Transactions of the American Fisheries Society

ISSN: 0002-8487 print / 1548-8659 online
DOI: 10.1002/tafs.10225

ARTICLE

Genetic Family Reconstruction Characterizes Lake Sturgeon Use of Newly Constructed Spawning Habitat and Larval Dispersal

Robert D. Hunter*
Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, Michigan 48824, USA

Edward F. Roseman
U.S. Geological Survey, Great Lakes Science Center, 1451 Green Road, Ann Arbor, Michigan 48105, USA

Nicholas M. Sard
Department of Biological Sciences, State University of New York–Oswego, Centennial Drive, Oswego, New York 13126, USA

Robin L. DeBruyne
Department of Environmental Sciences, University of Toledo Lake Erie Center, 6200 Bayshore Road, Oregon, Ohio 43616, USA

Jinliang Wang
Zoological Society of London, Institute of Zoology, London NW1 4RY, UK

Kim T. Scribner
Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, Michigan 48824, USA and; Department of Integrative Biology, Michigan State University, East Lansing, Michigan 48824, USA

Abstract

Since 2004, seven spawning reefs have been constructed in the St. Clair–Detroit River system to remediate lost spawning habitat and increase recruitment of Lake Sturgeon *Acipenser fulvescens*. Assessment of management actions by collecting and enumerating eggs and larvae provided evidence of spawning Lake Sturgeon and survival of eggs until larval dispersal at constructed reef sites. However, the number of spawners contributing sampled offspring (N_s), effective number of breeders (N_b), and extent of larval dispersal was unknown. Genetic reconstruction of familial relationships assigned eggs and larvae ($n = 725$) collected in 2015 and 2016 to full- and half-sibling groups and estimated N_s, N_b, and genetic connectivity. We used a modified COLONY simulation module to simulate and convert 18 microsatellite loci (13 disomic and 5 polysomic) to 205 dominant present/absent markers to increase marker number and familial assignment accuracy in family reconstruction analysis. We assessed COLONY's ability to accurately infer familial relationships across small ($n = 50$), moderate ($n = 125$), and large ($n = 750$) larval sample sizes using two assumed allele frequency distributions for polysomic loci. We found that with fewer offspring sampled, COLONY underestimated N_s and with large sample sizes overestimated N_s. However, estimates were usually within 12–16% of the simulated true N_s. Across reefs, estimates of N_s were 151 in 2015 and 208 in 2016, and N_b was similar (158 in 2015 and 198 in 2016). Evidence of full- and half-sibling larvae collected at multiple locations indicated that

*Corresponding author: hunterr2@msu.edu
Received November 4, 2019; accepted January 9, 2020

individual Lake Sturgeon spawned at multiple locations within years and larvae dispersed considerable distances. Estimating N_s, N_b, larval dispersal, and inferred genetic connectivity between locations provides managers with population demographic parameters to assess habitat remediation projects. Continued monitoring, including genetic family reconstruction, may provide insight into the long-term effects of constructed spawning habitat on recruitment and population-level genetic diversity.

Conservation efforts for threatened and endangered species are often constrained by aspects of a species' ecology and abundance and by the complexity of habitats they occupy. One example is Lake Sturgeon *Acipenser fulvescens*, which was once abundant in the Great Lakes region and Mississippi River basin (Thomas and Haas 2002). However, overharvest and habitat degradation resulted in declines in Lake Sturgeon population abundance and distribution (Auer 1996; Roseman et al. 2011). Currently, most Lake Sturgeon populations are estimated to be at less than 1% of historic abundance (Auer 1996; Hay-Chmielewski and Whelan 1997) or have been extirpated.

The St. Clair–Detroit River system (SCDRS) is believed to contain one of the largest remnant populations of Lake Sturgeon in the Great Lakes (Thomas and Haas 2002). However, assessment work throughout the SCDRS identified a lack of spawning habitat as a limiting factor for conservation of lithophilic spawners, including Lake Sturgeon (Hondorp et al. 2014; Manny et al. 2015). In response to this conservation need, seven spawning reefs have been constructed in the SCDRS since 2004 to remediate lost spawning habitat (Manny et al. 2015). Assessments to verify use by spawning Lake Sturgeon detected eggs on the new reefs, where few or no eggs were detected prior to construction, indicating that Lake Sturgeon immediately used the artificial reefs for spawning following initial construction (Roseman et al. 2011; Bouckaert et al. 2014; Manny et al. 2015; Prichard et al. 2017; Fischer et al. 2018). Additional assessment during the larval drift period demonstrated that Lake Sturgeon eggs deposited on the reefs survived to the larval drift stage (Roseman et al. 2011; Bouckaert et al. 2014).

Recruitment to the larval stage is one measure of success for the constructed spawning reefs (McLean et al. 2015). However, if spawning adults formerly dispersed across the few remaining spawning sites become spatially and numerically distributed over a larger number of smaller spawning sites there is a potential for depensatory effects. Genetic diversity may be reduced by limiting the number of mates that individuals encounter if comparatively fewer individuals spawn at each of the expanded number of individual reefs. Additionally, Lake Sturgeon females and males are polygamous broadcast spawners, and a reduction in individual reproductive success could result from density-dependent reductions in fertilization rates (e.g., sperm limitation), which have been observed in other broadcast spawners (Levitan and Petersen 1995; Levitan 2004). However, depensatory effects may be overcome if adults spawn in more than one location in a single spawning season. Spawning in multiple locations may contribute to a beneficial spawning portfolio effect (Schindler et al. 2010). Spawning portfolio effects may buffer against population-level loss of recruits and genetic diversity by mitigating stochastic mortality events at individual spawning locations. Additionally, concern exists regarding the fate of larvae following hatch and larval dispersal. Little is known about the dispersal of Lake Sturgeon larvae in this or other large river systems or of habitats required following drift for growth and survival to the juvenile stage (Boase et al. 2014). Therefore, questions remain regarding the following: (1) how many spawners contribute offspring at a reef, (2) whether adults spawn in more than one location in a single year, (3) and what distance larvae disperse through the SCDRS.

Genetic techniques such as family reconstruction can provide important information to address management questions (Baetscher et al. 2018; Steele et al. 2019) by identifying individuals and family groups to estimate ecological and evolutionary parameters required for assessment and monitoring (Schwartz et al. 2007). Using genetic data to reconstruct familial relationships, the number of spawners contributing sampled offspring (N_s), the effective number of breeders (N_b), and the mean and variance in individual reproductive success of adults contributing offspring at a location can be estimated, within and across years (e.g., Williamson et al. 2010; Duong et al. 2013; Sard et al. 2015). Estimates of N_s, N_b, and reproductive success (mean and variance) can quantify levels of use of constructed reefs by spawning Lake Sturgeon and are essential means of measuring postconstruction recruitment. The N_s quantified in a sample of progeny (eggs or larvae) can provide a relative index of reef use that can be compared among reefs and over time, assuming equal sampling and catchability between sites and years. The effective number of breeders is the breeding population size within a single spawning season, where the rate of change in allelic diversity is equal to the rate of change by genetic drift alone (Wright 1931). Despite that N_b refers to a population and N_s refers to a sample drawn from the population, N_b is

typically lower than N_s due to skewed sex ratios and variance in individual reproductive success (Waples 1990). When N_b is low, the risk of population-level losses of genetic diversity increases due to a large proportion of offspring being contributed by a relatively small number of spawning adults. Estimates of variation in reproductive success and measures of genetic diversity are also important metrics to inform management and help guide future habitat remediation efforts.

Knowledge of larval dispersal distances from spawning sites, residence times in different river sections, and habitat features associated with larval growth and survival are also critical information needs. In large barrier-free river systems such as the SCDRS, sampling is difficult, and it is possible for larvae to disperse great distances from their spawning location. Without the application of genetic family reconstruction, estimating the degree to which larvae disperse from local spawning reefs in large barrier-free river systems such as the SCDRS would be infeasible. Here we used inferred sibship (full-sibling, half-sibling, and nonsibling) relationships of larvae collected as eggs from spawning reefs and dispersing larvae subsequently captured in downstream nets to describe patterns of larval dispersal. Quantifying larval dispersal distance may inform placement of future habitat remediation efforts in large river systems and allow for identification of juvenile rearing habitat within the dispersal range of current or proposed constructed spawning sites.

The overall objective of this study was to estimate N_s and N_b to quantify spawning success and recruitment dynamics of adults contributing offspring at constructed reefs. Specifically, we used genetic family reconstruction (1) to estimate the minimum number of adults contributing sampled offspring at constructed spawning reefs (N_s), (2) to estimate the effective number of breeders (N_b) and the mean and variance in individual reproductive success, (3) to determine if adults spawned at multiple locations within a spawning season, and (4) to describe patterns in larval dispersal between reef sites.

METHODS

Study area.— The SCDRS is a 145-km, barrier-free waterway flowing from Lake Huron into Lake Erie that includes the St. Clair River, Lake St. Clair, and the Detroit River (Figure 1). The system has been heavily modified by dredging of shipping channels, draining of wetlands, and hardening of riparian areas to accommodate anthropogenic uses. The St. Clair and Detroit rivers were declared Great Lakes Areas of Concern in 1987 due in part to the loss of fish and wildlife habitat. Construction of shipping channels in the SCDRS has resulted in the removal of 46.2 million cubic meters of substrate, and disposal of those dredge spoils has resulted in the burial of approximately 4,000 ha of benthic habitat (Bennion and 2011). Subsequently, the number of known Lake Sturgeon spawning sites was reduced from 15 to 3 (Goodyear et al. 1982; Manny and Kennedy 2002; Nichols et al. 2003). Despite impairments, the SCDRS is believed to contain one of the largest remnant populations of Lake Sturgeon (Thomas and Haas 2002) in the Great Lakes. The largest naturally occurring spawning location is thought to be at the head of the St. Clair River under the Blue Water Bridge, Port Huron, Michigan (Manny and Kennedy 2002). Additional spawning sites have been detected in the North Channel of the St. Clair River (Thomas and Haas 1999; Manny and Kennedy 2002) and near Zug Island in the Detroit River (Manny and Kennedy 2002; Caswell et al. 2004). The results of these extensive efforts to locate natural spawning areas for Lake Sturgeon in the SCDRS identified lack of spawning habitat as a factor limiting Lake Sturgeon spawning success (Manny et al. 2015).

