Fisheries Research 230 (2020) 105646

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

Detecting fine-scale population structure in the age of genomics: a case study of lake sturgeon in the Great Lakes

J.M. Whitaker[a,*], L.E. Price[b], J.C. Boase[c], L. Bernatchez[d], A.B. Welsh[b]

[a] Department of Biological Sciences, Nicholls State University, 906 East 1st Street, P.O. Box 2021, Thibodaux, LA, 70301, United States
[b] School of Natural Resources, West Virginia University, PO Box 6125, Morgantown, WV, 26506-6125, United States
[c] U.S. Fish and Wildlife Service - Alpena Fish and Wildlife Conservation Office, 480W. Fletcher Street, Alpena, MI, 49707, United States
[d] Institut de Biologie Intégrative des Systèmes (IBIS), Université Laval, Québec, QC, G1Y 2T8, Canada

ARTICLE INFO

Handled by J Viñas

Keywords:
single nucleotide polymorphisms
microsatellites
lake sturgeon
fine-scale structure

ABSTRACT

Great Lakes-wide population structure analyses using neutral markers have provided an understanding of broad-scale genetic structure of lake sturgeon. To assess the fine-scale genetic structure of lake sturgeon populations in two different rivers, we combined both microsatellites and genome-wide SNP markers. The St. Clair-Detroit River System (SCDRS) is entirely freshwater with no known impediments to sturgeon movement. In contrast, the St. Lawrence River (SLR) outlets into the Atlantic Ocean and is fragmented by hydroelectric dams. Both microsatellites and SNPs provided evidence of differentiation between the rivers. When applied to fine-scale structure, microsatellites failed to detect population structure using a Bayesian approach for within either river and F_{ST} values using microsatellites identified only a low level of differentiation between the upper and lower St. Clair River. Using the full set of SNPs for each comparison yielded similar results to the microsatellite results. Discriminant analysis of principal components using both markers partitioned the samples into spatially structured clusters. The SNP datasets filtered for high F_{ST} values had greater success for detecting fine-scale population structure and had the highest accuracy for reassignment to prior populations. This reduced SNP dataset may represent a more meaningful set of loci that can be used to estimate lake sturgeon fine-scale population structure, which is complicated by their long-generation times.

1. Introduction

Microsatellites have been the most widely used genetic marker for population structure (Morin et al., 2004; Narum et al., 2008; Väli et al., 2010) and when compared to single nucleotide polymorphisms (SNPs), effectiveness of the markers to delineate structure varies. SNPs increase the power of statistical analyses due to the high number of loci identified; however, in some cases, microsatellites have proven more effective at differentiating fine-scale structure, most likely due to higher allelic diversity (Hess et al., 2011; Narum et al., 2008; Vignal et al., 2002). Alternatively, SNPs are often valuable for differentiating between populations on a broader scale (e.g. Beacham et al., 2020; Davey et al., 2011). More importantly, SNPs provide the further advantage of identifying adaptive divergence, which has been found in Nujiang catfish *Creteuchiloglanis macropterus* (Kang et al., 2017), the Pacific leaping blenny *Alticus amoldorum* (Morgans et al., 2014), striped marlin *Kajikia audax* (Mamoozadeh et al., 2019), and rainbow trout *Oncorhynchus mykiss* (Van Doornik et al., 2013).

In addition to connectivity and differentiation, including adaptive markers into genetic monitoring may enable the management of the evolutionary potential of populations and alter the way that management units are delineated (Allendorf et al., 2010; Bernatchez et al., 2017; Funk et al., 2012; McMahon et al., 2014). Adaptive markers provide insight into genetic variation that affects fitness and may serve as a proxy for local adaptation (Allendorf et al., 2010; Bernatchez, 2016; Funk et al., 2012). Stocking based on the evolutionary potential of organisms could provide the diversity necessary to be resilient in changing or stochastic environments.

Lake sturgeon (*Acipenser fulvescens*) are a good case study for a fine-scale analysis of population structure using both microsatellites and SNPs because populations across their distribution have declined drastically and despite stocking efforts, many populations have yet to return to historic numbers (Auer, 1996). Great Lakes-wide population structure analyses using neutral markers showed that lake sturgeon are genetically diverse throughout their range and most spawning populations are genetically distinct (DeHaan et al., 2006; Welsh et al., 2008).

* Corresponding author.
E-mail address: justine.whitaker@nicholls.edu (J.M. Whitaker).

https://doi.org/10.1016/j.fishres.2020.105646
Received 25 November 2019; Received in revised form 13 May 2020; Accepted 21 May 2020
Available online 10 June 2020

Analysis of 22 Great Lakes spawning populations, using eight microsatellite loci, resulted in the designation of six genetic stocking units, some of which encompass multiple rivers (Welsh et al., 2010). However, analyses on a finer geographic scale have generated mixed results. Within-river population structure has been identified in lake sturgeon in the Winnipeg and Ottawa Rivers (McDougall et al., 2017; Wozney et al., 2011), but fine-scale analysis of lake sturgeon in the Namakan River revealed the presence of only a single population (Welsh and McLeod, 2010). If cryptic population structure exists within a river, but is not being managed properly, a fine-scale analysis of genetic structure may provide a better understanding of why populations are not recovering to historic numbers.

Two rivers, the St. Clair-Detroit River system (SCDRS) and the St. Lawrence River (SLR), were chosen for this study because lake sturgeon are currently managed as single populations in both rivers, but have the potential for unrecognized genetic structure. The SCDRS consists of the Detroit River and the St. Clair River connecting Lake Huron to Lake Erie. Welsh et al. (2008) determined that the SCDRS, which has no dams, appears to contain one panmictic population based on microsatellite data. However, the aim of the Welsh et al. (2008) study was not to identify fine-scale structure, which was reflected in the sampling scheme. In contrast to genetic data, movement data of lake sturgeon in the St. Clair-Detroit River System suggest that there may be multiple populations (Kessel et al., 2018; Thomas and Haas, 2002). Lake sturgeon exhibit site fidelity (Auer, 1999: Barth et al., 2011; Gunderman and Elliott, 2004), which can result in high degrees of population structure. It is reasonable to hypothesize that there are multiple populations within the SCDRS that are associated with the five known spawning locations in the St. Clair and Detroit Rivers. Although genetic analyses have been performed on the SCDRS (Welsh et al., 2008), a refined analysis using SNPs may reveal fine-scale structure and be beneficial to the long-term management of lake sturgeon.

The SCDRS is managed as a single population due to the lack of impediments. Fishing for lake sturgeon in the SCDRS is currently limited to recreational fishing in Michigan waters only with a restricted season and limited recreational harvest. Due to the size and stability of the SCDRS lake sturgeon population, stocking is not a priority (Hayes and Caroffino, 2012).

In contrast to the SCDRS, the SLR has been modified by dam construction, including the Beauharnois–Les Cèdres (Beauharnois, QC, Canada) and Moses-Saunders Power Dams (Cornwall, ON, Canada). The Moses-Saunders dam was constructed in 1958 and Beauharnois-Les Cèdres construction took place from 1912-1961. Lake sturgeon may be trapped between the Moses-Saunders dam and the Beauharnois-Les Cèdres, causing fragmentation from fish downstream of the Beauharnois-Les Cèdres (Mailhot et al., 2011). Additionally, the SLR extends from Lake Ontario to the Gulf of St. Lawrence, eventually connecting with the Atlantic Ocean, with an established fluvial to salt and a no tide to pronounced tide gradients (Caron et al., 2002) that may contribute to population structure. The fluvial estuary extends from Lake Ontario to Orleans Island, where a brackish estuary begins (Hatin et al., 2002).

Despite fragmentation, the Ontario Ministry of Natural Resources and Forestry includes the entire Canadian SLR as one designatable unit (Unit 8) but divides that unit into three management units (Pratt, 2008). The New York State Department of Environmental Conservation manages the SLR as two management units: the upper SLR (from Cape Vincent to the Moses-Saunders Dam) and lower SLR (Lake St. Francis, Akwesasne, Grasse, Raquette, St. Regis, Salmon rivers), but has identified a research goal to better understand population structure in the lower unit (Holst, 2018). According to the 2018 Lake Sturgeon Recovery Plan (Holst, 2018), stocking has occurred most recently in 2013-2015 in the middle SLR (15,750 individuals over three years) and Lake St. Francis (10,500 individuals over three years) and recommendations suggest using genetically similar fish (Welsh et al., 2010). A closely monitored commercial fishery exists on parts of the SLR (Mailhot et al., 2011), with commercial fishing taking place in the lower part of the river system, mainly in Lake St-Louis, Lake St-Pierre, and the upper estuary (Mailhot et al., 2011).

Determining the level of genetic structure in these rivers could alter management strategies, including delineation of management units, stocking programs and fishing regulations. Smaller populations are at a greater risk of extinction (Gilpin and Soulé, 1986) due to the potential for inbreeding, and the deterioration of genetic diversity and evolutionary potential (Allendorf and Leary, 1986; Amos and Balmford, 2001). Stocking, which is valuable for introducing and maintaining genetic diversity, relies upon accurate assessment of population fragmentation and connectivity. If multiple populations exist within a river, maintaining that diversity may be important for developing sustainable stocking and fishing strategies. The main goals of this study are to identify and compare fine-scale population structure of lake sturgeon populations using both microsatellites and SNPs within two Great Lakes river systems that vary in terms of their potential for genetic structure.

2. Methods

2.1. Study sites

The SCDRS consists of two major lakes and a corridor connecting them, with no damns or impediments to movement (Fig. 1). The corridor is made up of the St. Clair River, Lake St. Clair and the Detroit River. The St. Clair River, which is 64 kilometers long, connects Lake Huron and Lake St. Clair (Boase et al., 2011). Depths in the river range from 8-24 meters, and velocities vary (Edsall et al., 1988), with the highest velocity occurring around the International Blue Water Bridge. The Detroit River, which is 51 kilometers long, connects Lake St. Clair with Lake Erie (Caswell et al., 2004). The SCDRS may have contained as many as 15 spawning sites at one time (Goodyear et al., 1982), but spawning habitat has been removed or degraded due to dredging for the shipping channel (Manny et al., 2007; Read and Manny, 2006). The SCDRS now consists of a few historic spawning sites and recently created artificial reefs built to support increased spawning of native fish (Read and Manny, 2006; Roseman et al., 2011).

