignment tests. Lake basins are delineated by dotted lines and labeled on the northern shore of the lake.

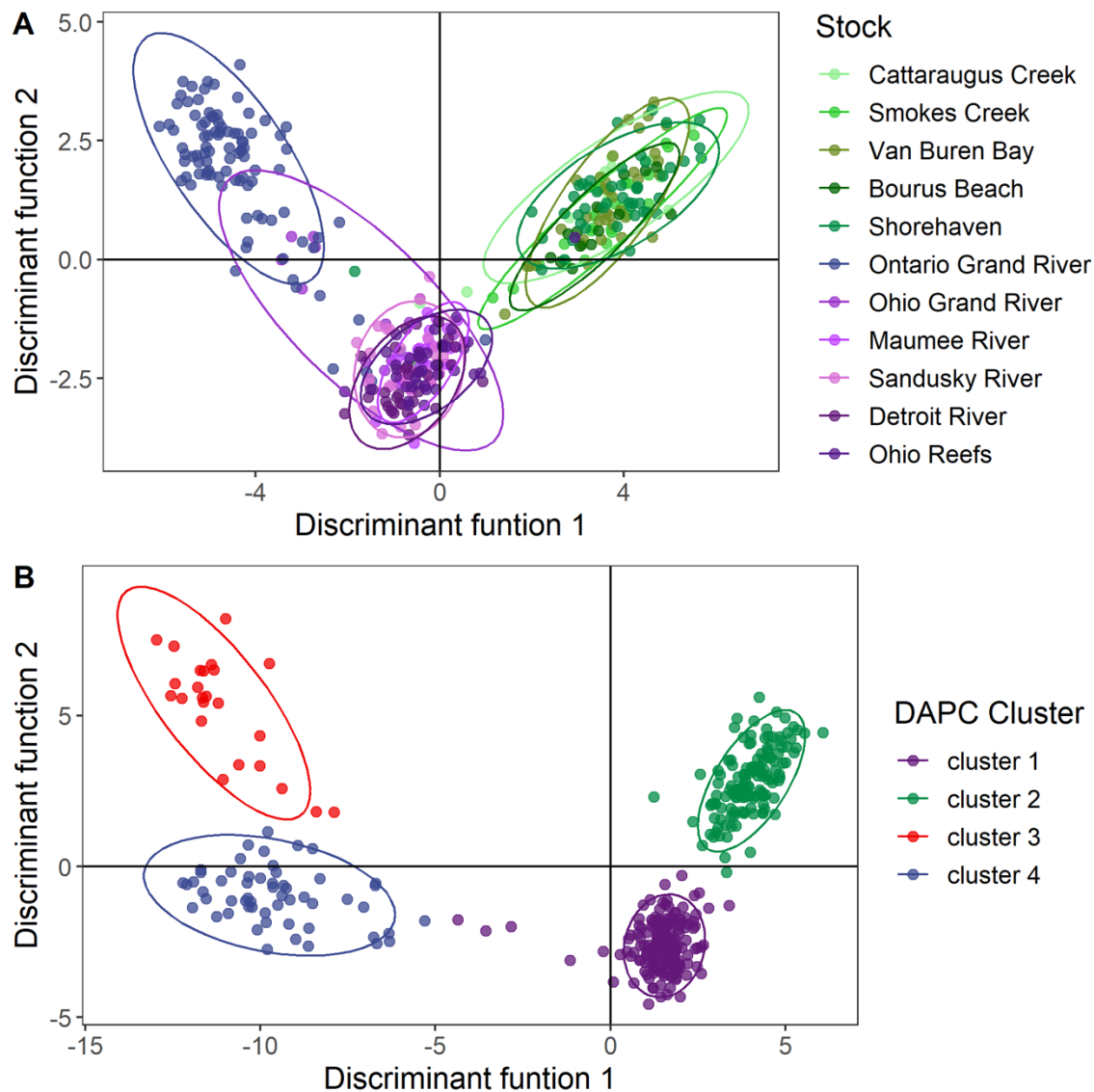


Fig. 2: Discriminant analysis of principal components (DAPC) of walleye (individual points) collected on spawning grounds during the spawning season 2012 – 2018. Individuals are grouped by spawning stocks (A) and into the 4 most parsimonious clusters based on Bayesian Information Criterion (B). Ellipses show the 95% confidence interval around each group. Colors in A correspond approximately to the reporting groups shown in Figure 1: west/central basin = purples; east basin = greens; Ontario Grand River = blue.