To increase suitable spawning habitat for lithophilic spawners such as Lake Sturgeon, seven spawning reefs were constructed with 8–15-cm-diameter fractured limestone throughout the SCDRS (Manny et al. 2015). To document the use of the newly constructed reefs by spawning Lake Sturgeon, eggs and drifting larvae were sampled postconstruction for 2 years. Three newly constructed reefs were sampled in 2015 and 2016. Harts Light reef, constructed in 2014 and sampled in 2015 and 2016, is the northernmost constructed reef site in the St. Clair River (Figure 1). Harts Light reef is 1.5 ha in size with water depths exceeding 16 m and stream flow at the time of drift up to 1.4 m^3/s. Downstream from Harts Light reef is Pointe Aux Chenes reef (Figure 1). Also constructed in 2014 and sampled in 2015 and 2016, this 0.6 ha reef was down to 15 m deep, with ~~stream flow~~ at the time of larval dispersal up to 1.2 m^3/s. Further downstream from Pointe Aux Chenes reef in the St. Clair River were the North Channel and Middle Channel control sites (Figure 1) that were sampled in 2015 and 2016 to assess spawning at alternate locations and larval dispersal from our study locations and other upstream spawning locations. Grassy Island reef was constructed in 2015 in the Detroit River (Figure 1) and sampled in 2016. Grassy Island reef is a 1.6 ha reef with depths down to 12 m and ~~stream flow~~ at the time of drift up to 0.8 m^3/s. Additional known natural spawning sites in the study area were in the North Channel (Maslinkas reef, a historic spawning site) and Middle Channel (Middle Channel reef, constructed in 2012) of the St. Clair River, and there is a suspected spawning site (exact location unknown) in the Trenton Channel of the Detroit River.

Sample collection.— All samples were collected in 2015 and 2016 in the St. Clair River and in 2016 in the Detroit River. Detailed information on all aspects of field sample

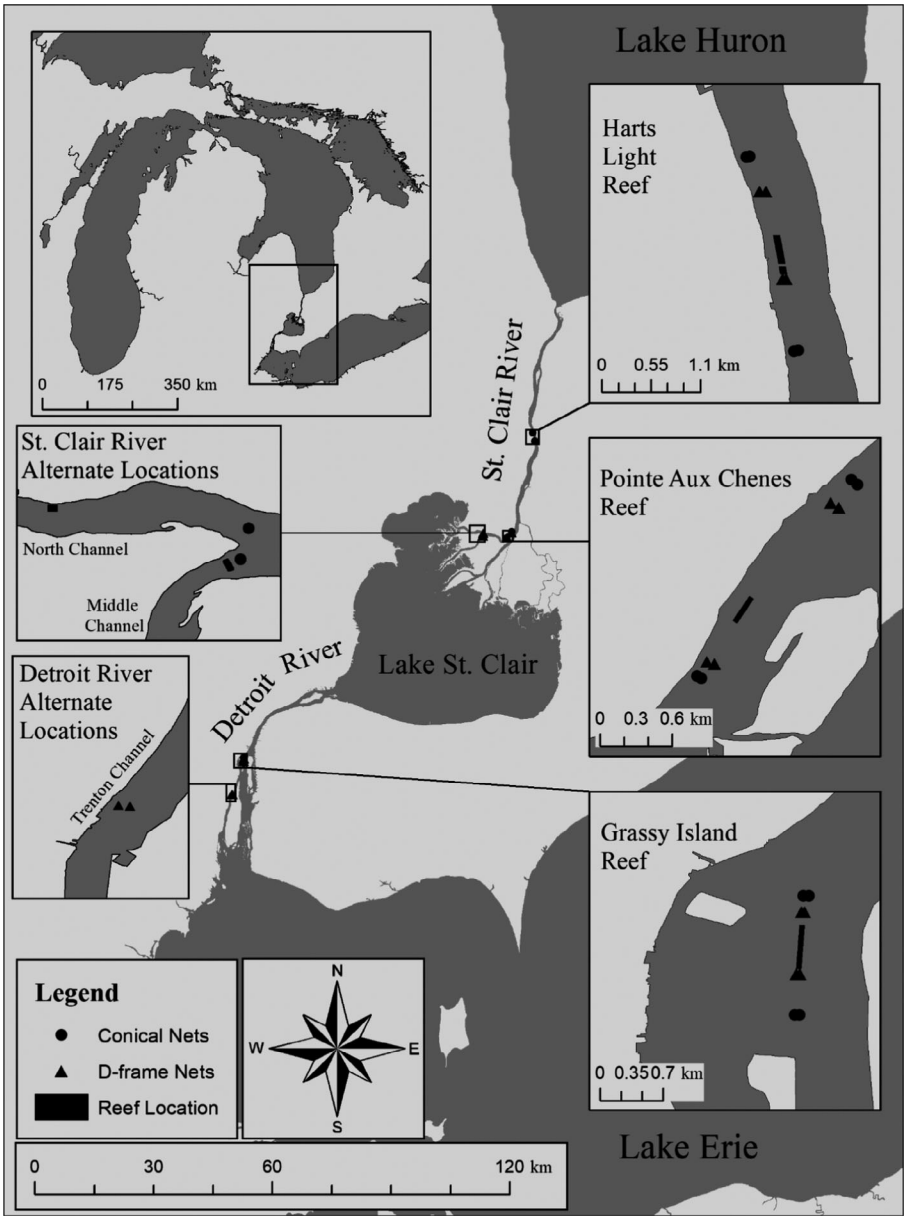

FIGURE 1. Map of the St. Clair–Detroit River system. The study site location relative to Michigan is shown in the upper left inset. The locations of each reef sampled are shown on the insets along the right of the figure, and alternate spawning locations are shown on the insets along the left.

collection are provided in the Supplement available in the online version of this article.

Genetic analyses.— Extraction of DNA followed manufacturer's protocol using the QIAGEN DNeasy Blood and Tissue Kit (QIAGEN). A NanoDrop ND-100 spectrophotometer (Nanodrop Technologies, Wilmington, Delaware) was used to quantify DNA. Prior to amplification using polymerase chain reaction (PCR), samples were diluted to 20 ng/μL of DNA suspended in sterile water.

The power to assign sibship and subsequently infer the likely number of parents contributing offspring is positively correlated with the number and allelic variation of loci analyzed (Ryman et al. 2006; Wang and Scribner 2014). The use of additional loci has the potential to provide marked increases in the accuracy of pedigree assignment. Although Lake Sturgeon are polyploid (Fontana et al. 2004), 13 disomic loci were available for pedigree analysis. In this study collected individuals were genotyped by amplifying DNA for 18 microsatellite loci, including the 13 disomic markers *LS-68* (May et al. 1997), *Afu68b* (McQuown et al. 2002), *Spl120* (McQuown et al. 2000), *Aox27* (King et al. 2001), *AfuG9*, *AfuG56*, *AfuG63*, *Afu*

G74, *AfuG112*, *AfuG160*, *AfuG195*, *AfuG204* (Welsh et al. 2003), *Atr113* (Rodzen and May 2002) and 5 polysomic markers *Atr100*, *Atr114*, *Atr117* (Rodzen and May 2002), *AciG35*, and *AciG110* (Börk et al. 2008) that have been adapted from previous use by our group in similar pedigree reconstruction analyses with White Sturgeon *Acipenser transmontanus* (Jay et al. 2014).

The PCRs were conducted in 25-μL reactions with 5 μL of 20 ng/μL DNA suspended in sterile water. Reactions were performed using 1× PCR Buffer (0.1 M Tris-HCl, 0.1 M $MgCl_2$, 0.1 M KCl, 1.0% gelatin, 1.0% NP-40, 1.0% Triton-X), with additional $MgCl_2$ as called for based on optimizations (0.5 mM for *AciG35* and *Atr113*; 1 mM for *AfuG68*, *AfuG68b*, *AfuG204*, *Atr100*, *AciG110*, *AfuG112*, *Atr114*; 1.1 mM for *AfuG56* and *AfuG195*; and 1.5 mM for *AfuG9* and *AfuG63*), 0.2 mM of each dNTP (0.216 mM for *AfuG56* and *AfuG195*), 0.4 mM each forward and reverse primer, and 5 units/μL *Taq* polymerase (0.75 units for *AfuG204*; 0.9 units for *AfuG160* and *AfuG74*; 1 unit for *AfuG68*, *AfuG120*, *AfuG56*, *AfuG195*, *Atr100*, *AciG110*, *Atr117*, *Atr114*, *AciG35*, *AfuG112*; 1.25 units for *AfuG68b* and *Aox27*; and 1.5 units for *AfuG9* and *AfuG63*). The PCR products were then poolplexed and diluted to standardized concentrations optimal for analysis at the Michigan State University Research Technology Support Facility (https://rtsf.natsci.msu.edu/) on an ABI 3730xl DNA analyzer. Allele sizes were determined using size standards obtained from MapMarker and BioVentures as well as the incorporation of three samples of Lake Sturgeon DNA with known genotypes with the analysis of each locus. Results were visualized using Gene Marker (Softgenetics, State College, Pennsylvania). All genotypes were independently determined by two experienced lab personnel to ensure consistency. Loci that were not identically scored were reanalyzed or were eliminated from further analysis by recording them as missing data. Individuals that failed to amplify at >50% of dominant presence/absence loci were excluded from further analysis. Approximately 10% of individuals were regenotyped for all loci as a further means of quality control. Mean allelic error rates were calculated at 0.5% and 1.8% in 2015 and 2016, respectively, by dividing the number of allelic differences between the original genotypes and the 10% error-check genotypes by the total number of alleles.

Current likelihood methods for pedigree reconstruction are developed for codominant or dominant markers in diploid species but not for nondisomic markers in polyploid species. In the latter case, a genotype can be ambiguous in terms of the number of copies of each unique allele it contains. To apply current methods for relatedness and pedigree inferences in polyploid species such as Lake Sturgeon that are octoploid (Blacklidge and Bidwell 1993), genotypes of nondisomic markers can be transformed to pseudo-diploid-dominant genotypes, as demonstrated by Rodzen et al. (2004) and Wang and Scribner (2014). Here, allele scores (base pairs) at a codominant locus were converted to dominant loci, where presence of the band is indexed by 1 and absence indexed by 2. For example, a phenotype that has bands at the 3rd, 5th, and 8th allele at a polysomic locus with 12 alleles would be converted to an array of phenotypes (2, 2, 1, 2, 1, 2, 2, 1, 2, 2, 2, 2) at 12 dominant marker "loci." A total of 205 alleles were analyzed across 18 microsatellite loci (13 disomic loci with 124 alleles and 5 polysomic loci with 81 alleles) in 2015 and 2016 for Lake Sturgeon in the St. Clair and Detroit rivers, resulting in the creation of a vector with a length of 205 phenotypes for each of 725 individuals. Resulting vectors were populated using a presence or absence score of 1 or 2, respectively, and 0 represented missing data. Vectors for individual genotypes were verified by first generating the series of vectors for each locus for each individual through the use of an automated function created in R (R Core Team 2017) and then hand coding the same vectors for each locus for a subset of those same individuals to ensure genotypes were generated accurately.

Simulations to assess familial assignment accuracy.— Simulations were used to assess the power of the full-likelihood method implemented in the program COLONY (Jones and Wang 2010) for reconstructing pedigrees and in estimating parameters of interest (familial sibling relationships, N_s) from the reconstructed pedigrees, given the empirical number of alleles observed ($n = 205$). A modified simulation program (available in GitHub repository, see below) previously implemented by Wang and Scribner (2014) was used in conjunction with custom R scripts to generate sets of simulated offspring with known polysomic genotypes and sibship relationships, using the 18 loci with the number of alleles observed per locus among SCDRS offspring collected. When offspring are created via COLONY's simulation module (including that of the polysomic model in the GitHub repository), all simulated offspring IDs contain information about the identities of the simulated parents that produced them. However, to assess the power of assignment in COLONY analyses like our empirical data (i.e., no available parent genotypes in the empirical data), the simulated parents were assumed unknown and unsampled (i.e., simulated parents were excluded from simulation analyses). Rather, the simulated offspring genotypes were converted to pseudo-dominant presence or absence data and evaluated in COLONY to assess the program's ability to accurately infer full-sibling, half-sibling, and unrelated dyads, and to estimate N_s without the help of any parental information. From the Best Configuration (pedigree) reconstructed by COLONY from simulated offspring genotypes, we compared the inferred and actual parents of each dyad to identify whether their genealogical

relationship (full-sibs, half-sibs, nonsibs) were correctly inferred or not.