The SLR (Fig. 1) connects Lake Ontario to the Atlantic Ocean, outputting to the Gulf of St. Lawrence (Mailhot et al., 2011). Lake St. Francis, Lake St. Louis, and Lake St. Pierre are contained within the SLR system. This river flows over 1,500 km, is comprised of two saline estuaries and a freshwater estuary, and flows at an average of 12,600 m^3/sec at Quebec City (Mailhot et al., 2011). The SLR main channel was dredged for shipping access and is typically between 10-15 m deep (Basu et al., 2000; Hudon et al., 1996). Other modifications to the SLR include dam construction, channeling, and flow regulation (Mailhot et al., 2011). The Beauharnois–Les Cèdres (Beauharnois, QC, Canada) and Moses-Saunders Power Dams (Cornwall, ON, Canada) occur within this study area.

2.2. Sampling

In the SCDRS, lake sturgeon were captured in the upper St. Clair River (n = 56), the lower St. Clair River (n = 44), and the Detroit River (n = 22) during the spawning season (May-June) from 2011 to 2015 as part of population surveys conducted by the Michigan Department of Natural Resources in the lower St. Clair, and by the U.S. Fish and Wildlife Service in the upper St. Clair and Detroit rivers. Sturgeon collections were supplemented with fish captured by commercial fishers in southern Lake Huron, approximately 15 km NNE of the Blue Water Bridge. Most lake sturgeon were captured using baited set lines deployed from a research vessel in the afternoon and retrieved the following morning (Thomas and Haas, 2002). Sturgeon supplied by commercial fishers were collected as by-catch in trapnets.

In the SLR, samples (n = 30) were collected from the Quebec City lake sturgeon area (mouths of the Chaudière River and Montmorency

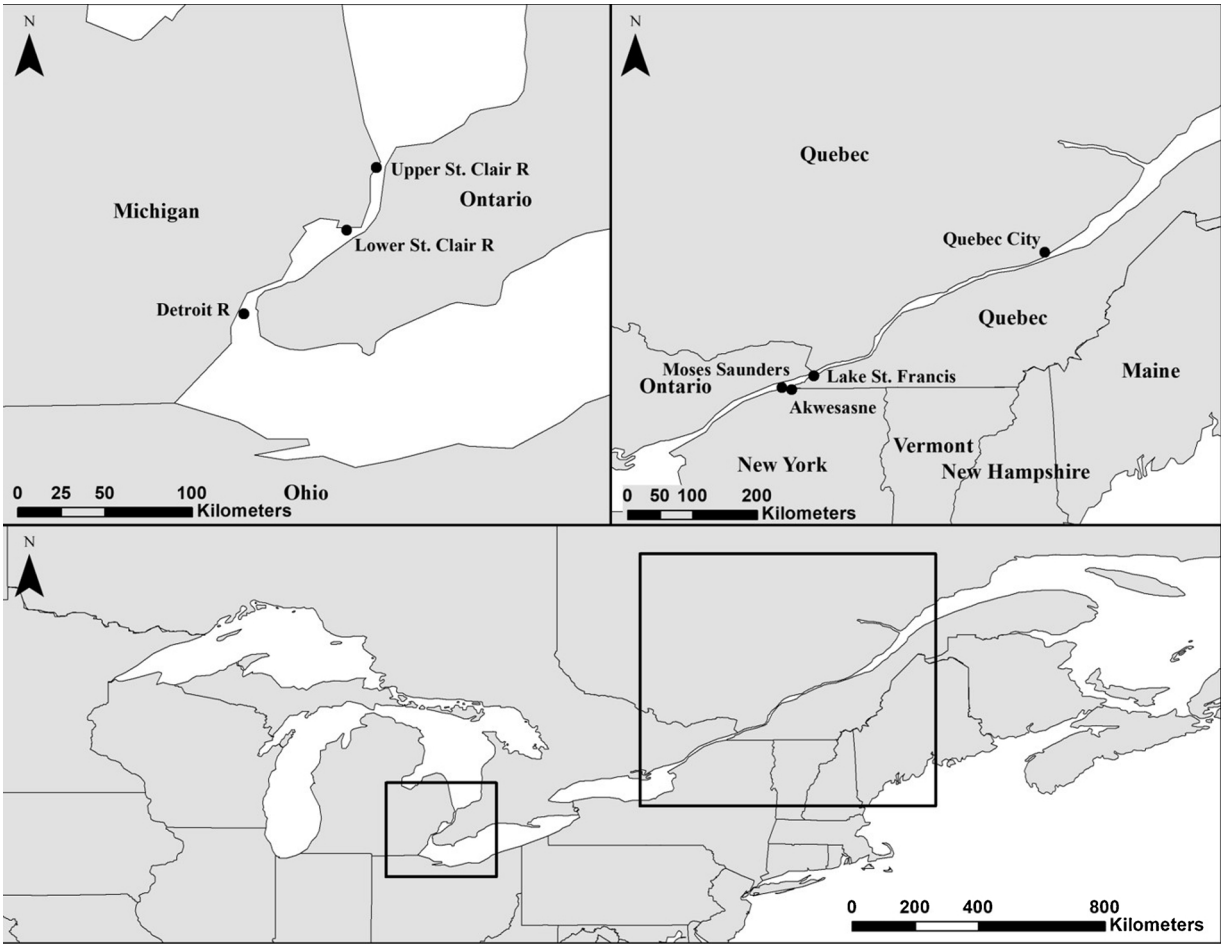

Fig. 1. This map provides the location of both study sites: the St. Clair Detroit River System (SCDRS) and the St. Lawrence River (SLR), including enlarged regions detailing the lake sturgeon sampling locations (black dots). The two black bars represent the hydroelectric dams located on the SLR.

Table 1
Lake sturgeon sampling locations and year for samples used for genotyping. The samples identified as Quebec City were taken from the Chaudiere River and Montmorency River in the St. Lawrence estuary.

		Sample Size	
Location	Year	Microsatellites	SNPs
St. Lawrence River			
Moses Saunders Dam	1994	53	24
Lake St. Francis	2006	28	-
Akwesasne	2012	22	-
Quebec City	2012	30	26
St. Clair-Detroit River System			
Upper St. Clair River	2012	57	5
Lower St. Clair River	2012	43	19
Detroit River	2012	22	23

River) by the Ministry of Natural Resources and Wildlife in Quebec. Lake sturgeon were captured using stretch mesh gill nets of 20.5, 23, and 25.5 cm. Samples were also collected from Lake St. Francis (N = 28) and an Akwesasne sampling site (N = 22) (Table 1). For microsatellite genotyping, these data were combined with previously analyzed samples from below Moses-Saunders Dam (N = 54) (Welsh et al., 2008).

2.3. Microsatellite Genotyping

DNA was extracted from fin clips collected from the SCDRS (n = 122) and SLR (n = 133) using the Gentra Puregene Tissue Kit (©Qiagen, Germantown, MD, USA) according to manufacturer's protocol and extracts were quantified on the Nanodrop spectrophotometer (©Thermoscientific, Wilmington, DE, USA). Extracted DNA was analyzed at 12 disomic microsatellite loci: *Afu*G9, *Afu*G56, *Afu*G63, *Afu*G74, AfuG112, *Afu*G160, *Afu*G195, *Afu*G204, *Afu*68, *Afu*68b, *Spl*120, Aox27 (Welsh and May, 2006). PCR amplification was done using the Qiagen Multiplex PCR Kit® (Germantown, MD, USA). Four primer mixes were prepared. Primer mix 1 (*Afu*G 56, *Spl*120 and *Afu*G 195), primer mix 2 (*Afu*G204, *Afu*G74, and *Afu*G9), and primer mix 3 (*Afu*G160, *Aox*27 and *Afu*G112) contained 2μM final concentration of each primer. Primer mix 4 included Afu68 and AfuG68b primers (final concentration of 1μM) and AfuG63 primers (final concentration of 2 μM). The Mulitplex PCR Master Mix (final concentration of 1X) was used in a total reaction volume of 10 μl, with 20 ng of extracted DNA. The PCR conditions were performed as in Whitaker et al. (2018). PCR products were then visualized on a Beckman Coulter GeXP Genetic Analysis System.

2.4. Single Nucleotide Polymorphism Genotyping

Subsets of each population were used for SNP genotyping. In the SLR, samples from the population below the Moses-Saunders dam (n = 24) and the estuary near Quebec City (n = 26) were used for genotyping-by-sequencing based on the protocol of Poland et al. (2012). In the SCDRS, samples from the Detroit River (n = 23) and St. Clair River (n = 19) were used. Extracted DNA was digested with two restriction enzymes (*PstI* and *MspI*); adaptors and individual barcodes

were then ligated onto the cut ends. Samples were pooled, fragments between 250 and 450 bp were selected using a Pippin Prep (Sage Science), and PCR amplification was conducted on the resulting fragments. Quality and quantity were assessed using an Agilent 2100 BioAnalyzer and the library was sequenced on an Illumina HiSeq 2500 generating single-end 100bp reads. If < 1 million reads/individual were obtained, additional runs were conducted after normalization to correct for underrepresented individuals.