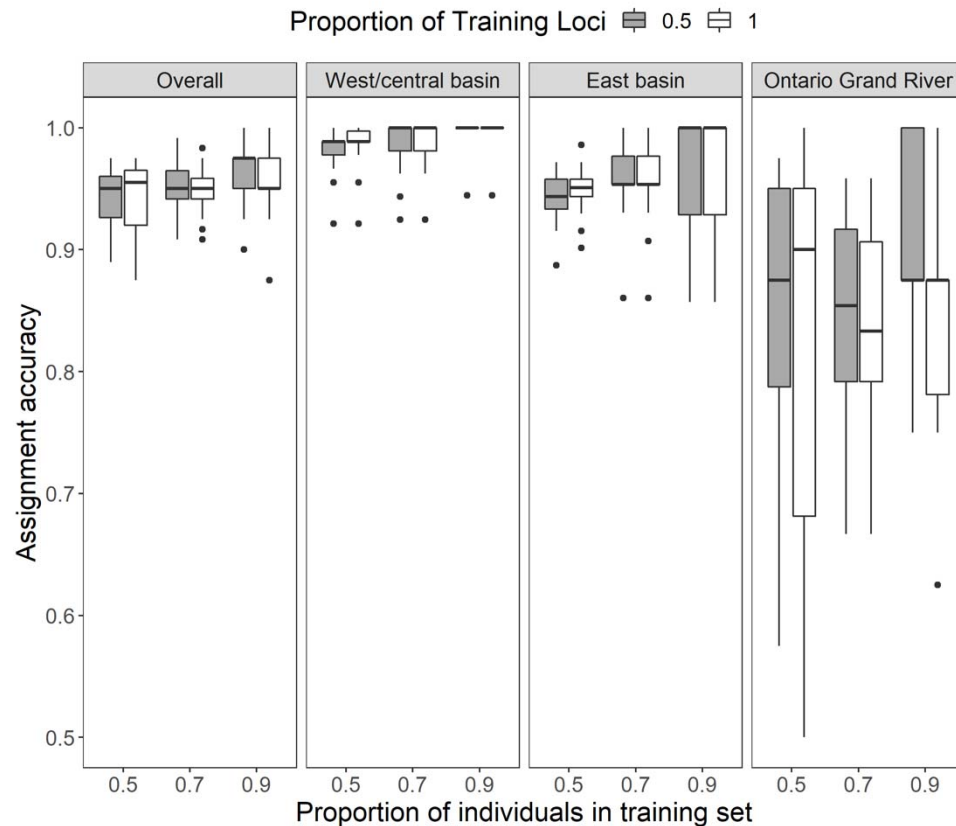


Fig. 3: Reassignment accuracies of three reporting groups identified based on 395 adult Walleye collected from 11 Lake Erie spawning sites during the spawning season between 2012–2017. The reporting groups are (1) West/Central Basin: Maumee River, Sandusky River, Detroit River, the Ohio reef complex, Ohio Grand River; (2) East Basin: Shorehaven, Bournes Beach, Van Buren Bay, Cattaraugus Creek, Lackawanna Shore; and (3) Ontario Grand River. Reassignment accuracy was determined using either 0.5 or 1 proportion of training loci (gray and white bars, respectively) and a support-vector machine algorithm (Chen et al. 2020), with training samples for each grouping consisting of 0.5, 0.7, or 0.9 proportion of the collected individuals (chosen randomly). The remainder of individuals (0.5, 0.3, or 0.1) was used as the test (holdout) data set to determine reassignment accuracy. Box plots portray medians (thick black line), interquartile ranges (ends of boxes), and outliers (black dots).