Given that we could not estimate allele frequencies from the ambiguous polysomic genotypes of larvae sampled, we simulated allele frequencies assuming two widely different allele frequency distributions (uniform and triangular). With a triangular distribution, each locus had frequencies 1, 2, 3, 4,… for alleles 1, 2, 3, 4,…, which were then normalized such that allele frequencies summed to 1. Except when loci have different numbers of alleles, loci have the same allele frequencies. With a uniform allele frequency distribution, the frequency of an allele was a random number drawn from a uniform distribution in the range (0, 1). Again, the allele frequencies at a locus were normalized such that they summed to 1. In the uniform distribution, allele frequencies were different between loci, even when they had the same number of alleles. The idea of employing the two different allele frequency distributions in simulations was to show that, given the polymorphisms (number of alleles per locus, number of loci) and individual sampling intensity, and the possible population structure of our empirical data, sibship can be reliably inferred independent of the allele frequencies which are unknown (and difficult to calculate) from the empirical polysomic genotype data. Simply, if COLONY inferred similar pedigree results from analyses run using the two different distributions, given the same set of 18 markers and other sampling and population parameters, then the results were deemed reliable and insensitive to allele frequency distributions. Otherwise, the results are dependent on allele frequencies and our analysis of empirical data would need to be interpreted with caution.

To further assess whether the marker information is sufficient or not for pedigree reconstruction and thus whether the results were reliable or not, the simulated (and empirical) data were analyzed with and without using a maternal and paternal sibship prior of average sibship size = 1. Like other likelihood or Bayesian methods, priors tend to have a decreasingly small effect on the final result with an increasing amount of marker information available to a pedigree analysis (Wang and Santure 2009). The inference errors of both types, false siblings and false nonsiblings, were compared for the same simulated data analyzed by COLONY with and without using the sibship prior.

A central component to the COLONY simulations was a breeding matrix, constructed with known females (columns j) and males (rows i), and the known number of offspring from the mating between female j and male i. The matrix size was determined by the number of females (N_{females}) and males (N_{males}) that successfully reproduced in each simulation. We evaluated a range of potential values of N_{females} and N_{males} expected to use constructed spawning reefs. Thus, we sampled the number of parents from a uniform distribution ranging from 10 to 500 based on estimates from empirical data (Hunter 2018). Given previous studies (Craig et al. 2009; Chiotti et al. 2016), there is evidence that the operational spawning sex ratio in the St. Clair River is approximately three to one (males to female). Thus, we divided the number of parents (N_{parents}) into males and females to represent the above male-biased sex ratio. Once the size of the breeding matrix was established, for j in 1 to N_{females} columns in the matrix, the number of males (N_{mates}) that mated with the jth female was randomly drawn from a Poisson distribution (mean and variance in reproductive success, $\lambda = 3$) given that there is genetic evidence of polygyny and polyandry in Lake Sturgeon (Duong et al. 2013; Dammerman et al. 2019). Unique males (N_{mates}) were then randomly identified in the breeding matrix and were considered the mates for the jth female. The number of eggs fertilized ($N_{\text{offspring}}$) for the jth female was sampled from a uniform distribution in the range of 50,000 to 700,000. The $N_{\text{offspring}}$ were then nonrandomly distributed among mated males by allocating the largest proportion of offspring to one male, and all following males were allocated fewer offspring. For example, in the case of a female that produced 100,000 offspring and mated with three males, the first male produced 50,000 offspring with the female, and the second and third males produced 33,333 and 16,667 offspring, respectively, with the female. The random process of determining mates per female, described above, occasionally resulted in a limited number of males not mating with any females. Thus, the above process was repeated with any unmated males, except that N_{mates} were females rather than males to maintain the desired sex ratio. After the breeding matrix was created, 50, 125, or 750 offspring were randomly sampled from the total number of offspring produced by the breeding matrix (0.95 to 48.5 million fertilized eggs, depending on the number of females in the breeding matrix). These offspring sampling sizes correspond to the lower, mean, and upper bounds of sample sizes collected and used for COLONY analysis as part of this study (i.e., the minimum of 60 offspring from Pointe Aux Chenes in 2016, mean of 128 offspring across all individual reef locations in all years, and a maximum of 725 offspring from all locations combined for 2015–2016). Subsampling the entire matrix was an important aspect of the simulation because it reflects the incomplete sampling of full- and half-sibling families in the wild for highly fecund species like Lake Sturgeon. The expectation was that, if the breeding matrix is held constant, the number of full- and half-sibling dyads would decline as fewer offspring are sampled, limiting the sibship information available to inform the full-maximum likelihood method and making it a more challenging scenario to infer relationships correctly. For samples of 50, 125, and 750 offspring, 100 independent simulated datasets were generated

for each of two allele frequency distributions (uniform and triangular), and each data set was analyzed by the program COLONY with and without a sibship prior (totaling 1,200 simulated data sets and analyses). All COLONY analyses were conducted using the Linux installation (colony2s-ifort.out) in parallel on the Michigan State University High Performance Computing Cluster provided by the Institute for Cyber Enabled Research. Each independent COLONY run was initiated using a different random number seed. The COLONY parameters outlined above and used in simulations matched those for COLONY runs with genotypes determined from larvae collected in the field (see below).

The postprocessing of all Best Configuration pedigrees for simulated data sets has been described in Sard et al. (in press), and its associated Github repository provides a diagram depicting how preprocessing occurred: https://github.com/ScribnerLab/SeaLampreyRapture/tree/master/analysis/pedigrees/Simulations. In short, we enumerated the number of parents inferred by counting the number of unique parents reported in the Best Configuration pedigree output files, following the completion of each simulation. The Best Configuration pedigree output files are the final inferred pedigree outputs (containing the following: genotyped OffspringID, putative FatherID, putative MotherID) inferred by COLONY. Given that unique parents were identified by COLONY, we were able to identify inferred full-sibling, half-sibling, and nonsibling (unrelated) dyads from the Best Configuration pedigree as well. Known (simulated) parent ID's for each offspring were compared to inferred parent ID's assigned to each sampled (simulated) offspring, allowing assessment of the power to estimate N_s and correctly infer dyadic relationships.

We created an assignment matrix for each pedigree that enumerated all nine possible dyadic relationships. We determined if known (simulated) full siblings, half siblings, and nonsiblings were correctly assigned (diagonal boxes in the matrix), and if not, how they were misassigned (off-diagonal boxes in the matrix). The assignment matrix identified potential bias in each of 100 replicate simulated pedigrees. Importantly, we also compared the estimated number of unique parents to the known number that was simulated (N_s) for each pedigree. The ratio of estimated N_s to known (simulated) N_s provides a measure of expected bias across the parameter space evaluated.

Pedigree analyses of empirical data.— Individual multilocus genotypes were analyzed using program COLONY (Wang 2004) to assign full-sibling, half-sibling, and unrelated groups and to infer the most likely number of parents that contributed sampled offspring (N_s) and the effective number of parents (N_b) contributing offspring in egg and larval samples. In COLONY's maximum-likelihood approach, the relationships among all individuals and all possible relationships (i.e., full siblings, half siblings, and nonsiblings) between each pair of individuals are considered such that a consistent pedigree that maximizes the probability of the entire observed genotype data is found. Unlike the cut-off (threshold) probability approach that tends to yield much more type II errors (false negative, false nonsibs) than type I errors (false positive, false sibs) when power is not high, the full-likelihood method in COLONY balances type I and II errors and thus the total number of both types of errors tends to be minimized in the likelihood inference framework. Estimates of N_b and confidence intervals were calculated by COLONY using the "sibship assignment method" described in Wang (2009) and are based on the frequencies of estimated full- and half-sibling dyads. Estimates of N_s were calculated by enumerating the unique putative parents that were inferred to have contributed offspring at sampled sites. Sex of putative parents in this analysis was unknown, but the number of larvae assigned to each parent could be summed. Individual reproductive success was calculated as the number of eggs or larvae contributed by a unique putative individual.

Measures were taken to minimize the possibility of errors due to amplification, scoring, typing, and data entry. Program COLONY accounts for data errors described as class 1 error: allelic dropout during PCR amplification, and class 2 error: mutation, genotyping, contamination, and data entry (Wang 2004). Mutation rates (i.e., germ-line mutations in offspring), also accounted for in COLONY, may be as high as 1.4×10^{2} for microsatellite loci (Talbot et al. 1995; Wang 2004). However, mutation rates for Lake Sturgeon are not described in the literature (McDougall et al. 2017) so our analyses assumed 1% for class 1 errors and 1% for class 2 errors. These error rates were used to account for data errors and the divergence from Mendelian inheritance created by data conversion (from polysomic genotypes to pseudo-diploid-dominant genotypes; Wang and Scribner 2014). All COLONY runs allowed for multiple mates for both sexes (polyandry for females and polygyny for males). Allele frequencies were considered unknown and were not updated. Simulated and empirical analyses were conducted with and without use of a sibship prior, and results were reported for both. Pedigrees were reconstructed using the full-likelihood approach during a "long" run with "high" precision. The COLONY runs were conducted using genotyped offspring from all years and all reefs ($n = 725$), each year and all reefs ($n = 307$ for 2015 and $n = 418$ for 2016), and each year and each individual reef ($n = 60$–200). Best Configuration pedigrees from COLONY runs that used the full data set ($n = 725$) were then subset by collection location and collection gear type to infer the number of full and half siblings and the number of unique parents that contributed offspring between locations or patterns of larval dispersal. Finally, no candidate

parent (female or male) genotypes were included in the analysis.