SNP identification was conducted using the pipeline STACKS v2.1 (Catchen et al., 2013). Sequences were cleaned, low-quality reads were discarded, and individuals were separated by barcodes using *process_radtags*. Sequences were aligned and stacks were created using a minimum depth (-m) of 3 and an allowable distance between stacks of 5 (-M) in *ustacks*. A catalog was created, allowing a distance of 5 between catalog loci (-n) in *cstacks* and individual stacks were aligned to the catalog using *sstacks*. SNPs were initially identified using *gstacks* and population comparisons were done separately for the two systems and between the two systems using *populations*. To achieve a balance between amount of missing data and the number of identified SNPs, various percentages of individuals required to have the loci in each population (-r) were tested (70%, 75%, 80%, 85%, 90%). The percentage with the greatest gain in number of SNPs and minimal increase in missing data was 75%. This level was used for the remainder of the analyses, along with the requirement that the loci had to be present in all populations. Individuals with 70% missing data or more were discarded. Lake sturgeon are polyploid, making paralogs a serious concern. Paralogs were identified based on the method of McKinney et al. (2017), as implemented in *paralog-finder* (Ortiz, 2018). Paralogs were blacklisted and removed from a second *populations* run.

To increase the power of detecting fine-scale population structure and improving assignment accuracy, a second dataset was created containing SNPs with significant F_{ST} values ($p < 0.05$) between the compared populations. This resulted in 1,436 SNPs with F_{ST} values ranging between 0.03 – 0.19 (average: 0.06) for the STL-SCDRS comparison; 603 SNPs ranging between 0.09 – 0.35 (average: 0.13) for the SCDRS, and 1,415 SNPs ranging between 0.04 – 0.21 (average: 0.13) for the SLR.

2.5. Population Clustering and Genetic Differentiation

2.5.1. Microsatellites

Hardy-Weinberg equilibrium (HWE) was tested for each locus and population using the software GENEPOP 4.2 (Raymond and Rousset, 1995), which utilizes an exact probability test estimated by Markov chain algorithms. The following parameters were used: 10,000 burn-in, 100 batches, 1,000 iterations, and $\alpha < 0.05$. Significance of the HWE tests was determined after performing a sequential Bonferroni correction (Rice, 1989). Linkage disequilibrium was tested for pairs of loci using probability tests, via GENEPOP 4.2 (Raymond and Rousset, 1995). A burn-in length of 10,000 was used for 100 batches and 1,000 iterations.

A Bayesian approach was implemented to determine the most likely number of populations in each river, using the software STRUCTURE 2.3.4 (Pritchard et al., 2000), which estimates the likelihood of population clusters based on Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium. A burn-in length of 1,000,000 and 1,000,000 MCMC steps were used for 10 iterations under the admixture model for K = 1 to K = 5. It was assumed that the allele frequencies were correlated. Since samples were taken during spawning season, the sampling location was used as prior knowledge.

A Discriminant Analysis of Principal Components (DAPC) was performed using the *adegenet* package in R (cran mirror: USA PA1) to identify variation in genotypes between population locations (Jombart, 2008). The DAPC uses Principal Component Analysis prior to the Discriminant Analysis (Jombart et al., 2010). For the SCDRS, there were three locations (upper St. Clair, lower St. Clair, and Detroit rivers) used in the analysis; therefore, two DAPC axes were retained. Twenty-five PCs were retained, as 25 had the lowest root mean squared error (10%) and had a decreased likelihood of overfitting the data. For the SLR, four DAPC axes were retained, as there were five sampling locations, and 35 PCs were retained. Cross-validation was used to assess the validity of discriminant function and principal component retention. An estimate of the number of clusters was determined by choosing the lowest *BIC* score.

Microsatellite genotypes were used to determine genetic differentiation between the sampling locations for each system. The F_{ST} value between the sampling sites was calculated using the program FSTAT 2.9.3.2 (Goudet, 1995), which implements the method outlined in Weir and Cockerham (1984), to determine if sampling sites within the rivers were genetically distinct. F_{ST} values range from 0 to 1 and larger numbers correspond to increased genetic differentiation.

2.5.2. Single Nucleotide Polymorphisms

Hardy-Weinberg equilibrium and linkage disequilibrium were tested for and filtered out using the dartR Package in R (Gruber et al., 2018). Population genetic structure was estimated from the final full set of SNPs for each river (SCDRS, n = 51,747; SLR, n = 30,034). DAPC was performed as described above using the *adegenet* package in R (cran mirror: USA PA1, Jombart, 2008). FastStructure (Raj et al., 2014) was used to evaluate the potential number of populations. The default parameters were used in fastStructure, with a simple prior and testing K = 1 through 5, to determine the model complexity that maximizes the marginal likelihood. To determine genetic differentiation between the sampling locations for each system, F_{ST} values were calculated using the function *gl.fst.pop* in the program *dartR*, with significance assessed through bootstrapping.

Two methods were used to identify sites that may be under selection. Using BayeScan 2.1 (Foll and Gaggiotti, 2008), sites that may be under selection were identified by analyzing patterns of differentiation in allelic frequencies that diverged from neutral expectations. Prior odds for the neutral model were set to 10:1. The program was run with 20 pilot runs at a length of 5,000 for 5,000 iterations. The burn-in (after the pilot runs) was 50,000 iterations, and the thinning interval was set to 10. F_{IS} values ranged from 0 to 1 and $\alpha < 0.05$ was considered to be significant. If the alpha value was positive, the selection was considered to be diversifying, whereas if the alpha was negative, then the selection was considered to be balancing or purifying. Model choice was performed using Bayes Factors, which were calculated for the individual restriction sites. Default parameters and a 5% trim fraction were used to run OutFLANK, through the R package, dartR (Gruber et al., 2018). Prior to outlier detection, SNP datasets were filtered to remove loci with minor allele frequency less than 0.1 (Whitlock and Lotterhos, 2015). Outliers identified through these analyses and putatively neutral loci were analyzed separately for population structure using the methods described above. Private alleles were identified using default parameters of the *private_alleles* function in the R package *radiator* (Gosselin et al., 2020).

2.5.3. Population Assignment

For both microsatellites and SNPs, population assignment was performed using the *assignPOP* package in R (Chen et al., 2018). Monte Carlo cross-validation was performed using the assign.MC and accuracy.MC functions to subset samples (proportions: 0.5, 0.7, 0.9) and assign them to populations, using 0.1, 0.25, 0.5, 1% of the loci, performed for 100 iterations.

3. Results

3.1. Microsatellites

3.1.1. St. Clair Detroit River System

No loci showed evidence of linkage disequilibrium in the samples

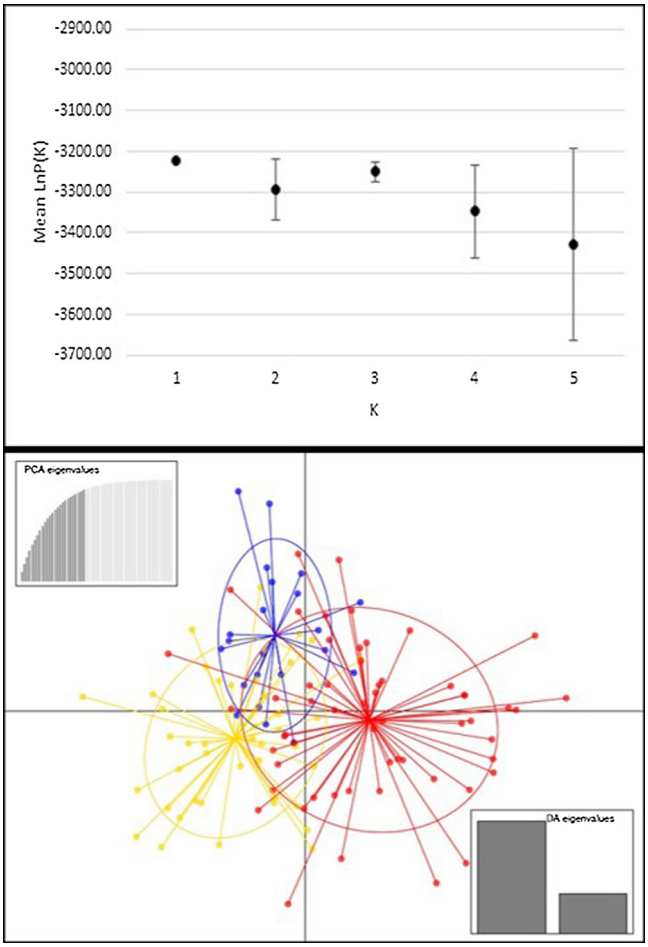

Fig. 2. Population structure analysis using microsatellite genotyped samples from the St. Clair-Detroit River system using A) a clustering analysis, performed using a Bayesian approach implemented in the software program STRUCTURE; and B) a Discriminant Analysis of Principal Components (DAPC), providing a visualization of the location of the individuals and clusters along the first two axes of the DAPC. Geographic sampling sites were used as priors for genetic clusters, and ellipses are 95% inertia ellipses with 1 = Upper St. Clair River, 2 = Lower St. Clair River, and 3 = the Detroit River.

from the SCDRS. After a sequential Bonferroni correction, two populations were out of HWE at multiple loci. The upper St. Clair river was out of HWE at Afu68 ($P = 0.000$), Afu68B ($P = 0.000$), AfuG160 ($P = 0.000$), and Spl120 ($P = 0.003$). The lower St. Clair river was out of HWE at Afu68 ($P = 0.001$), Aox27 ($P = 0.004$), and AfuG56 ($P = 0.014$). The upper St. Clair ($P = 0.000$) and the lower St. Clair ($P = 0.0001$) also exhibited a significant heterozygote deficiency. Loci that are consistently out of HWE are often discarded from subsequent analyses. Only Afu68 was out for both rivers. Removing Afu68 did not affect the STRUCTURE analyses; therefore, all loci were retained because discarding loci might cause subpopulation structure to be overlooked (Dharmarajan et al., 2013).

Bayesian analysis, implemented by the program STRUCTURE, revealed that the most likely number of populations in the SCDRS is one ($K = 1$) (Fig. 2). The upper St. Clair River is significantly differentiated from the lower St. Clair, albeit a low level of differentiation ($F_{ST} = 0.014$) (Table 2). None of the other F_{ST} values were significant, suggesting that neither the upper, nor lower St. Clair River, is genetically differentiated from the Detroit River. The DAPC partitioned samples into three spatially structured clusters (Fig. 2). Using cross validation, the lowest *BIC* value was used to choose K = 2.