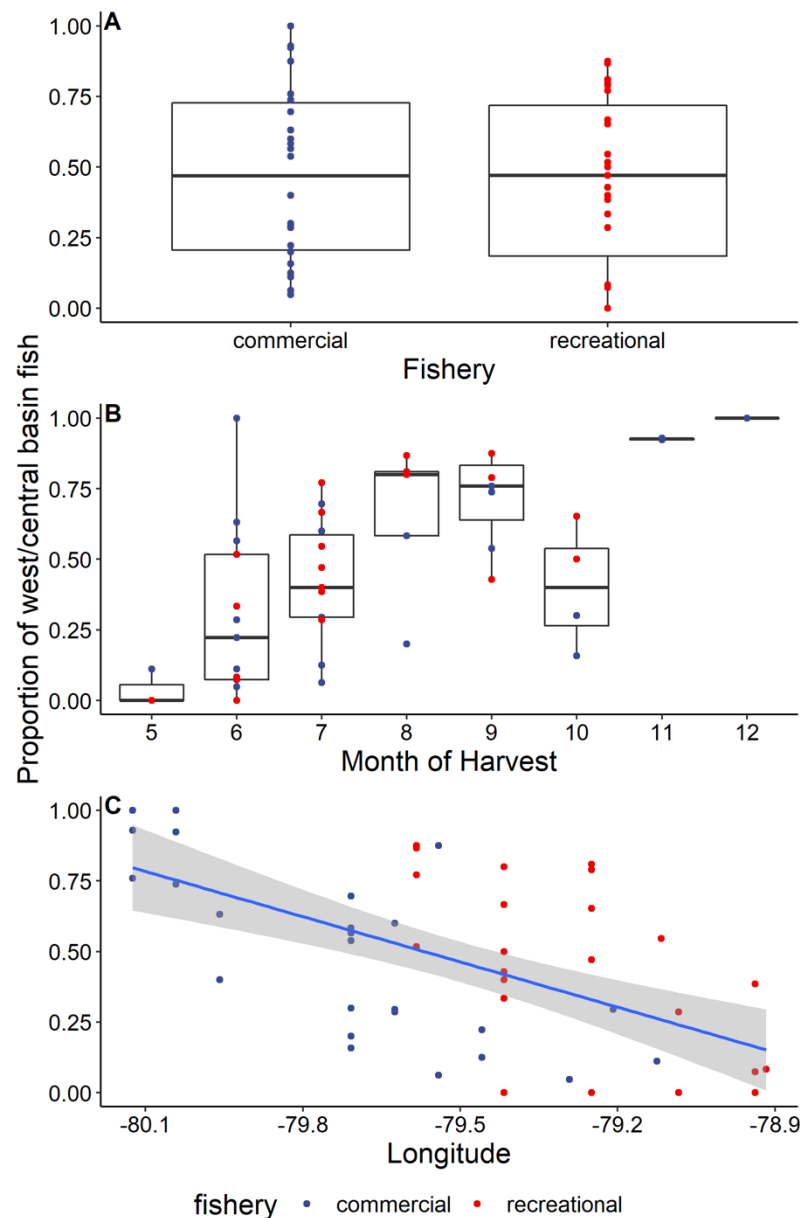


Fig. 4: Proportion of walleye harvested in eastern Lake Erie's recreational and commercial fisheries during 2017 assigned to the west/central basin reporting group (A), broken down by harvest month (B) and location (longitude; C). Each colored point represents the proportion of fish assigned to the west/central basin from a particular sampling event (grid/date combination). Blue line and gray background in C represent correlation of generalized linear model and 95% confidence interval. Only sampling events that contained greater than 6 assigned individuals were included. Note higher sample sizes for core months of harvest, June – September.