RESULTS

Power of Assignment

We conducted simulations to evaluate how accurately the pedigrees were reconstructed and how accurately the key parameters associated with Lake Sturgeon life history were inferred from the reconstructed pedigrees. We found, regardless of sample sizes, assumed allele frequency distributions (uniform or triangular), numbers of parents that produced the offspring sampled, and the inclusion or exclusion of a sibship prior, that the major conclusions based on COLONY family reconstructions remain unchanged (Figures S1–S4 available in the Supplement in the online version of this article). Thus, for brevity, we report statistics (and associated figures) for simulations that assumed a triangular allele frequency distribution that did not use a sibship prior. Across the offspring sample sizes evaluated that reflect the empirical range in this study (50, 125, 750), on average, 98–100% of unrelated dyads were classified as unrelated (Figures 2–4). That is, few unrelated dyads were falsely inferred as related. In addition, regardless of sample size, most known full siblings and half siblings were correctly inferred; however, the average proportion of dyads correctly inferred increased with larval sample size ($N_{offspring} = 50$: full sibling = 75% correct assignment, half sibling = 71% correct assignment; $N_{offspring} = 125$: full sibling = 90% correct assignment, half sibling = 79% correct assignment; $N_{offspring} = 750$: full sibling = 99% correct assignment, half sibling = 88% correct assignment; Figures 2–4). The vast majority of pedigree assignment errors, reported here as the average proportion of dyads, were associated with known full-sibling dyads assigned as half siblings ($N_{offspring} = 50$: 25%; $N_{offspring} = 125$: 10%; $N_{offspring} = 750$: 1%; Figures 2–4). Additional errors, reported here as the average proportion of dyads, were associated with half-sibling dyads falsely considered unrelated ($N_{offspring} = 50$: 28%; $N_{offspring} = 125$: 20%; $N_{offspring} = 750$: 12%). Importantly, the proportion of incorrectly inferred dyads decreased as the true number of parents that produced the collection of offspring increased (Figures 2–4), reflecting a decline, on average, in sibship size, and thus the power for COLONY to correctly infer sibship. Taken together, simulations demonstrate that results from dyadic relationship inferences do not substantially change conclusions about adult spawning location and larval dispersal downstream from reefs, regardless of allele frequency distributions and the inclusion or exclusion of a sibship prior. Specifically, larvae collected in downstream D-frame nets that were inferred to be related to larvae that were collected as an egg on an upstream spawning reef likely shared at least one parent (i.e., were either a full- or half-sibling), and thus reliably reflected dispersal of siblings from the upstream reef spawning location at least as far as the

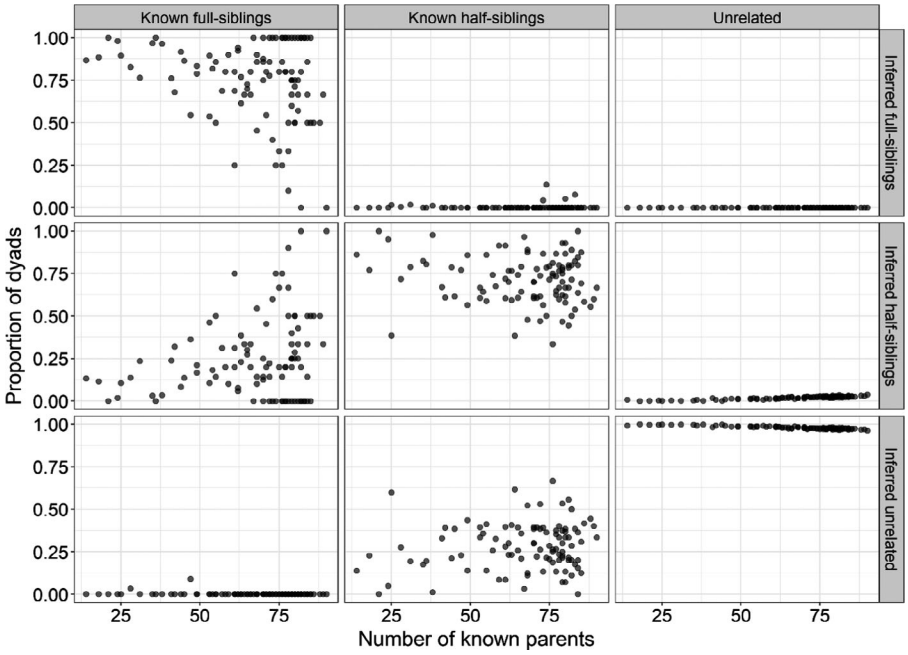

FIGURE 2. Matrix showing the proportion of true to inferred full-sibling, half-sibling, and nonsibling (unrelated) offspring dyads for the number of known parents (N_s) included in 100 simulations that sampled 50 offspring from a simulated breeding matrix assuming a variable allele frequency distribution and no sibship prior.

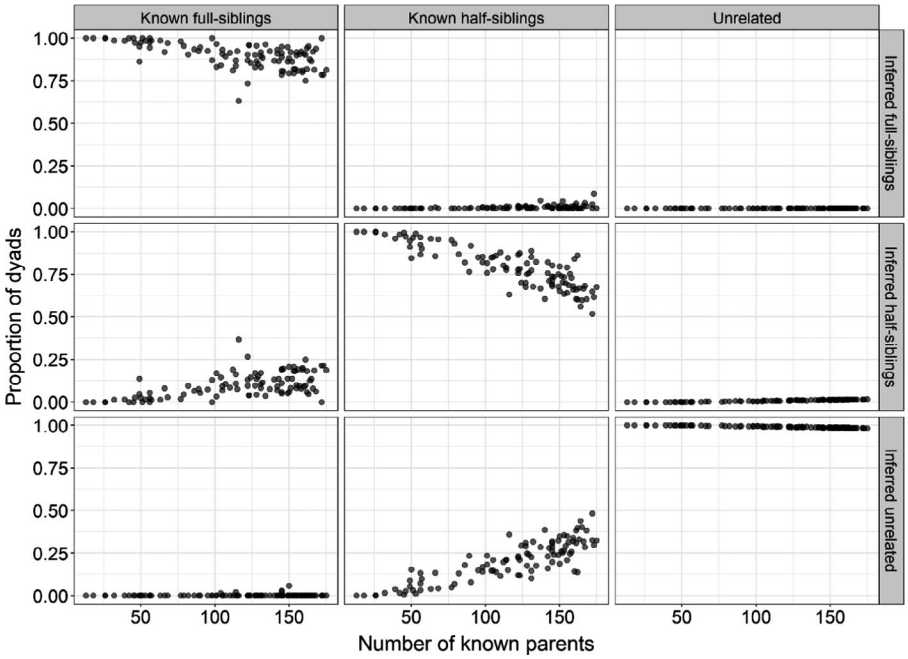

FIGURE 3. Matrix showing the proportion of true to inferred full-sibling, half-sibling, and nonsibling (unrelated) offspring dyads for the number of known parents (N_s) included in 100 simulations that sampled 125 offspring from a simulated breeding matrix assuming a variable allele frequency distribution and no sibship prior.

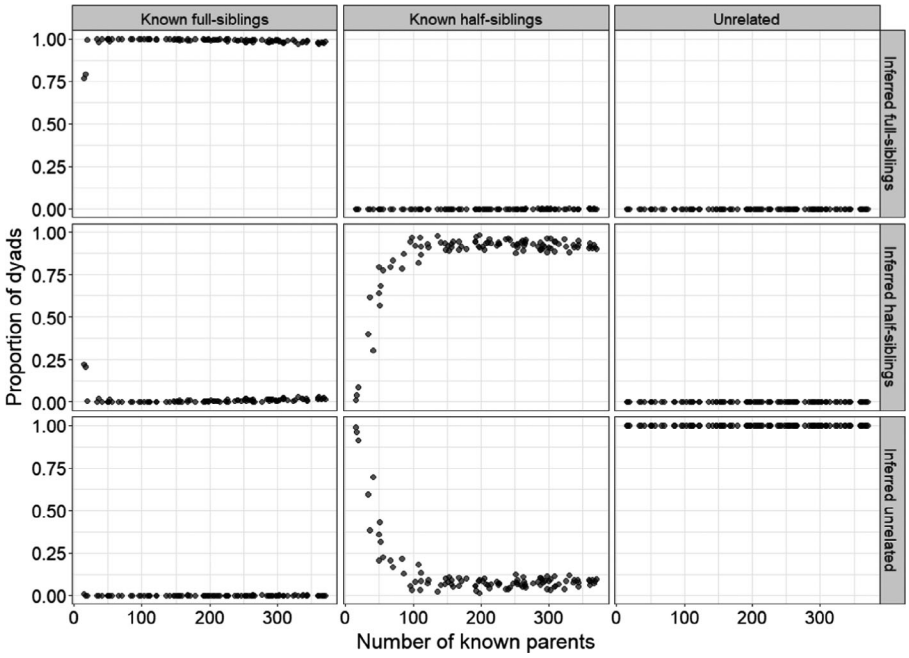

FIGURE 4. Matrix showing the proportion of true to inferred full-sibling, half-sibling, and nonsibling (unrelated) offspring dyads for the number of known parents (N_s) included in 100 simulations that sampled 750 offspring from a simulated breeding matrix assuming a variable allele frequency distribution and no sibship prior.

downstream D-frame net collection location. In addition, the inference errors associated with dyadic relationships did not substantially bias estimation of N_s. That is, across the parameter space evaluated, COLONY inferred, on average, 16% percent fewer parents compared with the known number of parents when 50 or 125 offspring were

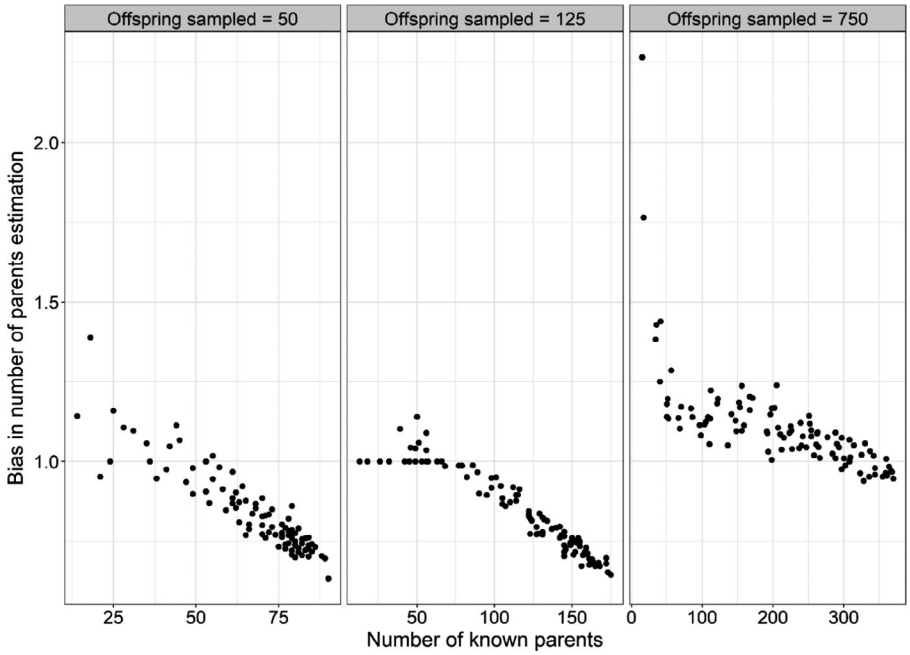

FIGURE 5. Scatter plot characterizing the relationship of between bias (the ratio of estimated N_s to known [simulated] N_s) in the estimated and number of known parents that produced simulated egg and larval collections of 50, 125, and 750 offspring assuming a variable allele frequency distribution and no sibship prior. The number of known parents in the simulations is plotted on the *x*-axis, and the bias in the number of estimated parents is plotted on the *y*-axis.

sampled (Figure 5). When 750 offspring were sampled, COLONY inferred, on average, 12% more parents as compared with the known number of parents. Given that the true number of parents that produced the collection of offspring is dependent on sample size and will never be known on reefs evaluated as part of this study, we consider the empirical results as approximations of the minimum number of adults using the reefs during the spawning on the reefs. Thus, based on simulated data sets, we conclude that genotyping the 18 loci provided sufficient power for the accurate pedigree reconstruction of eggs or larvae, allowing for the reliable estimation of N_s, dispersal, interannual reproduction, and spawning at multiple reefs within a year.