Population assignment using MC cross-validation in the assignPOP package resulted in moderate assignment success, with 79% of the fish from the Detroit and lower St. Clair rivers assigning to the upper St. Clair River. Assignment for upper St. Clair River individuals to the upper St. Clair populations was 83%.

3.1.2. St. Lawrence River

After sequential Bonferroni correction, no loci were deemed to be out of HWE and there was no evidence of linkage disequilibrium. Bayesian analysis, implemented by the program STRUCTURE, showed that the most likely number of populations in the St. Lawrence River (SLR) was one ($K = 1$) (Fig. 3). This was the most consistent estimate with the highest likelihood of occurrence.

According to the pairwise F_{ST} values, the highest level of differentiation was detected between Quebec City and Lake St. Francis ($F_{ST} = 0.016$, $P = 0.01$, Table 2b). Such F_{ST} values suggest that sturgeon from both sites belong to genetically distinct populations connected by a relatively high amount of gene flow. The DAPC partitioned samples into three spatially structured clusters (Fig. 3). As with the SCDRS, the sampling locations were used as *a priori* information. The 95% genotpyic intertia ellipses associated with the Moses Suanders Dam, Lake St. Francis, and Quebec City do not overlap; however, there is overlap between the ellipses associated with the Moses Saunders Dam and the Akwesasne sampling sites. Using cross validation, the lowest *BIC* value was used to choose K = 2.

Population assignment using MCMC in the assignPOP package resulted in assignment (96-97%) of all samples to the population associated with Quebec City.

3.1.3. St. Lawrence River and St. Clair Detroit River System

When including samples from both rivers, STRUCTURE Harvester showed that the most likely number of populations with the lowest amount of variance was two (K = 2). F_{ST} values for the SCDRS and the STL showed statistically significant differentiation ($F_{ST} = 0.051$, $P < 0.0014$). DAPC partitioned samples into three spatially structured clusters with high overlap between inertia ellipses between most clusters. Using cross validation, the lowest *BIC* value was used to choose K = 7. Assignment accuracy using microsatellites for both systems ranged from 0-62%, with the greatest assignment accuracy of fish to the Lake St. Francis sampling location (33-62%) and the upper St. Clair River (34-54%).

3.2. Single Nucleotide Polymorphisms

3.2.1. St Clair-Detroit River System

For the SCDRS, 51,747 SNPs were retained after quality filtering and used for population genomic analyses. No loci were deemed to be out of Hardy-Weinberg or under linkage disequilibrium. When Bayesian analysis, implemented by the program fastStructure, was performed on the entire set of loci, the most likely number of populations in the SCDRS was one ($K = 1$). The F_{ST} values also suggested that there was no differentiation between these populations (Table 2a). Comparison of the two rivers using SNPs showed differentiation between the SCDRS and the SLR ($F_{ST} = 0.010$, $P < 0.0001$). Similar to the microsatellite analyses, the DAPC on the whole SNP dataset partitioned samples into spatially structured clusters (Fig. 4a). Using cross validation, the lowest *BIC* value was used to choose K = 2. Assignment accuracy estimated using the assignPOP package on the full set of SNPs ranged from 35-65% with the greatest accuracy (65%) of assignment to the upper St. Clair River.

When analyzing the full dataset, no outlier loci were identified using BayeScan, but 303 loci were identified as outliers using OutFLANK. Results from Bayesian analysis using both fastStructure and STRUCTURE on the outlier dataset alone indicated that the most likely number of populations was still one. The DAPC (Fig. 4) provided similar results to those generated using the whole dataset. None of the F_{ST} values for the comparison of the populations using only outlier loci were significant.

Table 2
The F_{ST} values for microsatellites (presented below the diagonal) and SNPs (presented above the diagonal line) were generated using the method outlined in Weir and Cockerham (1984) for lake sturgeon in (a) the St. Clair-Detroit River System and (b) the St. Lawrence River. The bolded F_{ST} values are statistically significant ($\alpha = 0.05$); corresponding p-values are in parentheses. Significant *P*-values are bolded.

(a)

	Detroit	Lower St. Clair	Upper St. Clair
Detroit		−0.0025 (0.79)	0.017 (0.10)
Lower St. Clair	0.006 (0.88)		−0.0001 (0.55)
Upper St. Clair	0.008 (0.55)	**0.014 (0.02)**	

(b)

	Moses Saunders	Akwesasne	Lake St. Francis	Quebec City
Moses Saunders				**0.002 (P < 0.0001)**
Akwesasne	0.0093 (0.15)			
Lake St. Francis	0.0025 (0.17)	0.0143 (0.02)		
Quebec City	0.0036 (0.08)	0.0256 (0.13)	**0.0164 (0.01)**	

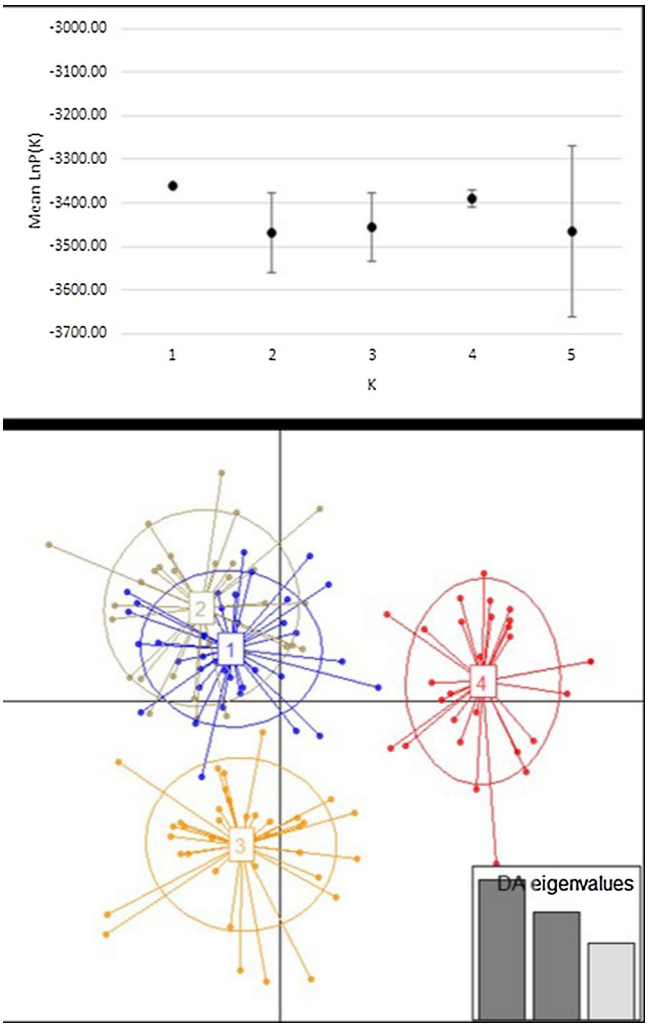

Fig. 3. Population structure analysis using microsatellite genotyped samples from the St. Lawrence River using A) a clustering analysis, performed using a Bayesian approach implemented in the software program STRUCTURE; and B) a Discriminant Analysis of Principal Components (DAPC), providing a visualization of the location of the individuals and clusters along the first two axes of the DAPC. Geographic sampling sites were used as priors for genetic clusters, and ellipses are 95% inertia ellipses with 1 = Moses Saunders Dam, 2 = Akwesasne, 3 = Lake St. Francis, and 4 = Quebec City.

When the high F_{ST} SNP dataset was used, greater population structure was detected, along with higher assignment accuracy. Using fastSTRUCTURE and by choosing the lowest *BIC* value, cross-validation

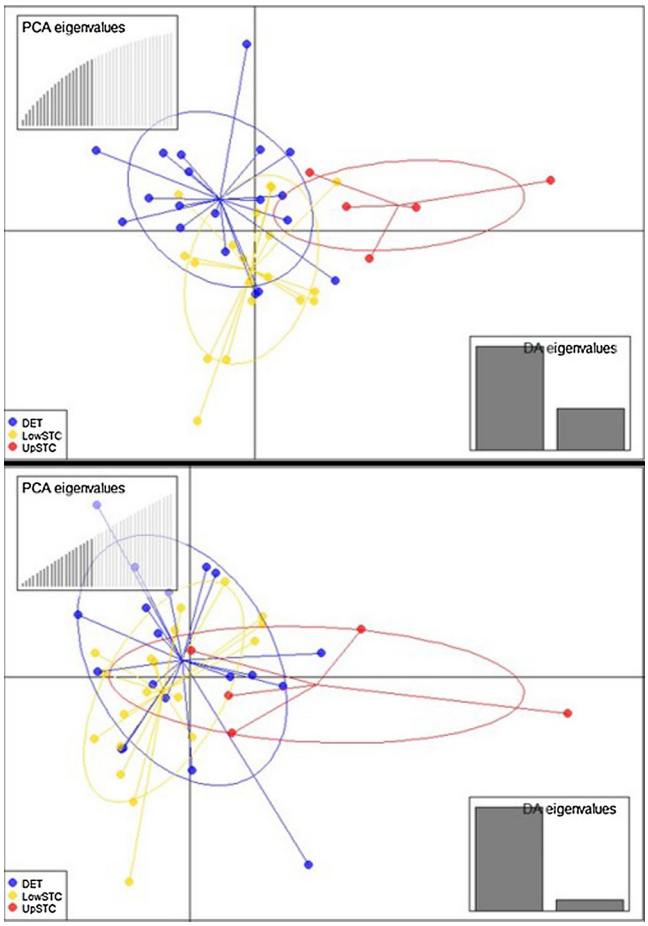

Fig. 4. Discriminant Analysis of Principal Components (DAPC), using A) outlier SNP genotypes and B) neutral SNP genotypes from the St. Clair Detroit River System, providing a visualization of the location of the individuals and clusters along the first two axes of the DAPC. Geographic sampling sites were used as priors for genetic clusters, and ellipses are 95% inertia ellipses with Red = Upper St. Clair River, Yellow = Lower St. Clair River, and Blue = the Detroit River.

in the DAPC resulted in two clusters (K = 2) (Fig. 5). Assignment accuracy was improved with 98% accuracy in assignment of fish to the St. Clair River and 88% accuracy in assignment of fish to the Detroit River. There were 63 private alleles identified for the Detroit River and 128 for the St. Clair River. No outliers were identified in the reduced SNP dataset.