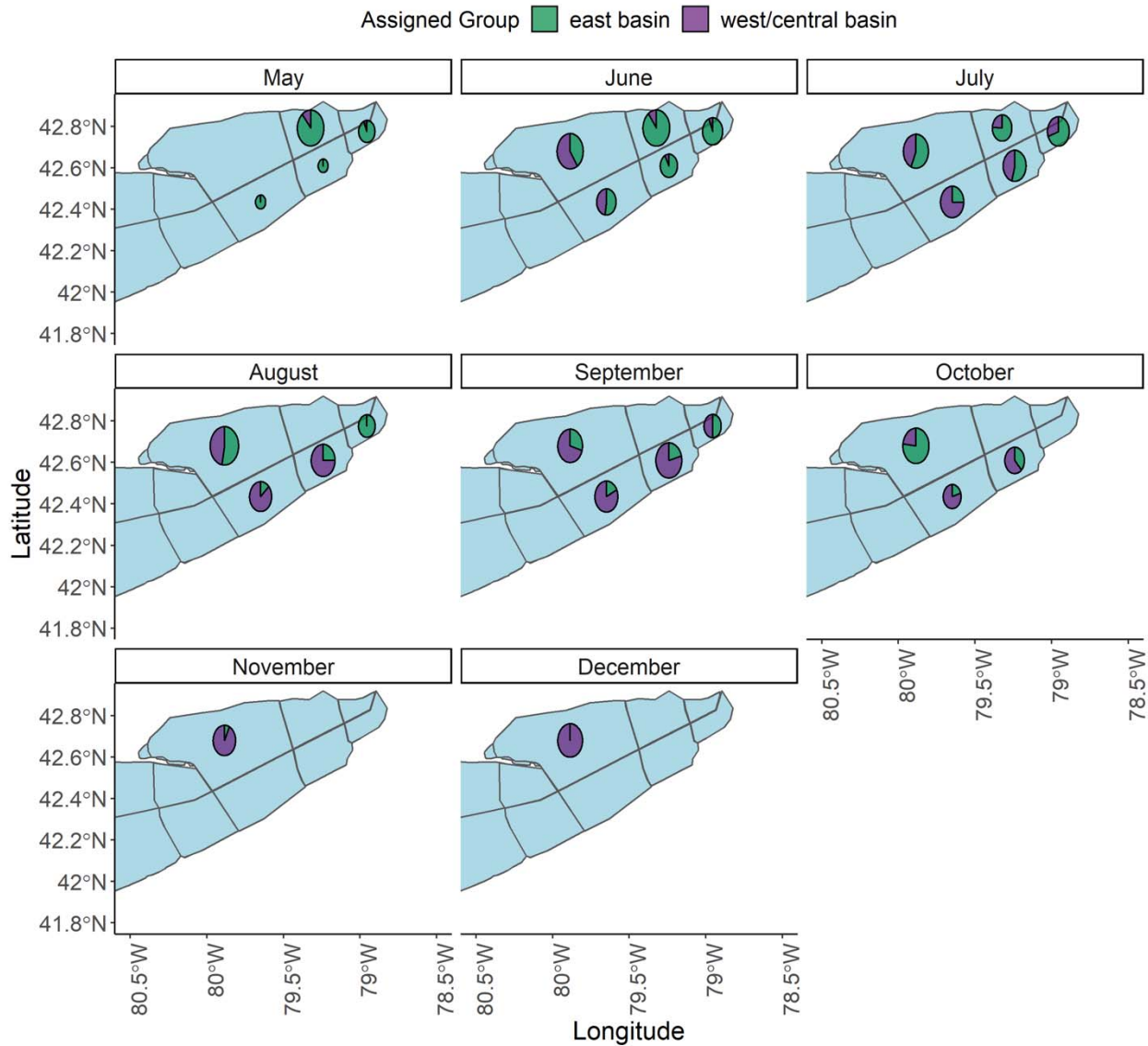


Fig. 5: Monthly assignment of walleye to basin of origin sampled from the commercial and recreational harvest in Lake Erie's east basin in 2017 from eight harvest zones. The size of each pie corresponds to  $\log_{10}$  normalized mean number of fish harvested from creel sampling locations with centroids located within each harvest zone. Only two colors are shown because no harvested individuals were assigned to the Ontario Grand River reporting group.

## Supporting Information

Suppl. Table 1: Locus specific diversity estimates (Locus ID, number of alleles, effective number of alleles, observed heterozygosity, inbreeding coefficient ( $G_{IS}$ ), and genetic distance estimate ( $G_{ST}$ ) for the 8,482 microhaplotypes used for individual assignments.

Suppl. Table 2: Metadata collected for all mixed-stock individuals including assigned basin.

Suppl. Fig. 1: Reassignment accuracies of 11 spawning stocks identified based on 395 adult Walleye collected from 11 Lake Erie spawning sites between 2012–2017 calculated in assignPOP: West/Central Basin – Maumee River (MA), Sandusky River (SA), Detroit River (DE), the Ohio reef complex (RE), Ohio Grand River (OHG), Ontario Grand River– Ontario Grand River (ONG), East Basin – Shorehaven (SH), Bourus Beach (BB), Van Buren Bay (VB), Cattaraugus Creek (CC), Smokes Creek (SC). Reassignment accuracy was determined using either 0.5 or 1 proportion of training loci (colors) and a support-vector machine algorithm, with training samples for each grouping consisting of 0.5, 0.7, or 0.9 proportion of the collected individuals (chosen randomly). The remainder of individuals (0.5, 0.3, or 0.1) was used as the test (holdout) data set to determine reassignment accuracy. Box plots portray medians (thick black line), interquartile ranges (ends of boxes), and outliers (black dots).

Suppl. Fig. 2: Summary of rubias assignment accuracy of three Lake Erie reporting groups Ontario Grand River (ONG), west/central basin (West\_Basin) and east basin (East\_Basin). Mean posterior mixing proportions of 100% mixtures for each reporting group across the first 25 iterations (top) and distribution of individual posterior reporting group proportion for simulated collections from 100% mixtures of each reporting group.

Suppl. Fig. 3: Total walleye harvest by commercial and recreational fishing in 2017 across the 98 commercial harvest reporting grids and 13 recreational reporting grids located in the study area as reported in the annual Lake Erie Walleye Task Group report for 2017.