Quantifying Spawning Success

Results of simulations and empirical analysis, for brevity, are reported and discussed hereafter for COLONY runs that did not include a sibship size prior. However, results of replicate COLONY analyses with and without sibship priors are reported in referenced tables, figures, and the Supplement. A total of 725 reared eggs and larvae were genotyped and used for analysis, 307 in 2015 and 418 in 2016 (Table 1). Concordance between replicate COLONY analyses was high for all estimates at individual sampling locations within years (Table S1 available in the Supplement in the online version of this article). Estimates of the number of adults producing sampled offspring genotyped (N_s) averaged across COLONY runs were 151, 208, and 330 in 2015, 2016, and 2015–2016 combined, respectively (Table 1). Estimates of the effective number of breeding adults (N_b) over all locations was 158 (95% CI = 127–199) and 198 (95% CI = 163–245) and over all samples was 317 (95% CI = 271–372) in 2015, 2016, and 2015–2016 combined, respectively (Table 1). Estimates of N_s and N_b varied across years and among individual reef sites and were positively correlated with sample size (Table 1). As more larvae were genotyped, more spawning adults were detected and estimates of effective breeding size increased, indicating that sample sizes do not likely capture the total number of spawning Lake Sturgeon contributing offspring. Estimated mean and variance in adult reproductive success (number of offspring produced) also varied with sample size. Adult reproductive success was estimated at 3.54 (mean) and 3.62 (variance) larvae per spawning adult, across all years and locations (Table 1).

Adult Spawning Site Selection

Identification of adults using specific spawning reefs was based on family reconstruction analysis using the entire data set ($n = 725$ offspring) and resulting sibship assignments between larvae reared from eggs collected on egg mats between spawning reef locations. Genetic family

TABLE 1. Mean estimates of the number of spawning adults (N_s), the effective number of breeders (N_b), including the lower and upper 95% confidence intervals, and the mean and variance in individual reproductive success (R_s) at constructed reefs. Estimates for N_s, N_b, confidence intervals, and R_s mean and variance are presented for COLONY runs without and with a sibship prior, where the number in front of the semicolon is for runs without a sibship prior and the number after the semicolon is for runs with a sibship prior. Sample size (n) is the number of larvae genotyped from each location included in the analysis.

Location	Year	n	N_s	N_b	Lower CI	Upper CI	Mean R_s	Variance R_s
All locations	2015	307	151; 157	158; 158	127; 128	199; 199	3.93; 4.09	3.87; 3.33
Harts Light	2015	160	97; 96	110; 97	82; 72	146; 133	3.34; 3.30	3.19; 2.01
Pointe Aux Chenes	2015	106	66; 66	70; 74	50; 52	98; 106	3.22; 3.24	2.16; 2.59
All locations	2016	418	208; 206	198; 198	163; 161	245; 243	4.06; 4.03	4.77; 4.85
Harts Light	2016	200	122; 124	132; 134	106; 103	172; 172	3.23; 3.29	2.44; 2.40
Pointe Aux Chenes	2016	60	50; 50	61; 63	41; 42	92; 95	2.40; 2.41	1.14; 1.34
Grassy Island	2016	117	63; 62	54; 53	37; 37	79; 79	3.82; 3.72	6.07; 6.20
All locations	2015–2016	725	330; 334	317; 317	271; 271	372; 372	4.35; 4.41	5.33; 5.13

reconstruction allowed estimates of the number of unique parents that contributed offspring at multiple SCDRS reefs. Additionally, detection of full- and half-sibling larvae reared from eggs collected on egg mats placed at different spawning reefs provided evidence of Lake Sturgeon spawning at multiple locations in the same year (Figure 6). The inferred number of parents contributing sampled offspring at multiple reef locations ranged from 2 to 17 parents across pairs of sampling sites in 2015 (Figure 6; Table S2). The number of parents contributing sampled offspring at multiple reefs sites ranged from 2 to 23 in the St. Clair River in 2016 (Figure 6; Table S2). For both years, the greatest number of adults using multiple sites were spawning at Harts Light reef and Pointe Aux Chenes reef in the St. Clair River main channel (Figure 6). Collections of sibling larvae reared from eggs collected on egg mats at multiple sites revealed that adults spawned at multiple locations within the same spawning year in the St. Clair River in 2015, and further evidence of adults spawning in multiple locations was documented in 2016. Family reconstruction of individuals sampled on Grassy Island reef in the Detroit River in 2016 revealed evidence of Lake Sturgeon spawning in multiple rivers in a single spawning season. Analysis documented larvae reared using eggs collected on egg mats at Grassy Island reef that were half-siblings with larvae reared from eggs collected on egg mats at Harts Light reef, Pointe Aux Chenes reef, and the Maslinka and Middle Channel reefs (Figure 6). The inferred minimum number of parents contributing sampled offspring in both rivers in 2016 was estimated at 14 spawning adults (Figure 6).

Evidence for Spawning in Consecutive Years

To determine if adults were spawning in consecutive years, COLONY was run combining all larvae from 2015 and 2016 and from all sites ($n = 725$). From the 262,450 possible dyads, COLONY analysis inferred 226 between-years half-sibling dyads concordantly between two replicate colony runs. Detection of half-sibling larvae between years provides evidence that a limited number of adults spawned in consecutive years at the same and on different reefs in the St. Clair River and between the St. Clair and Detroit rivers.

Patterns of Larval Dispersal

Larval dispersal was characterized based on collections of larvae reared from eggs collected on egg mats and larvae captured in D-frame and vertically stratified conical nets placed downstream of egg collection (spawning) sites (Figure 7). Collection of inferred full- and half-sibling larvae from upstream (spawning) sites and during dispersal downstream indicated that larvae may disperse large distances in the SCDRS (Table S3). Of particular note were the large collections of inferred siblings sampled between egg mats on Harts Light reef and nets at Pointe Aux Chenes reef and between main-channel and alternate spawning locations (Table S3). In 2016, small numbers of half-sibling dyads were detected between eggs collected on egg mats in the St. Clair River and dispersing larvae collected in drift nets at Grassy Island in the Detroit River (Figure 7).

DISCUSSION

Based on pedigree reconstruction data for Lake Sturgeon eggs and larvae collected over a large area in the SCDRS, results illustrate the demographic and genetic benefits of habitat remediation for Lake Sturgeon. Companion simulations provided compelling evidence for accurate full- and half-sibship assignments based on joint use of Mendelian and polysomic markers and reliable estimates of N_s and N_b. Empirical collections from multiple and widely dispersed spawning sites during consecutive years provided estimates of the number of spawning adults per site (N_s), the effective number of spawning adults per site (N_b), and the mean and variance in individual

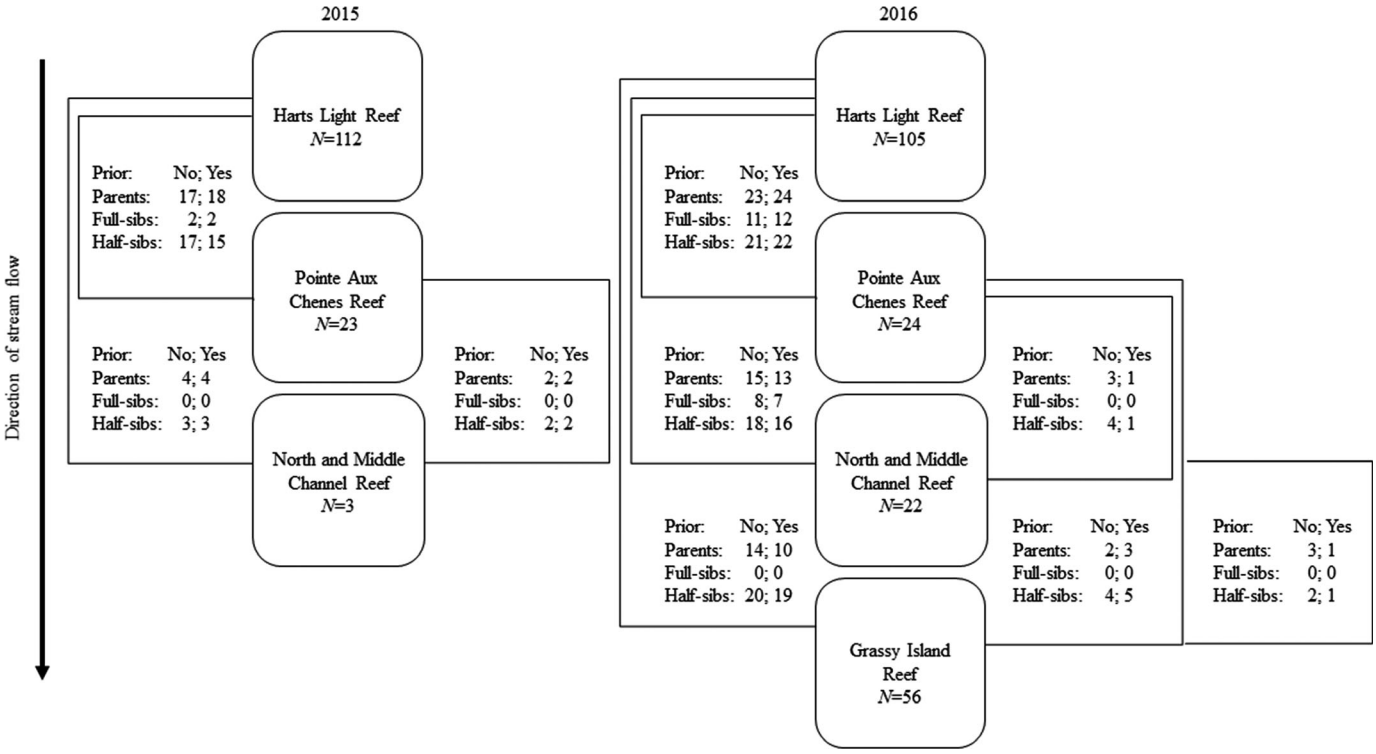

FIGURE 6. Diagram showing the number of unique parents that contributed offspring, number of full-sibling dyads, and number of half-sibling dyads between constructed reef sites in 2015 and 2016. A sibling dyad refers to each full- or half-sibling relationship between two genotyped offspring. To ensure that the spawning location of offspring was known, sibling and parent relationships for larvae reared from eggs collected on egg mats were summed by sub setting the pedigrees generated from the full data set by year and collection location for comparison between sites and years. Results from COLONY runs without and with sibship size priors are reported, where the number in front of the semicolon is for runs without a sibship prior and the number after the semicolon is for runs with a sibship prior. Brackets indicate the sibship and parent comparisons being made between locations (e.g., for the 2015 diagram, the bracket from Pointe Aux Chenes reef to Harts Light reef indicates that from 23 larvae reared from eggs collected on egg mats at Pointe Aux Chenes reef there were 2; 2 full-sibling dyads and 17; 15 half-sibling dyads contributed by 17; 18 unique parents when compared to 112 larvae reared from eggs collected on egg mats at Harts Light reef). Arrows on the left indicate the direction of stream flow relative to collection location order.

reproductive success and identified whether (and where) adults spawned at multiple locations in the same or consecutive years, as well as characterized patterns in larval dispersal. Results of this study provide important information for managers to consider when assessing current habitat remediation efforts for Lake Sturgeon in the SCDRS and can inform future habitat remediation efforts for fishes in large river systems.