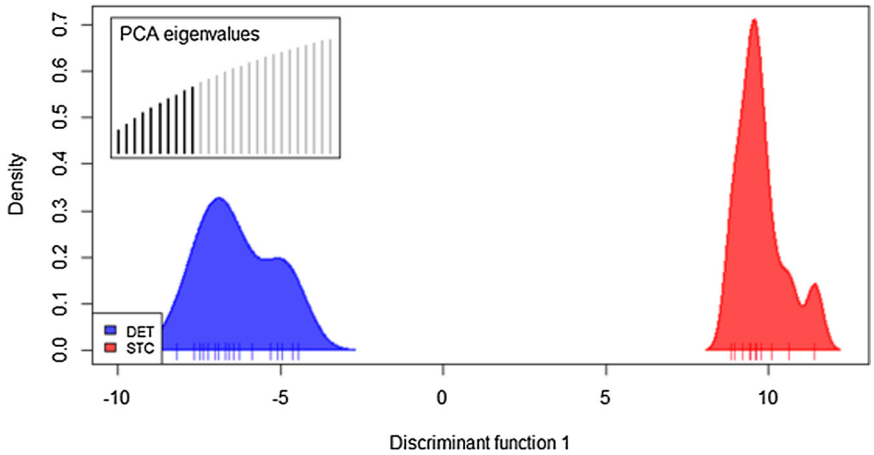

Fig. 5. Discriminant Analysis of Principal Components (DAPC) using SNP genotypes filtered for high F_{ST} values from St. Clair-Detroit River System samples was performed using the adegenet package in R to identify variation in SNP genotypes among population locations (Jombart, 2008).

3.2.2. St. Lawrence River

For the SLR, 30,034 SNPs were retained after quality filtering and used for population genomic analyses. No loci were deemed to be out of Hardy-Weinberg equilibrium or under linkage disequilibrium. Bayesian analysis, implemented by the program fastStructure, showed that the most likely number of populations in the SLR was one ($K = 1$). The F_{ST} value for the comparison of the two sampling locations was 0.002 ($P < 0.0001$), suggesting that there is weak, albeit significant, population structure accompanied by pronounced gene flow between the two sampled populations. The DAPC partitioned samples into spatially structured clusters (Fig. 6). Using cross-validation, two population clusters (K = 2) was chosen based on the lowest *BIC* value. Assignment accuracy using the assignPOP package to analyze the full set of SNPs ranged from 29-71% with the greatest accuracy (71%) of assignment to the sampling location below the Moses-Saunders dam. No outlier loci were identified using either BayeScan or OutFLANK.

As with the SCDRS, the SNP dataset that included only high F_{ST} values (n = 1415 SNPs) showed greater evidence of population structure. Using fastSTRUCTURE and by choosing the lowest *BIC* value, cross-validation in the DAPC resulted in K = 2 (Fig. 6). Assignment accuracy was improved with 97% accuracy in assignment of fish to the Quebec sampling location and 88% accuracy in assignment of fish to the population below the Moses-Saunders dam. There were 206 private alleles identified for the population sampled below the Moses-Saunders dam and 172 identified in Quebec. No outliers were identified in the reduced SNP dataset.

3.2.3. St. Lawrence River and St. Clair Detroit River System

For the comparison between the two rivers, 23,741 SNPs were retained after quality filtering. No loci were deemed to be out of Hardy-Weinberg equilibrium or under linkage disequilibrium. Using Bayesian inference implemented in the program fastSTRUCTURE, the most likely number of populations was one ($K = 1$). The F_{ST} value for the comparison of the two rivers was 0.011 ($P < 0.0001$). By choosing the lowest *BIC* value, cross-validation in the DAPC resulted in $K = 2$. Using assignPOP to perform population assignment, the full SNP dataset had an assignment accuracy of 76% for the SLR and 77% for the SCDRS

The dataset filtered for high F_{ST} values, resulting in a total of 1,415 SNPs, performed similarly for the DAPC with $K = 2$. However, using Bayesian inference, the most likely number of populations was two ($K = 2$) and the F_{ST} value was 0.091 ($P < 0.0001$). Additionally, the dataset filtered for high F_{ST} values for the STL and SCDRS had a greater assignment accuracy of 91% and 93%, respectively. No outliers were identified.

4. Discussion

Marker choice is an important factor in designing genetic studies to delineate populations for management. In this case study, microsatellites and the full SNP dataset performed similarly in the detection of fine-scale population structure in lake sturgeon. However, when filtered for high F_{ST} values, the SNP dataset performed better at detecting population structure and had improved assignment accuracy. Improved

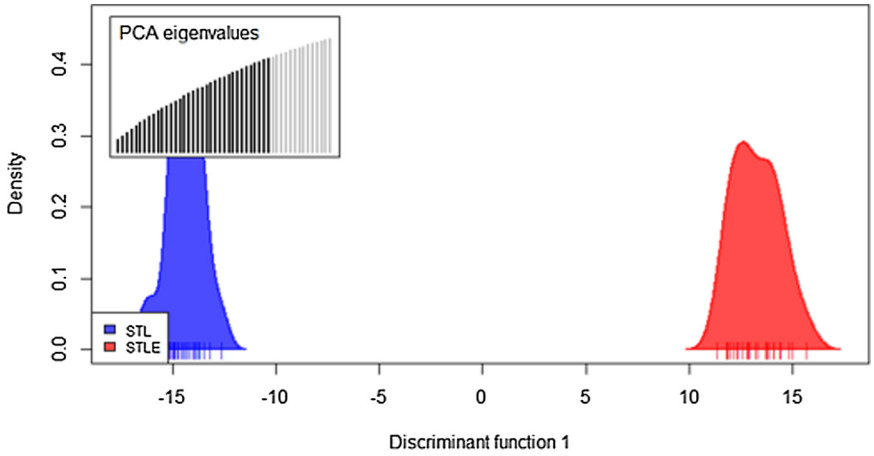

Fig. 6. Discriminant Analysis of Principal Components (DAPC) using SNP genotypes filtered for high F_{ST} values from St. Lawrence River samples was performed using the adegenet package in R to identify variation in SNP genotypes among population locations (Jombart, 2008).

genetic assignment provides more accurate metrics that are valuable for conservation and management, including spatial distribution, migratory patterns, stock-specific exploitation, inbreeding, and parentage (e.g. Allendorf et al., 2010; André et al., 2016). For example, SNPs provided refined estimates of relatedness in *Acipenser sturio*, despite extremely low genetic diversity (Roques et al., 2019). For lake sturgeon specifically, accurate assignment is important for evaluating stocking and reintroduction programs, as well as estimating straying rates and potential impacts to neighboring populations (e.g. outbreeding depression). Additionally, reliable estimates of interspecific hybridization, which occurs fairly commonly in sturgeon (e.g. Tranah et al., 2004; Zhang et al., 2013), could be achieved. Further, due to the long distances that sturgeon are capable of traveling (Auer, 1996), the efficiency of assignment using SNPs could prove highly valuable to identify non-spawning movement patterns that are not yet well understood. SNPs have been used in sturgeon conservation to differentiate between species (e.g. Eichelberger et al., 2014; Ogden et al., 2013) and to differentiate between geographically distant populations (Rastorguev et al., 2013), but to date, there are no published studies of fine-scale genetic structure detected in sturgeon using SNPs.

Adaptive differences are considered especially informative for delineating conservation units, specifically evolutionarily significant units (Fraser and Bernatchez, 2001; Funk et al., 2012). Adaptive markers may be useful in fisheries research by providing a better understanding of anthropogenic pressures, such as climate- and fishing-induced changes to populations, and for conserving the adaptive potential of a species that may encounter unforeseen or unpredictable changes. Further, adaptive divergence may be especially important for the SCDRS and SLR both of which host populations that are large enough that selection may be a strong driving force for differentiation more so than drift, which is more of a concern in smaller populations (André et al., 2011; O'Malley et al., 2007; Perrier et al., 2017; Westgaard and Fevolden, 2007). For instance, Atlantic cod (*Gadus morhua*), a well-studied fishery, shows differentiation along temperature clines with outlier SNPs, a pattern that was overlooked using neutral markers (Bradbury et al. 2010). Loci under selection were identified for the SCDRS, but without a reference genome, it is challenging to determine whether these loci play an important role in adaptive differentiation. Considering the lack of physical barriers in the SCDRS, differentiation between these rivers could be due to known differences in migratory behaviors (Kessel et al., 2018; Whitaker et al., 2018). It should be noted that there are pitfalls associated with the use of adaptive loci, such as the assumption of the future benefit or consequence of such loci, and often the function of adaptive loci is unknown in non-model organisms. Inheritance, recombination, gene interactions and expression can all affect the relationship of the genotype to the phenotype in studies of natural populations. Direct knowledge of the genotype to phenotype relationship eases the translation of genomics into management applications. Future research should focus on examining environmental factors that could potentially contribute to local adaptation despite gene flow in lake sturgeon of the SCDRS.

No loci were determined to be under selection for the SLR. The ease of identification of an adaptive signal can be confounded by, for instance, demography, gene interactions, and expression regulation (Pardo-Diaz et al., 2015; Schoville et al., 2012). Reduced populations, or a reduction in the effective population size, may also affect the ability to detect loci under selection (Jensen and Bachtrog, 2011; Poh et al., 2014). Further, loci with a small effect on phenotypic variation are especially difficult to detect (Brieuc et al., 2018). As more genomic data become available for non-model organisms, alternative approaches, such as genome-wide association studies and a search for candidate pathways, may provide more useful information on adaptive evolution (Pardo-diaz et al., 2015). Alternatively, a landscape genomics approach, where a relationship between genomics and environmental variables is used to identify an adaptive signal, has proven useful in non-model species (e.g. Bourret et al., 2013; Grummer et al., 2019), including in the St. Lawrence system (e.g. Leclerc et al., 2008; Ouellet-Cauchon et al., 2014).

Collectively, the analyses performed in this study suggest that there is evidence of fine-scale population differentiation. While the Bayesian analysis using microsatellites and the full SNP dataset failed to detect structure, the DAPC did show evidence of spatially separated populations and F_{ST} values corroborated some population differentiation. Further, significant population structure was evident in the analyses of the SNP dataset filtered for high F_{ST} values. These SNPs may be more meaningful when detecting population structure in lake sturgeon.

The long generation time of lake sturgeon should be considered when interpreting the results. Identifying high levels of differentiation and fragmentation may require long time periods when the species exhibits late maturation and spawning periodicity (Gibbs, 2001; O'Grady et al., 2008; Wozney et al., 2011). STRUCTURE is known to produce unreliable clustering under some scenarios (Kalinowski, 2011; Waples and Gaggiotti, 2006). Specifically, STRUCTURE may not perform as well when the sampled populations have low to moderate genetic differentiation (0.005-0.136) or they have high gene flow, resulting in the failure to detect the correct number of populations (Waples and Gaggiotti, 2006). This may be the case with both the SCDRS and SLR, resulting in an underestimation of populations using STRUCTURE. In addition, F_{ST} values can be biased by demographic changes, such as decreased population sizes, resulting in underestimation of differentiation (Pearse and Crandall, 2004). Therefore, the low levels of genetic differentiation detected may increase over time, making genetic monitoring of these populations important for their long-term sustainability.

The heterozygote deficiency detected in the microsatellite genotypes from the upper and lower populations of the St. Clair River may be due to sampling multiple subpopulations that are perceived as one (i.e. a Wahlund effect; Tracey et al., 1975). The findings of slight differentiation in the SCDRS using both microsatellites and outlier SNPs support the hypothesis that multiple subpopulations could have been sampled. Historically, the SCDRS has had as many as 15 spawning sites, so it is reasonable to suppose that all possible populations or subpopulations were not independently sampled. Additionally, as many as 14 distinct movement patterns have been identified in the SCDRS (Kessel et al., 2018), which may also be contributing to the slight differentiation of the upper St. Clair River.

Once isolated by dams, populations may experience the loss of diversity from genetic drift, creating differentiation between the populations. Genetic differentiation of upstream populations from other populations has been seen in brown trout (*Salmo trutta*, Heggenes and Røed, 2006; Horreo et al., 2011), Roanoke logperch *Percina rex* (Roberts et al., 2013), and greenside darter *Etheostoma blennioides* (Beneteau et al., 2009). While construction of the Beauharnois–Les Cèdres generating station began in 1912, due to the long generation time of sturgeon, detecting differentiation between these populations may require that the barrier be in place longer. Studies such as Yamamoto et al. (2004) and Haponski et al. (2007) observed that genetic differentiation may increase corresponding to the length of time that the dam has been in place. A variety of variables can affect the level of differentiation, including degree of hydrological modification (Roberts et al., 2013), permeability of the barrier (Reid et al., 2008), habitat size above the dam (Reid et al., 2008; Yamamoto et al., 2004), time since fragmentation (Yamamoto et al., 2004), historical population size (Yamamoto et al., 2004), and migratory capabilities of the species (Rolls et al., 2014).

Population structure in Acipenserids varies by species and management should reflect this variation. Bayesian analysis, isolation-by-distance, and F_{st} values of Lake sturgeon genotypes have failed to identify population structure in other fragmented systems (Welsh and McLeod, 2010; Wozney et al., 2011). Similarly, studies of Atlantic sturgeon (*A. oxyrinchus oxyrinchus*) showed no evidence of genetic structure in fragmented drainages (Grunwald et al., 2008; Waldman

et al., 2002). Within-river analyses of White sturgeon (*A. transmontanus*) in the Columbia-Snake drainage showed evidence of isolation-by-distance, possibly due to downstream movement, but no upstream movement (Schreier et al., 2013). Gulf sturgeon (*A. oxyrinchus destoi*) in the Pascagoula River show evidence of population structure, while the remainder of the distribution only shows evidence of structure between drainages (Dugo et al., 2004). In most cases, the authors of these studies (e.g. Schreier et al., 2013; Wozney et al., 2011) warn that lack of genetic structure in sturgeon should not be interpreted as habitat connectivity, primarily for the reasons mentioned above regarding the long generation time of lake sturgeon and the impact on measures of genetic differentiation.

This study provides important baseline data of population structure between and within both river systems, and shows the potential of SNPs for identifying fine-scale population structure and improved genetic assignment testing in lake sturgeon. Based on the analyses of SNPs filtered for high F_{ST} values, the evidence of within-river population structure suggests there may be a need to manage separate populations within the rivers to maintain previously unrecognized cryptic diversity. Long-term observations of genetic diversity may be necessary to determine an appropriate delineation of management units and to ensure that the management strategies (e.g. stocking and fishing regulations) remain appropriate. Genetic monitoring has become more cost-effective and reliable (Schwartz et al., 2006), allowing for the implementation in ongoing monitoring programs. Long term datasets allow for better interpretation of trends and may be necessary with long-lived and late-maturing species like lake sturgeon exhibiting a weak genetic signal due to long generation times (Schwartz et al., 2006).

CRediT authorship contribution statement

J.M. Whitaker: Data curation, Investigation, Formal analysis, Writing - original draft. **L.E. Price:** Formal analysis, Writing - original draft. **J.C. Boase:** Resources, Validation, Writing - review & editing. **L. Bernatchez:** Validation, Writing - review & editing. **A.B. Welsh:** Conceptualization, Methodology, Formal analysis, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Samples were provided by Justin Chiotti (U.S. Fish and Wildlife Service), Andrew Briggs (U.S. Fish and Wildlife Service), Margaret Hutton (U.S. Fish and Wildlife Service), Jeremy Moore (U.S. Fish and Wildlife Service), Karen Soper (Ontario Ministry of Natural Resources and Forestry), Chris Gignac (Ontario Ministry of Natural Resources and Forestry), Jennifer Smith (University of Windsor), Mike Thomas (Michigan Department of Natural Resources), Doug Carlson (New York State Department of Environmental Conservation), Guy Verreault, and the U.S. Geological Survey – Great Lakes Science Center. Eric Normandeau and Simon Bernatchez (IBIS, Université laval) assisted with data analysis. Funding was provided by the Great Lakes Fishery Trust (Project No. GLFT-1539). This work was also supported by the USDA National Institute of Food and Agriculture, Hatch project WVA00690, and the West Virginia Agricultural and Forestry Experiment Station. We would like to acknowledge the WVU Genomics Core Facility, Morgantown WV for support provided to help make this publication possible.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fishres.2020.105646.

References

Allendorf, F.W., Leary, R.F., 1986. Heterozygosity and fitness in natural populations of animals. In: Soulé, M.E. (Ed.), Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, MA, pp. 57–76.
Allendorf, F.W., Hohenlohe, P.A., Luikart, G., 2010. Genomics and the future of conservation genetics. Nature Reviews Genetics 11, 697.
Amos, W., Balmford, A., 2001. When does conservation genetics matter? Heredity 87 (3), 257–265.
André, C., Larsson, L.C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G.R., Dahlgren, T.G., Hutchinson, W.F., Mariani, S., Mudde, K., Ruzzante, D.E., 2011. Detecting population structure in a high gene-flow species, Atlantic herring (Clupea harengus): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity 106 (2), 270–280.
André, C., Svedäng, H., Knutsen, H., Dahle, G., Jonsson, P., Ring, A.K., Sköld, M., Jorde, P.E., 2016. Population structure in Atlantic cod in the eastern North Sea-Skagerrak-Kattegat: early life stage dispersal and adult migration. BMC Res. Notes 9 (1), 63.
Auer, N.A., 1996. Importance of habitat and migration to sturgeons with emphasis on lake sturgeon. Canadian Journal of Fisheries and Aquatic Sciences. 53 (1), 152–160.
Auer, N.A., 1999. Population characteristics and movements of lake sturgeon in the Sturgeon River and Lake Superior. Journal of Great Lakes Research. 25 (2), 282–293.
Barth, C.C., Anderson, W.G., Henderson, L.M., Peake, S.J., 2011. Home range size and seasonal movement of juvenile lake sturgeon in a large river in the Hudson Bay drainage basin. Transactions of the American Fisheries Society. 140 (6), 1629–1641.
Basu, B.K., Kalff, J., Pinel-Alloul, B., 2000. Midsummer plankton development along a large temperate river: the St. Lawrence River. Canadian Journal of Fisheries and Aquatic Science 57 (S1), 7–15.
Beacham, T.D., Wallace, C.G., Jonsen, K., McIntosh, B., Candy, J.R., Rondeau, E.B., Moore, J.S., Bernatchez, L., Withler, R.E., 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 (ja).
Beneteau, C.L., Mandrak, N.E., Heath, D.D., 2009. The effects of river barriers and range expansion of the population genetic structure and stability in Greenside Darter (Etheostoma blennioides) populations. Conservation Genetics. 10 (2), 477.
Bernatchez, L., 2016. On the maintenance of genetic variation and adaptation to environmental change: considerations from population genomics in fishes. Journal of Fish Biology. 89, 2519–2556.
Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D.T., Barth, J.M.I., Beacham, T.D., Maes, G.E., Martinsohn, J.Th, Miller, K.M., Naish, K., Ovenden, J.R., Primmer, C.R., Suk, H.Y., Therkildsen, N.O., Withler, R.E., 2017. Harnessing the power of genomics to secure the future of seafood. Trends in Ecology and Evolution. 32, 665–680.
Boase, J.C., Diana, J.S., Thomas, M.V., Chiotti, J.A., 2011. Movements and distribution of adult Lake Sturgeon from their spawning site in the St. Clair River, Michigan. Journal of Applied Ichthyology 27 (2), 58–65.
Bourret, V., Kent, M.P., Primmer, C.R., Vasemägi, A., Karlsson, S., Hindar, K., et al., 2013. SNP-array reveals genome-wide patterns of geographical and potential adaptive divergence across the natural range of Atlantic salmon (*Salmo salar*). Molecular Ecology 22, 532–551.
Bradbury, I.R., Hubert, S., Higgins, B., et al., 2010. Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. Proc R Soc B: Biological Sciences 277, 3725–3734.
Brieuc, M.S., Waters, C.D., Drinan, D.P., Naish, K.A., 2018. A practical introduction to Random Forest for genetic association studies in ecology and evolution. Molecular Ecology Resources 18 (4), 755–766.
Caron, F., Hatin, D., Fortin, R., 2002. Biological characteristics of adult Atlantic sturgeon (*Acipenser oxyrinchus*) in the St Lawrence River estuary and the effectiveness of management rules. Journal of Applied Ichthyology 18 (4-6), 580–585.
Caswell, N.M., Peterson, D.L., Manny, B.A., Kennedy, G.W., 2004. Spawning by lake sturgeon (*Acipenser fulvescens*) in the Detroit River. Journal of Applied Ichthyology 20 (1), 1–6.
Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22 (11), 3124–3140.
Chen, K.Y., Marschall, E.A., Sovic, M.G., Fries, A.C., Gibbs, H.L., Ludsin, S.A., 2018. assign POP: 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.
Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics 12 (1), 499.
DeHaan, P.W., Libants, S.V., Elliott, R.F., Scribner, K.T., 2006. Genetic population structure of remnant lake sturgeon populations in the Upper Great Lakes Basin. Transactions of the American Fisheries Society 135 (6), 1478–1492.
Dharmarajan, G., Beatty, W.S., Rhodes, O.E., 2013. Heterozygote deficiencies caused by a Wahlund effect: dispelling unfounded expectations. The Journal of Wildlife Management 77 (2), 226–234.
Dugo, M.A., Kreiser, B.R., Ross, S.T., Slack, W.T., Heise, R.J., Bowen, B.R., 2004. Conservation and management implications of fine-scale genetic structure of Gulf