Number of Spawning Adults Contributing Offspring

Lake Sturgeon are highly fecund (Bruch and Binkowski 2002), and as such, an entire sample collected from an egg mat or net could contain eggs or larvae from a single female. Accordingly, enumeration of total egg or larval counts at spawning reefs are not a viable estimator of adult spawner number or reproductive contributions. Additionally, if construction of spawning habitat disperses the finite number of spawning adults in a river system, there is potential for depensatory effects due to mate limitation that could lead to accrual of levels of relatedness among offspring that in turn can have negative effects on long-term population genetic diversity. However, genetic family reconstruction techniques used here provided compelling evidence that many spawning adult Lake Sturgeon were contributing to the sampled offspring at each reef site sampled.

We caution that the number of spawners detected is positively correlated with sample size, and it is highly unlikely that all spawning adults were detected in this study. Thus, cumulative and site-specific estimates of N_s and N_b (Table 1) are likely downwardly biased as they represent the number of spawners contributing the offspring that were sampled at a spawning reef and not the total number of spawners contributing offspring at a spawning reef. The N_s estimates (Table 1) demonstrate that offspring were recruited by >50 spawning adult Lake Sturgeon at all locations in both years. These results reduce concerns that constructed spawning habitat may cause depensatory effects by reducing spawner abundance at individual spawning locations.

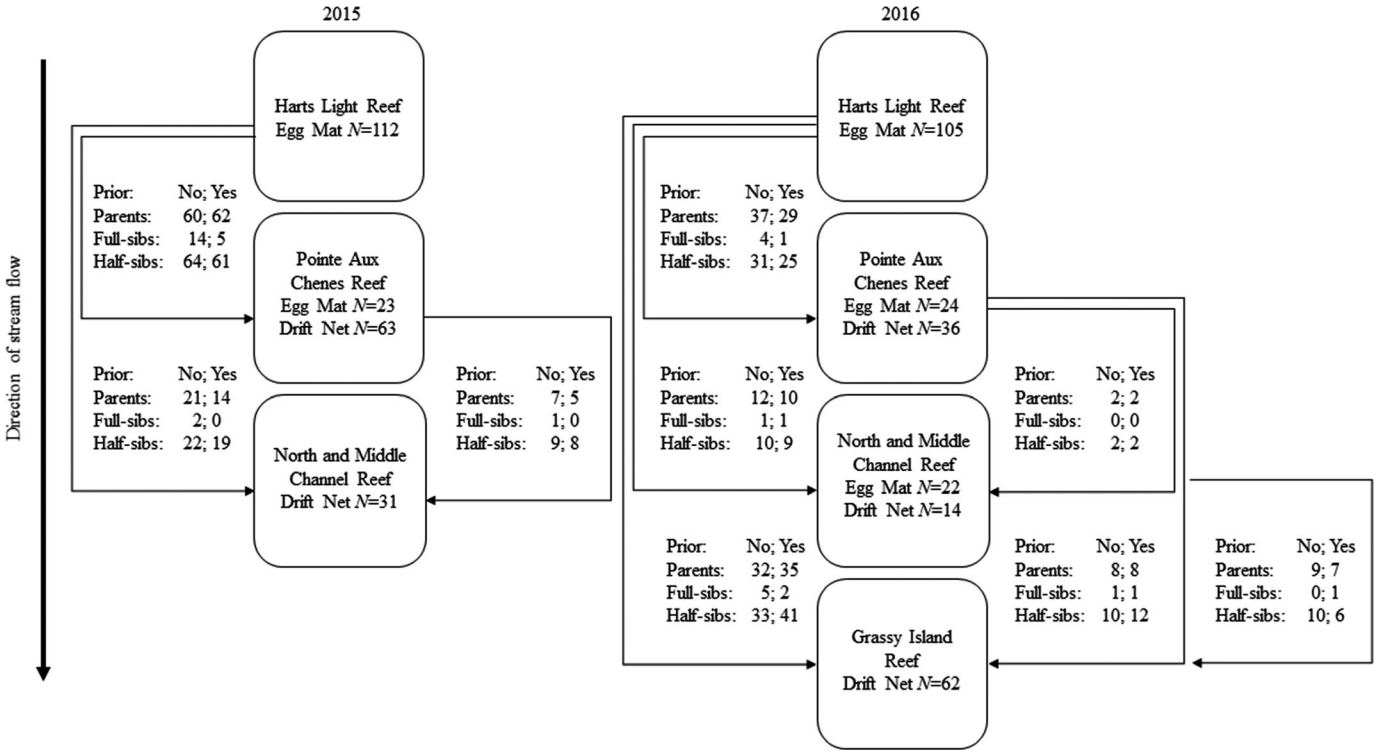

FIGURE 7. Diagram showing the number of unique parents that contributed offspring, number of full-sibling dyads, and number of half-sibling dyads between larvae reared from eggs collected on egg mats at an upstream site and larvae collected in drift nets (D-frame nets and vertically stratified conical nets) at a downstream site. Results from COLONY runs without and with sibship size priors for the full data set are reported (the number in front of the semicolon is for runs without a sibship prior and the number after the semicolon is for runs with a sibship prior). Sibling and parent relationships were determined by sub setting and comparing the pedigrees generated from the full data set by year and collection location. Full- and half-sibling relationships between offspring spawned at an upstream site and offspring dispersing to a downstream location are used to infer patterns of larval dispersal in the system. Arrows indicate the direction of stream flow relative to collection location order.

Effective Number of Breeders

Prior to this study little was known about the effective number of breeding adults in the system as a source of reference collectively across the SCDRS or on a site-by-site basis before the constructed spawning reefs were established. Here, pedigree reconstruction estimated the effective numbers of breeding adults (N_b) based on frequencies of full- and half-sibling dyads (Wang 2009) that used constructed spawning habitat in the SCDRS. Estimates of N_b ranged from 54 to 132 across all sampled locations (Table 1) and at each site was relatively consistent in 2015 and 2016. Data indicate that adult Lake Sturgeon use new spawning areas immediately and in fairly large numbers (i.e., at least 50–100 adults per reef per year) given the limitations of sample size and duration. Other Lake Sturgeon research indicates that N_b decreases as a function of increasing sex ratio skew (Duong et al. 2013; Dammerman et al. 2019). Results of this study estimated that mean and variance in reproductive success of adults that contributed offspring sampled were nearly Poison distributed, explaining comparable N_s and N_b estimates (Table 1). Marranca et al. (2015) also estimated N_b in the SCDRS but on different spawning reefs and years. Sample sizes in Marranca et al. (2015) were smaller and the authors used a different (linkage disequilibrium) estimator, but their N_b estimates were similar to estimates in this study.

Adults Spawned in Multiple Locations

Detection of full-sibling and half-sibling larvae reared from eggs collected at multiple upstream reef sites in the St. Clair River provides evidence of flexibility in adult spawning behavior, whereby adult Lake Sturgeon are participating in multiple spawning events at multiple locations in the same year (Table S2). Spawning at multiple locations may contribute to a resilient population recruitment portfolio (Schindler et al. 2010; Dufour et al. 2015) for Lake Sturgeon in the SCDRS by buffering the effects of site-specific mortality in a single year, including a reduction in the potential losses of genetic diversity. Successful reproduction by adults at multiple reefs contributed to relatively low estimated variance in reproductive success. Determining levels of connectivity and adult use of different reefs could help inform future placement of new spawning reefs.

Evidence of adults spawning at multiple locations within years provides compelling evidence for spawning habitat connectivity in the St. Clair River in 2015–2016 and limited connectivity between the St. Clair and Detroit rivers in 2016 (Figure 6; Table S2). When designing habitat remediation projects with limited resources, managers frequently must compromise between size, number, location, and juxtaposition of restoration sites to be constructed. Diamond (1975) outlined principles for management of habitat patches, including prioritization of large protected areas and connectivity that may be applied to spawning habitat restoration in river systems. Large sites may attract more spawning adults and reduce concerns regarding depensatory effects. However, a single large site may reduce population resiliency by negating the potential benefits of a diverse spawning portfolio if site conditions vary spatially within a season (i.e., efficacy of single large or several small areas for conservation). Additionally, hydrology and anthropogenic factors, such as commercial shipping channels, may not allow for large-scale reef designs. Connectivity between locations as indicated in this study based on evidence that adults produced viable offspring at multiple locations (Figure 6; Table S2) suggests that construction of many, smaller spawning habitats may be desirable.

Patterns of Larval Dispersal

Larval dispersal behavior, timing of larval dispersal, and time to initiation of the benthic foraging life stage varies among sturgeon species and systems (Gisbert and Ruban 2003; Braaten et al. 2008; Duong et al. 2011). In this study, related dyadic offspring pairs were detected at large distances. Collection of larvae in drift nets that were full or half siblings with larvae reared from eggs collected on egg mats at upstream locations indicates that larvae in the SCDRS may have moved distances > 40 km in the St. Clair River in 2015 and 2016 (Figure 7). Additionally, collection of half-sibling larvae between larvae reared from eggs collected on egg mats at Harts Light reef and larvae caught in downstream D-frame nets at Grassy Island reef suggest that larvae may disperse > 120 km between the St. Clair and Detroit rivers. While it may seem unlikely that larvae are dispersing > 120 km from Harts Light reef in the St. Clair River to Grassy Island reef in the Detroit River, there is evidence in the literature for transport of Deepwater Sculpin *Myoxocephalus thompsonii* and Burbot *Lota lota* (Roseman et al. 1998; Lantry et al. 2007; McCullough et al. 2015) between the two rivers.

However, this study also provides evidence of gene flow between the two rivers (Figures 6–7; Tables S2–S3) by adults spawning in multiple locations. Previous genetic analysis of Lake Sturgeon in the St. Clair and Detroit rivers indicated low genetic differentiation between rivers (e.g., nonsignificant F_{ST} values and contingency test results), citing possible successful spawning by adults moving between locations (Welsh et al. 2008). Genetic family reconstruction results from this study are also consistent with findings from telemetry studies that indicate some adult Lake Sturgeon move between the St. Clair and Detroit rivers (Caswell et al. 2004; Kessel et al. 2018). Additionally, while full- and half-sibling relationships between larvae reared from eggs on upstream egg mats and larvae captured in downstream D-frame nets suggests downstream larval dispersal, this pattern is confounded by adults observed spawning in multiple locations (Figures 6–7; Tables S2–S3). Therefore, full- and half-sibling relationships may be representative of some combination of larval dispersal and adults spawning in multiple locations. Results are suggestive of a pattern that warrants further investigation rather than compelling evidence for trends in larval dispersal. Future research to quantify movements of spawning Lake Sturgeon, dispersing offspring between rivers, and levels of gene flow per generation across reefs and the entirety of the 160-km system would be useful in assessing the effects of increasing spawning locations in both river populations.