sturgeon in the Pascagoula River. Mississippi. Journal of Applied Ichthyology 20 (4), 243–251.
Eichelberger, J.S., Braaten, P.J., Fuller, D.B., Krampe, M.S., Heist, E.J., 2014. Novel single-nucleotide polymorphism markers confirm successful spawning of endangered Pallid Sturgeon in the upper Missouri River basin. Transactions of the American Fisheries Society. 143 (6), 1373–1385.
Edsall, T.A., Manny, B.A., Raphael, C.N., 1988. The St. Clair River and Lake St. Clair, Michigan: an ecological profile. US Fish and Wildlife Service, Biological Report. 85 (7.3). Slidell, LA. .
Foll, M., Gaggiotti, O.E., 2008. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics. 180, 977–993.
Fraser, D.J., Bernatchez, L., 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology. 10 (12), 2741–2752.
Funk, W.C., McKay, J.K., Hohenlohe, P.A., Allendorf, F.W., 2012. Harnessing genomics for delineating conservation units. Trends in Ecology & Evolution. 27 (9), 489–496.
Gibbs, J.P., 2001. Demography versus habitat fragmentation as determinants of genetic variation in wild populations. Biological Conservation. 100 (1), 15–20.
Gilpin, M.E., Soulé, M.E., 1986. Minimum viable populations: processes of species extinction. In: Soulé, M.E. (Ed.), Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, MA, pp. 19–34.
Goodyear, C.S., Edsall, T.A., Ormsby, Dempsey, D.M., Moss, G.D., Polanski, P.E., 1982. Atlas of spawning and nursery areas of Great Lakes fishes: Detroit River. 8. US Fish and Wildlife Service, Washington, DC FWS/OBS 82/52.
Gosselin, T., Lamothe, M., Devloo-Delva, F., Grewe, P., 2020. radiator: RADseq Data Exploration, Manipulation and Visualization using R. https://doi.org/10.5281/zenodo.3687060. https://thierrygosselin.github.io/radiator/.
Goudet, J., 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. Journal of Heredity. 86 (6), 485–486.
Gunderman, B.J., Elliott, R.F., 2004. Assessment of remnant lake sturgeon populations in the Green Bay Basin, 2000–2003. Report to the Great Lakes Fishery Trust, Project 2001.113. Lansing, Michigan. .
Gruber, B., Unmack, P.J., Berry, O.F., Georges, A., 2018. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Molecular Ecology Resources. 18 (3), 691–699.
Grummer, J.A., Beheregaray, L., Bernatchez, L., Hand, B., Luikart, G., Narum, S.R., Taylor, E.B., 2019. Aquatic landscape genomics and environmental effects on genetic variation. Trends in Ecology & Evolution. 34 (7), 641–654.
Grunwald, C., Maceda, L., Waldman, J., Stabile, J., Wirgin, I., 2008. Conservation of Atlantic sturgeon Acipenser oxyrinchus oxyrinchus: delineation of stock structure and distinct population segments. Conservation Genetics 9 (5), 1111–1124.
Haponski, A.E., Marth, T.A., Stepien, C.A., 2007. Genetic divergence across a low-head dam: a preliminary analysis using logperch and greenside darters. Journal of Great Lakes Research 33 (S2), 117–126.
Hatin, D., Fortin, R., Caron, F., 2002. Movements and aggregation areas of adult Atlantic sturgeon (*Acipenser oxyrinchus*) in the St Lawrence River estuary, Quebec, Canada. Journal of Applied Ichthyology. 18 (46), 586–594.
Hayes, D.B., Caroffino, D.C., 2012. Michigan's lake sturgeon rehabilitation strategy. Michigan Department of Natural Resources, Fisheries Division Special Report. 62.
Heggenes, J., Røed, K.H., 2006. Do dams increase genetic diversity in brown trout (Salmo trutta)? Microgeographic differentiation in a fragmented river. Ecology of Freshwater Fish 15 (4), 366–375.
Hess, J.E., Matala, A.P., Narum, S.R., 2011. Comparison of SNPs and microsatellites for fine-scale application ofgenetic stock identification of Chinook salmon in the Columbia River Basin. Molecular Ecology Resources. 11 (1), 137–149.
Holst, L., 2018. Lake Sturgeon Recovery Plan 2018-2024. New York State Department of Environmental Conservation.
Horreo, J.L., Martinez, J.L., Ayllon, F., Pola, I.G., Monteoliva, J.A., Heland, M., Garcia-Vazquez, E.V.A., 2011. Impact of fragmentation on the genetics of populations in dendricitic landscapes. Freshwater Biology. 56 (12), 2567–2579.
Hudon, C., Paquet, S., Jarry, V., 1996. Downstream variations of phytoplankton in the St. Lawrence River (Québec, Canada). Hydrobiologia. 337 (1), 11–26.
Jensen, J.D., Bachtrog, D., 2011. Characterizing the influence of effective population size on the rate of adaptation: Gillespie's Darwin domain. Genome Biology and Evolution. 3, 687–701.
Jombart, T., 2008. Adegenet: an R package for the multivariate analysis of genetic markers. Bioinformatics. 24 (11), 1403–1405.
Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics. 11 (1), 94.
Kalinowski, S.T., 2011. The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. Heredity. 106 (4), 625.
Kang, J., Ma, X., He, S., 2017. Population genetics analysis of the Nujiang catfish Creteuchiloglanis macropterus through a genome-wide single nucleotide polymorphisms resource generated by RAD-seq. Scientific reports 7 (1), 1–13.
Kessel, S.T., Hondorp, D.W., Holbrook, C.M., Boase, J.C., Chiotti, J.A., Thomas, M.V., Wills, T.C., Roseman, E.F., Drouin, R., Krueger, C.C., 2018. Divergent migration within lake sturgeon (*Acipenser fulvescens*) populations: Multiple distinct patterns exist across an unrestricted migration corridor. Journal of Animal Ecology. 87 (1), 259–273.
Leclerc, E., Mailhot, Y., Mingelbier, M., Bernatchez, L., 2008. The landscape genetics of yellow perch (*Perca flavescens*) in a large fluvial ecosystem. Molecular ecology 17 (7), 1702–1717.
Mailhot, Y., Dumont, P., Vachon, N., 2011. Management of the lake sturgeon *Acipenser fulvescens* population in the lower St Lawrence River (Quebec, Canada) from the 1910s to the present. Journal of Applied Ichthyology. 27 (2), 405–410.
Mamoozadeh, N.R., Graves, J.E., McDowell, J.R., 2019. Genome-wide SNPs resolve spatiotemporal patterns of connectivity within striped marlin (*Kajikia audax*), a broadly distributed and highly migratory pelagic species. Evolutionary Applications.
Manny, B.A., Kennedy, G.W., Allen, J.D., French III, J.R.P., 2007. First evidence of egg deposition by walleye (*Sander vitreus*) in the Detroit River. Journal of Great Lakes Research. 33 (2), 512–516.
McDougall, C.A., Nelson, P.A., Macdonald, D., Kroeker, D., Kansas, K., Barth, C.C., MacDonell, D.S., 2017. Habitat quantity required to support self-sustaining lake sturgeon populations: An alternative hypothesis. Transactions of the American Fisheries Society. 146 (6), 1137–1155.
McKinney, G.J., Waples, R.K., Seeb, L.W., Seeb, J.E., 2017. 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.
McMahon, B.J., Teeling, E.C., Höglund, J., 2014. How and why should we implement genomics into conservation? Evolutionary Applications. 7 (9), 999–1007.
Morgans, C.L., Cooke, G.M., Ord, T.J., 2014. How populations differentiate despite gene flow: sexual and natural selection drive phenotypic divergence within a land fish, the Pacific leaping blenny. BMC Evolutionary Biology. 14 (1), 97.
Morin, P.A., Luikart, G., Wayne, R.K., 2004. SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution. 19 (4), 208–216.
Narum, S.R., Banks, M., Beacham, T.D., Bellinger, M.R., Campbell, M.R., Dekoning, J., Elz, A., Guthrie III, C.M., Kozfkay, C., Miller, K.M., Moran, P., 2008. Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Molecular Ecology. 17 (15), 3464–3477.
Ogden, R., Gharbi, K., Mugue, N., Martinsohn, J., Senn, H., Davey, J.W., Pourkazemi, M., McEwing, R., Eland, C., Vidotto, M., Sergeev, A., 2013. Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. Molecular Ecology. 22 (11), 3112–3123.
O'Grady, J.J., Reed, D.H., Brook, B.W., Frankham, R., 2008. Extinction risk scales better to generations than to years. Animal Conservation. 11 (5), 442–451.
O'Malley, K.G., Camara, M.D., Banks, M.A., 2007. Candidate loci reveal genetic differentiation between temporally divergent migratory runs of Chinook salmon (Oncorhynchus tshawytscha). Mol. Ecol. 16 (23), 4930–4941.
Ortiz, E.M., 2018. paralog-finder v1.0: detect and blacklist paralog RAD loci. https://doi.org/10.5281/zenodo.1209328.
Ouellet-Cauchon, G., Mingelbier, M., Lecomte, F., Bernatchez, L., 2014. Landscape variability explains spatial pattern of population structure of northern pike (*E sox lucius*) in a large fluvial system. Ecology and evolution 4 (19), 3723–3735.
Pardo-Diaz, C., Salazar, C., Jiggins, C.D., 2015. Towards the identification of the loci of adaptive evolution. Methods in Ecology and Evolution. 6 (4), 445–464.
Pearse, D.E., Crandall, K.A., 2004. Beyond F ST: Analysis of population genetic data for conservation. Conservation Genetics. 5, 585–602.
Perrier, C., Ferchaud, A.-L., Sirois, P., Thibault, I., Bernatchez, L., 2017. Do tremendous genetic drift and accumulation of deleterious mutations preclude adaptation to temperature in a northern lacustrine fish? Molecular Ecology 26, 6317–6335.
Poh, Y.P., Domingues, V.S., Hoekstra, H.E., Jensen, J.D., 2014. On the prospect of identifying adaptive loci in recently bottlenecked populations. PloS one 9 (11) p.e110579.
Poland, J.A., Rife, T.W., 2012. Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5 (3), 92–102.
Pratt, T.C., 2008. Population status and threats of lake sturgeon in designatable unit 8 (Great Lakes/St. Lawrence River watersheds). Canadian Science Advisory Secretariat. .
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics. 155 (2), 945–959.
Raj, A., Stephens, M., Pritchard, J.K., 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics. 197, 573–589.
Rastorguev, S.M., Nedoluzhko, A.V., Mazur, A.M., Gruzdeva, N.M., Volkov, A.A., Barmintseva, A.E., Mugue, N.S., Prokhortchouk, E.B., 2013. High throughput SNP genotyping analysis of the relationships among Ponto-Caspian sturgeon species. Ecology and Evolution. 3 (8), 2612–2618.
Raymond, M., Rousset, F., 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity. 86 (3), 248–249.
Read, J., Manny, B.A., 2006. Monitoring element of the Belle Isle/ Detroit River Sturgeon Habitat Restoration, Monitoring, and Education Project. Research completion report to Michigan Sea Grant Program. Univ. of Michigan, Ann Arbor, MI, pp. 85.
Reid, S.M., Mandrak, N.E., Carl, L.M., Wilson, C.C., 2008. Influence of dams and habitat condition on the distribution of redhorse (Moxostoma) species in the Grand River watershed, Ontario. Environmental Biology of Fishes. 81 (1), 111–125.
Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution. 43 (1), 223–225.
Roberts, J.H., Angermeier, P.L., Hallerman, E.M., 2013. Distance, dams and drift: what structures populations of an endangered, benthic stream fish? Freshwater Biology. 58 (10), 2050–2064.
Rolls, R.J., Stewart-Koster, B., Ellison, T., Faggotter, S., Roberts, D.T., 2014. Multiple factors determine the effect of anthropogenic barriers to connectivity on riverine fish. Biodiversity and Conservation. 23, 2201–2220.
Roseman, E.F., Manny, B.A., Boase, J., Child, M., Kennedy, G.W., Craig, J., Soper, K., Drouin, R., 2011. Lake sturgeon response to a spawning reef constructed in the Detroit River. Journal of Applied Ichthyology 27 (S2), 66–76.
Roques, S., Chancerel, E., Boury, C., Pierre, M., Acolas, M.L., 2019. From microsatellites to single nucleotide polymorphisms for the genetic monitoring of a critically endangered sturgeon. Ecol. Evol. 9 (12), 7017–7029.
Schoville, S.D., Bonin, A., François, O., Lobreaux, S., Melodelima, C., Manel, S., 2012. Adaptive genetic variation on the landscape: methods and cases. Annual Review of Ecology, Evolution, and Systematics. 43, 23–43.

Schreier, A.D., Mahardja, B., May, B., 2013. Patterns of population structure vary across the range of the white sturgeon. Transactions of the American Fisheries Society. 142 (5), 1273–1286.
Schwartz, M.K., Cushman, S.A., McKelvey, K.S., Hayden, J., Engkjer, C., 2006. Detecting genotyping errors and describing American black bear movement in northern Idaho. Ursus. 17 (2), 138–148.
Thomas, M.V., Haas, R.C., 2002. Abundance, age structure, and spatial distribution of lake sturgeon (*Acipenser fulvescens*) in the St. Clair System. Journal of Applied Ichthyology. 18 (4-6), 495–501.
Tracey, M.L., Bellet, N.F., Gravem, C.D., 1975. Excess allozyme homozygosity and breeding population structure in the mussel Mytilus californianus. Marine Biology. 32 (3), 303–311.
Tranah, G., Campton, D.E., May, B., 2004. Genetic evidence for hybridization of pallid and shovelnose sturgeon. J. Heredity 95 (6), 474–480.
Väli, Ü., Saag, P., Dombrovski, V., Meyburg, B.U., Maciorowski, G., Mizera, T., Treinys, R., Fagerberg, S., 2010. Microsatellites and single nucleotide polymorphisms in avian hybrid identification: a comparative case study. Journal of Avian Biology 41 (1), 34–49.
Van Doornik, D.M., Berejikian, B.A., Campbell, L.A., 2013. Gene flow between sympatric life history forms of *Oncorhynchus mykiss* located above and below migratory barriers. PLoS One. 8 (11), e79931.
Vignal, A., Milan, D., SanCristobal, M., Eggen, A., 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution. 34 (3), 275.
Waldman, J.R., Grunwald, C., Stabile, J., Wirgin, I., 2002. Impacts of life history and biogeography on the genetic stock structure of Atlantic sturgeon Acipenser oxyrinchus oxyrinchus, Gulf sturgeon A. oxyrinchus desotoi, and shortnose sturgeon *A. brevirostrum*. Journal of Applied Ichthyology. 18 (46), 509–518.
Waples, R.S., Gaggiotti, O., 2006. INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular ecology 15 (6), 1419–1439.
Weir, B.S., Cockerham, C.S., 1984. Estimating F-statistics for the analysis of population structure. Evolution. 38 (6), 1358–1370.
Welsh, A.B., Elliott, R.F., Scribner, K.T., Quinlan, H.R., Baker, E.A., Eggold, B.T., Holtgren, J.M., Krueger, C.C., May, B., 2010. Genetic guidelines for the stocking of lake sturgeon (*Acipenser fulvescens*) in the Great Lakes basin. Great Lakes Fishery Commission Miscellaneous Publication 2010-01. .
Welsh, A.B., McLeod, D.T., 2010. Detection of natural barriers to movement of lake sturgeon (Acipenser fulvescens) within the Namakan River, Ontario. Canadian Journal of Zoology. 88 (4), 390–397.
Welsh, A.B., Hill, T., Quinlan, H., Robinson, C., May, B., 2008. Genetic assessment of lake sturgeon population structure in the Laurentian Great Lakes. North American Journal of Fisheries Management. 28 (2), 572–591.
Welsh, A.B., May, B., 2006. Development and standardization of disomic microsatellite markers for lake sturgeon genetic studies. Journal of Applied Ichthyology. 22 (5), 337–344.
Westgaard, J.I., Fevolden, S.E., 2007. Atlantic cod (Gadus morhua L.) in inner and outer coastal zones of northern Norway display divergent genetic signature at non-neutral loci. Fish. Res. 85 (3), 306–315.
Whitaker, J.M., Welsh, A.B., Hondorp, D.W., Boase, J.C., Merovich, G.T., Welsh, S., Krueger, C., 2018. Variation in DNA methylation is associated with migratory phenotypes of lake sturgeon *Acipenser fulvescens* in the St. Clair River, MI, USA. Journal of Fish Biology. 93 (5), 942–951.
Whitlock, M.C., Lotterhos, K.E., 2015. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of F ST. Am. Nat. 186 (S1), S24–S36.
Wozney, K.M., Haxton, T.J., Kjartanson, S., Wilson, C.C., 2011. Genetic assessment of lake sturgeon (*Acipenser fulvescens*) population structure in the Ottawa River. Environmental Biology of Fishes. 90 (2), 183–195.
Yamamoto, S., Morita, K., Koizumi, I., Maekawa, K., 2004. Genetic differentiation of white-spotted charr (*Salvelinus leucomaenis*) populations after habitat fragmentation: spatial–temporal changes in gene frequencies. Conservation Genetics. 5 (4), 529–538.
Zhang, X., Wu, W., Li, L., Ma, X., Chen, J., 2013. Genetic variation and relationships of seven sturgeon species and ten interspecific hybrids. Genetics Selection Evolution. 45 (1), 21.