Power Assessment

Levels of uncertainty have been evaluated via simulations in this study (Figures 2–5, S1–S4). The effects of this uncertainty on inferences from genetic family reconstruction techniques employed here deserve mention. As stated above, in COLONY's maximum likelihood approach, all possible relationships among all individuals sampled were considered such that a consistent pedigree that maximizes the probability of the observed genotype data is found. This is the maximum likelihood (point) estimate of the entire pedigree (i.e., the COLONY Best Configuration pedigree). Thorstensen et al. (2019) sought to evaluate the possibility of using cutoffs as part of a postprocessing procedure to minimize type II errors and ensure only correctly inferred dyads are used in subsequent analyses. However, such cutoffs should be performed with caution as the maximum likelihood approach used in COLONY seeks to balance type I and type II errors and the introduction of sibling cutoffs can introduce additional errors in the analysis. When marker information is not very high (due to limited marker number or polymorphisms) or/and the genetic structure is weak (nonsibs are much more frequent than sibs), the power of a sibship analysis can be low. In such a case, the cutoff approach could lead to many sibs being falsely inferred as nonsibs, inflating type II errors and thus the total number of errors of both types. This is not a problem if later analysis relies on inferred sibs (which are reliable) only and the inferred nonsibs (which are not reliable because of inflated type II errors) do not matter. Otherwise, however, the approach could lead to much reduced inference accuracy and

possibly incorrect conclusions in later analysis. Our analyses depend on both the inferred sibs and the inferred nonsibs, and thus we choose to use the Best Configuration pedigree to minimize the total number of errors of both types. Simulations presented here (Figures 2–5; Figures S1–S4) show that errors occurred in sibship assignments most frequently at small offspring sample sizes ($n = 50$ or 125), where error rates were as high as 25% for full siblings and 28% for half siblings. However, at the largest simulated sample size ($n = 750$) rarely were full siblings, half siblings, or unrelated individuals incorrectly assigned, and error rates were as low as 1% for full siblings and 12% for half siblings. Our simulation results demonstrate that with 18 loci (13 disomic and 5 polysomic) and a large sample size ($n = 750$) most related individuals are likely inferred as related individuals. Importantly, all conclusions regarding adult spawning location and larval dispersal discussed here were drawn from COLONY analysis that included 725 genotyped offspring, a sample size for which simulations demonstrate reliable sibship inferences can be made. However, there may be some error associated with specific types of relationships (full sibling or half sibling). The program COLONY could produce type I and II errors in the maximum likelihood pedigree; for example, a full-sibling pair might be inferred as a half siblings. Again, these two types of errors are minimized in the likelihood inference framework, and both types of errors decrease in frequency consistently with an increase in marker information (Figures 2–4, S1–S4). Not surprisingly our simulations showed that, under sampling conditions similar to our empirical data, the analysis results are accurate and robust, regardless of the allele frequency distributions used in generating the data and whether or not a sibship prior was used.

Previous research using simulated and empirical Lake Sturgeon data and presence or absence locus coding of alleles (Wang and Scribner 2014) indicated that, with the number of loci employed in this study ($n = 205$ alleles across 18 loci), familial relationships should be accurately characterized. However, when using fewer loci (i.e., 13 disomic loci) and lower genotyping error rates (i.e., 0.001), COLONY may split large full-sibling families into half-sibling groups, thus inflating N_s (Blankenship et al. 2017). With the additional marker information from the inclusion of five more polysomic loci, the N_s values were likely within 10–15% of the true number of parents that produced the collection of offspring. Simulations conducted as part of this study likewise indicate that the 18 loci used have fairly high discriminatory potential, particularly when large numbers of offspring are surveyed. Simulations conducted as part of this study (Figures 2, 3) were based on offspring sample sizes of 50, 125, and 750 individuals, which approximate minimum, mean, and maximum offspring numbers sampled from the lowest individual count at spawning reefs ($n = 60$, Point Aux Chenes reef in 2016) and overall locations and years, respectively. Nearly all reef–year samples exceeded 100 individuals (Table 1). Even when the simulated number of individuals sampled was low ($n = 50$; Figure 2), rarely were unrelated individuals misclassified as related individuals (full or half siblings).

Misclassifications, when present, occurred at low simulated sample sizes and were between full-sibling and half-sibling and half-sibling and unrelated individual assignments. Additionally, with sample sizes of 750, simulations show that full-sibling, half-sibling, and unrelated individual assignments can be made reliably. Thus, subsequent use of familial reconstruction inference, such as characterization of dispersal from areas of egg deposition to downstream larval collections or samples across years, was based on COLONY results using the full data set ($n = 725$); a scenario where simulations (Figure 4) demonstrate high familial assignment accuracy. We also note that most biological inferences (based on observed genotypic data) from inferred relationships were reproduced in replicate independent COLONY runs. We therefore believe that the familial assignments and inferences derived from them provide a useful step forward to quantify important aspects of Lake Sturgeon reproductive ecology and larval dispersal behaviors in a large river system that is not otherwise amenable to investigations of these questions using traditional sampling methodologies.

While simulations presented here provide compelling evidence that pedigrees can be reliably generated using COLONY and the 18 microsatellite loci available at this time, we acknowledge that even under the best circumstances, pedigrees were not likely generated without errors. Importantly for interpretations of results in this study, some estimates of N_s and N_b based on pedigree reconstructions at individual reef sites were generated using small sample sizes, and information should be interpreted in light of the simulations presented here. However, the major conclusions of this study, including documentation of large-distance larval dispersal and adult use of multiple spawning areas, were reconstructed with large sample sizes, for which simulations show high levels of accuracy in sibship assignment.

Moving forward, genomic techniques utilizing high throughput sequencing technologies will greatly increase the number of markers available for analysis at minimal additional cost. Utilizing genetic markers, such as single nucleotide polymorphisms, may make sufficient marker information available to substantially reduce misassignment rates for half siblings in particular. For example, Baetscher et al. (2018) describe the use of "microhaplotype" markers for pedigree analysis. This process involves genotyping multiple single nucleotide polymorphism loci that occur within a small sequencing region. The process

substantially increased the power of each locus and may allow for increased accuracy of sibship assignment.

ACKNOWLEDGMENTS

Support for this work was provided by the Great Lakes Restoration Initiative. We would like to thank the field and laboratory technicians at the U.S. Geological Survey Great Lakes Science Center and U.S. Fish and Wildlife Service for assistance with this project. All sampling and handling of fish during research were carried out according to "Guidelines for the Care and Use of Fishes in Research" by the American Fisheries Society (Use of Fishes in Research Committee 2014). Any trade, firm, or product names are used for descriptive purposes and do not imply endorsement by the U.S. Government. There is no conflict of interest declared in this article.

REFERENCES

Auer, N. A. 1996. Importance of habitat and migration to sturgeons with emphasis on Lake Sturgeon. Canadian Journal of Fisheries and Aquatic Sciences 53:152–160.

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 18:296–305.

Bennion, D. H., and B. A. Manny. 2011. Construction of shipping channels in the Detroit River: history and environmental consequences. U.S. Geological Survey Scientific Investigation Report 2011-5122.

Blacklidge, K. H., and C. A. Bidwell. 1993. Three ploidy levels indicated by genome quantification in Acipenseriformes of North America. Journal of Heredity 84:427–430.

Blankenship, S. M., G. Schumer, J. P. Van Eenennaam, and Z. J. Jackson. 2017. Estimating number of spawning White Sturgeon adults from embryo relatedness. Fisheries Management and Ecology 24:163–172.

Boase, J. C., B. A. Manny, K. A. L. Donald, G. W. Kennedy, J. S. Diana, M. V. Thomas, and J. A. Chiotti. 2014. Habitat used by juvenile Lake Sturgeon (*Acipenser fulvescens*) in the North Channel of the St. Clair River (Michigan, USA). Journal of Great Lakes Research 40:81–88.

Börk, K., A. Drauch, J. A. Israel, J. Pedroia, J. Rodzen, and B. May. 2008. Development of new microsatellite primers for Green and White sturgeon. Conservation Genetics 9:973–979.

Bouckaert, E. K., N. A. Auer, E. F. Roseman, and J. Boase. 2014. Verifying success of artificial spawning reefs in the St. Clair–Detroit River system for Lake Sturgeon (*Acipenser fulvescens* Raifinesque, 1817). Journal of Applied Ichthyology 30:1391–1401.

Braaten, P. J., D. B. Fuller, L. D. Holte, R. D. Lott, and W. Viste. 2008. Drift dynamics of larval Pallid Sturgeon and Shovelnose Sturgeon in a natural side channel of the upper Missouri River, Montana. North American Journal of Fisheries Management 28:808–826.

Bruch, R. M., and F. P. Binkowski. 2002. Spawning behavior of Lake Sturgeon (*Acipenser fulvescens*). Journal of Applied Ichthyology 18:570–579.

Caswell, N. M., D. L. Peterson, B. A. Manny, and G. W. Kennedy. 2004. Spawning by Lake Sturgeon (*Acipenser fulvescens*) in the Detroit River. Journal of Applied Ichthyology 20:1–6.

Chiotti, J. A., J. C. Boase, D. W. Hondorp, and A. S. Briggs. 2016. Assigning sex and reproductive stage to adult Lake Sturgeon using ultrasonography and common morphological measurements. North American Journal of Fisheries Management 36:21–29.

Craig, J. M., D. M. Papoulias, M. V. Thomas, M. L. Annis, and J. Boase. 2009. Sex assignment of Lake Sturgeon (*Acipenser fulvescens*) based on plasma sex hormone and vitellogenin levels. Journal of Applied Ichthyology 25:60–67.

Dammerman, K. J., M. A. Webb, and K. T. Scribner. 2019. Riverine characteristics and adult demography influence female Lake Sturgeon (*Acipenser fulvescens*) spawning behavior, reproductive success, and ovarian quality. Canadian Journal of Fisheries and Aquatic Sciences 76:1147–1160.

Diamond, J. M. 1975. The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. Biological Conservation 7:129–146.

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. Marshall, and C. M. Mayer. 2015. Portfolio theory as a management tool to guide conservation and restoration of multi-stock fish populations. Ecosphere [online serial] 6(12):296.

Duong, T. Y., K. T. Scribner, P. S. Forsythe, J. A. Crossman, and E. A. Baker. 2013. Inter-annual variation in effective number of breeders and estimation of effective population size in long-lived iteroparous Lake Sturgeon (*Acipenser fulvescens*). Molecular Ecology 22:1282–1294.

Duong, Y., K. T. Scribner, J. A. Crossman, P. Forsythe, and E. Baker. 2011. Environmental and maternal effects on timing and duration of dispersal of larval Lake Sturgeon (*Acipenser fulvescens*). Canadian Journal of Fisheries and Aquatic Sciences 68:643–654.

Fischer, J. L., J. J. Pritt, E. F. Roseman, C. G. Prichard, J. M. Craig, G. W. Kennedy, and B. A. Manny. 2018. Lake Sturgeon, Lake Whitefish, and Walleye egg deposition patterns with response to fish spawning substrate restoration in the St. Clair–Detroit River system. Transactions of the American Fisheries Society 147:79–93.

Fontana, F., M. Lanfredi, M. Chicca, N. Beltrami, and L. Congiu. 2004. Karyotype characterization of the Lake Sturgeon, *Acipenser fulvescens* (Rafinesque 1817) by chromosome banding and fluorescent in situ hybridization. Genome 47:742–746.

Gisbert, E., and G. I. Ruban. 2003. Ontogenetic behavior of Siberian Sturgeon, *Acipenser baerii*: a synthesis between laboratory tests and field data. Environmental Biology of Fishes 67:311–391.

Goodyear, C. S., T. A. Edsall, D. M. Ormsby Dempsey, G. D. Moss, and P. E. Polanski. 1982. Atlas of the spawning and nursery areas of Great Lakes fishes. U.S. Fish and Wildlife Service FWS/OBS-82/52.

Hay-Chmielewski, E. M., and G. E. Whelan. 1997. Lake Sturgeon rehabilitation strategy. Michigan Department of Natural Resources Fisheries Division Special Report 18.

Hondorp, D. W., E. F. Roseman, and B. A. Manny. 2014. An ecological basis for future fish habitat restoration efforts in the Huron–Erie corridor. Journal of Great Lakes Research 40:23–30.

Hunter, R. D. 2018. Assessing reproductive success of Lake Sturgeon (*Acipenser fulvescens*) associated with natural and constructed spawning reefs in a large river system using pedigree analysis. Master's thesis. Michigan State University, East Lansing.

Jay, K., J. A. Crossman, and K. T. Scribner. 2014. Estimates of effective number of breeding adults and reproductive success for White Sturgeon. Transactions of the American Fisheries Society 143:1204–1216.

Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10:551–555.

Kessel, S. T., D. W. Hondorp, C. M. Holbrook, J. C. Boase, J. A. Chiotti, M. V. Thomas, T. C. Wills, E. F. Roseman, R. Drouin, and C. C. Krueger. 2018. Divergent migration within Lake Sturgeon

(*Acipencer fulvescens*) populations: multiple distinct patterns exist across an unrestricted migration corridor. Journal of Animal Ecology 87:259–273.

King, T. L., B. A. Lubinski, and A. P. Spidle. 2001. Microsatellite DNA variation in Atlantic Sturgeon (*Acipenser oxyrinchus oxyrinchus*) and cross-species amplification in the Acipenseridae. Conservation Genetics 2:103–119.

Lantry, B. F., R. O'Gorman, M. G. Walsh, J. M. Casselman, J. A. Hoyle, M. J. Keir, and J. R. Lantry. 2007. Reappearance of Deepwater Sculpin in Lake Ontario: resurgence or last gasp of a doomed population? Journal of Great Lakes Research 33(Supplement):34–45.

Levitan, D. R. 2004. Density-dependent sexual selection in external fertilizers: variance in male and female fertilization success along the continuum from sperm limitation to sexual conflict in the sea urchin *Strongylocentrotus franciscanus*. American Naturalist 164:298–309.

Levitan, D. R., and C. Petersen. 1995. Sperm limitation in the sea. Trends in Ecology and Evolution 10:228–231.

Manny, B. A., and G. W. Kennedy. 2002. Known Lake Sturgeon (*Acipenser fulvescens*) spawning habitat in the channel between Lakes Huron and Erie in the Laurentian Great Lakes. Journal of Applied Ichthyology 18:486–490.

Manny, B. A., E. F. Roseman, G. Kennedy, J. C. Boase, J. M. Craig, D. H. Bennion, J. Read, L. Vaccaro, J. Chiotti, R. Drouin, and R. Ellison. 2015. A scientific basis for restoring fish spawning habitat in the St. Clair and Detroit rivers of the Laurentian Great Lakes. Restoration Ecology 23:149–156.

Marranca, J. M., A. B. Welsh, and E. F. Roseman. 2015. Genetic effects of habitat restoration in the Laurentian Great Lakes: an assessment of Lake Sturgeon origin and genetic diversity. Restoration Ecology 23:455–464.

May, B., C. C. Krueger, and H. L. Kinkaid. 1997. Genetic variation at microsatellite loci in sturgeon: primer sequence homology in *Acipenser* and *Scaphirhynchus*. Canadian Journal of Fisheries and Aquatic Sciences 54:1542–1547.

McCullough, D. E., E. F. Roseman, K. M. Keeler, R. L. DeBruyne, J. J. Pritt, P. A. Thompson, S. Ireland, J. E. Ross, D. Bowser, R. D. Hunter, and D. Castle. 2015. Evidence of the St. Clair–Detroit River system as a dispersal corridor and nursery habitat for transient larval Burbot. Hydrobiologia 757:21–34.

McDougall, C. A., A. B. Welsh, T. Gosselin, W. G. Anderson, and P. A. Nelson. 2017. Rethinking the influence of hydroelectric development on gene flow in a long-lived fish, the Lake Sturgeon *Acipenser fulvescens*. PLoS ONE [online serial] 12:e0174269.

McLean, M., E. F. Roseman, J. J. Pritt, G. Kennedy, and B. A. Manny. 2015. Artificial reefs and reef restoration in the Laurentian Great Lakes. Journal of Great Lakes Research 41:1–8.

McQuown, E., G. A. E. Gall, and B. May. 2002. Characterization and inheritance of six microsatellite loci in Lake Sturgeon. Transactions of the American Fisheries Society 131:299–307.

McQuown, E. C., B. L. Sloss, R. J. Sheehan, J. Rodzen, G. J. Tranah, and B. May. 2000. Microsatellite analysis of genetic variation in sturgeon: new primer sequences for *Acipenser* and *Scaphirhynchus*. Transactions of the American Fisheries Society 129:1380–1388.

Nichols, S. J., G. Kennedy, E. Crawford, J. Allen, J. French III, G. Black, M. Blouin, J. Hickey, S. Chernyak, R. Haas, and M. Thomas. 2003. Assessment of Lake Sturgeon (*Acipenser fulvescens*) spawning efforts in the lower St. Clair River, Michigan. Journal of Great Lakes Research 29:383–391.

Prichard, C. G., J. M Craig, E. F. Roseman, J. L Fischer, B. A. Manny, and G. W. Kennedy. 2017. Egg deposition by lithophilic-spawning fishes in the Detroit and St. Clair rivers, 2005–2014. U.S. Geological Survey Scientific Investigations Report 2017-5003.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation For Statistical Computing, Vienna. Available: https://www.R-project.org/. (February 2020).

Rodzen, J. A., T. R. Famula, and B. May. 2004. Estimation of parentage and relatedness in the polyploid White Sturgeon (*Acipenser transmontanus*) using a dominant marker approach for duplicated microsatellite loci. Aquaculture 232:165–182.

Rodzen, J. A., and B. May. 2002. Inheritance of microsatellite loci in White Sturgeon (*Acipenser transmontanus*). Genome 45:1064–1076.

Roseman, E. F., D. J. Jude, T. G. Coon, M. K. Raths, and W. W. Taylor. 1998. Occurrence of Deepwater Sculpin (*Myoxocephalus thompsoni*) in western Lake Erie. Journal of Great Lakes Research 24:479–483.

Roseman, E. F., B. A. Manny, J. Boase, M. Child, G. W. Kennedy, J. Craig, K. Soper, and R. Drouin. 2011. Lake Sturgeon response to a spawning reef constructed in the Detroit River. Journal of Applied Ichthyology 27:66–76.

Ryman, N., S. Palm, C. Andre, G. R. Carvalho, T. G. Dahlgren, P. E. Jorde, L. Laikre, L. C. Larsson, A. Palme, and D. E. Ruzzante. 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. Molecular Ecology 15:2031–2045.

Sard, N. M., K. G. O'Malley, D. P. Jacobson, M. J. Hogansen, M. A. Johnson, and M. A. Banks. 2015. Factors influencing spawner success in a spring Chinook Salmon (*Oncorhynchus tshawytscha*) reintroduction program. Canadian Journal of Fisheries and Aquatic Sciences 72:1390–1397.

Sard, N. M., S. R. Smith, J. J. Homola, J. Kanefsky, G. Bravener, J. V. Adams, C. M. Holbrook, P. J. Hrodey, K. Tallon, and K. T. Scribner. In press. RAPTURE (RAD capture) panel facilitates analyses characterizing Sea Lamprey reproductive ecology and movement dynamics. Ecology and Evolution. DOI: 10.1002/ece3.6001.

Schindler, D. E., R. Hilborn, B. Chasco, C. P. Boatright, T. P. Quinn, L. A. Rogers, and M. S. Webster. 2010. Population diversity and the portfolio effect in an exploited species. Nature 465:609–612.

Schwartz, M. K., G. Luikart, and R. S. Waples. 2007. Genetic monitoring as a promising tool for conservation and management. Trends in Ecology and Evolution 22:25–33.

Steele, C. A., M. Hess, S. Narum, and M. Campbell. 2019. Parentage-based tagging: reviewing the implementation of a new tool for an old problem. Fisheries 44:412–422.

Talbot, C. C. Jr., D. Avrampopoulos, S. Gerken, A. Chakravarti, J. A. Armour, N. Matsunami, R. White, and S. E. Antonarakis. 1995. The tetranucleotide repeat polymorphism D21S1245 demonstrates hypermutability in germline and somatic cells. Human Molecular Genetics 4:1193–1199.

Thomas, M. V., and R. C. Haas. 1999. Capture of Lake Sturgeon with setlines in the St. Clair River, Michigan. North American Journal of Fish Management 19:610–612.

Thomas, M. V., and R. C. Haas. 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.

Thorstensen, M., P. Bates, K. Lepla, and A. Schreier. 2019. To breed or not to breed? Maintaining genetic diversity in White Sturgeon supplementation programs. Conservation Genetics 20:997–1007.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166:1963–1979.

Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology 18:2148–2164.

Wang, J., and A. W. Santure. 2009. Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181:1579–1594.

Wang, J., and K. T. Scribner. 2014. Parentage and sibship inference from markers in polyploids. Molecular Ecology Resources 14:541–553.

Waples, R. S. 1990. Conservation genetics of Pacific salmon. II. Effective population size and the rate of loss of genetic variability. Journal of Heredity 18:267–276.

Welsh, A. B., M. Blumberg, and B. May. 2003. Identification of microsatellite loci in Lake Sturgeon, *Acipenser fulvescens*, and their variability in Green Sturgeon, *A. medirostis*. Molecular Ecology Notes 3:47–55.

Welsh, A., T. Hill, H. Quinlan, C. Robinson, and B. May. 2008. Genetic assessment of Lake Sturgeon population structure in the Laurentian Great Lakes. North American Journal of Fisheries Management 28:572–591.

Williamson, K. S., A. R. Murdoch, T. N. Pearsons, E. J. Ward, and M. J. Ford. 2010. Factors influencing the relative fitness of hatchery and wild spring Chinook Salmon (*Oncorhynchus tshawytscha*) in the Wenatchee River, Washington, USA. Canadian Journal of Fisheries and Aquatic Sciences 67:1840–1851.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

SUPPORTING INFORMATION

Additional supplemental material may be found online in the Supporting Information section at the end of the article.