Hatchery Strain Contributions to Emerging Wild Lake Trout Populations in Lake Huron

Kim Scribner[1,2], IyobTsehaye[1,3,*], Travis Brenden[1,3], Wendylee Stott[1,4], Jeannette Kanefsky[1], James Bence[1,3]

[1]Dept. Fisheries & Wildlife, Michigan State Univ., 13 Natural Resources Bldg, East Lansing, MI 48824

[2]Dept. Integrative Biology, Michigan State Univ., 201 Natural Sciences Bldg, East Lansing, MI 48824

[3]Quantitative Fisheries Center, Michigan State Univ., 13 Natural Resources Bldg , East Lansing, MI 48824

[4]Great Lakes Science Center, 1451 Green Rd., Ann Arbor, MI 48105-2807

*present address – Wisconsin Dept. Natural Resources, Madison, WI

Received xxxxxx; Accepted xxxxxxx

Subject area: Conservation genetics and biodiversity

Abstract

Recent assessments have indicated the emergence of naturally produced lake trout (*Salvelinus namaycush*) recruitment throughout Lake Huron in the North American Laurentian Great Lakes. By 2013, more than 50% of young fish (<7 yrs) were produced by naturally spawning fish. Because naturally produced fish derived from different stocked hatchery strains are not phenotypically distinct and are unmarked, managers cannot distinguish what strains are contributing to natural recruitment. We used 15 microsatellite loci to characterize the originating strains of naturally produced Lake Huron lake trout (N=1567) collected in assessment fisheries during an early (2002-2004) and late (2009-2012) sampling period. Individuals from 13 American and Canadian hatchery strains (N=1143) were genotyped to develop standardized baseline information for source strains. Strain contributions were estimated using a Bayesian inferential approach, and deviance information criteria was used to compare models evaluating strain contributions at different spatial and temporal scales. The best performing models were the most complex models, suggesting that hatchery strain contributions to naturally produced lake trout varied spatially among management districts and temporally across time periods. Contributions of Seneca strain lake trout were consistently high across management districts, with contributions increasing from early to late time periods. Strain contributions deviated from expectations based on historical stocking levels, suggesting strains differed with respect to survival, reproductive success, and/or dispersal. Knowledge of recruitment levels of strains stocked in different management districts within a lake, and how strain-specific recruitment varies temporally, spatially, and as a function of local or regional stocking is important to target strains to prioritize for future stocking and management of the transition process from primarily hatchery stocks to naturally produced stocks in Lake Huron and other Great Lakes.

Key words: Great Lakes, recruitment, restoration, *Salvelinus*, stocking

Introduction

In the Laurentian Great Lakes of North America, lake trout (*Salvelinus namaycush*) experienced considerable reductions in population abundance and distribution over the last two centuries (Hansen 1999). Historically, lake trout were a dominant predator in the lakes and were important drivers of human settlement around the basin (Muir et al. 2013). During the 19th and 20th centuries, native lake trout stocks declined in each lake owing to over-exploitation, parasitism by sea lamprey (*Petromyzon marinus*), predation and competition stemming from the introductions and spread of alewife (*Alosa pseudoharengus*) and rainbow smelt (*Osmerus mordax*), and anthropogenic effects on water quality (Hansen 1999; Muir et al. 2013). Management actions were undertaken in the 1950s to restore lake trout populations in the Great Lakes, including stocking of juvenile fish, closure of commercial fisheries, and reduction of sea lamprey populations through lamprey control efforts (Muir et al. 2013). After decades of considerable effort, however, lake trout restoration in the Great Lakes has still not been fully realized except in Lake Superior (Muir et al. 2013). In Lake Superior, detailed information on hatchery strain-specific recruitment and survival were not available when stocks were recovering, which are required to unambiguously quantify the relative contributions of environmental, ecological, and management actions to the restoration of self-sustaining lake trout populations. For example, debate exists surrounding the relative importance of recruitment from remnant wild stocks or hatchery strains to lake trout restoration (Schram et al. 1995; Guinand et al. 2004).

Lake Huron is presently in the early stages of a lake trout restoration success like that seen on Lake Superior (Johnson et al. 2015). Lake trout stocking began in Lake Huron in 1973, and thereafter annual stocking levels have ranged between 1.39 and 1.95 million yearlings (He et al. 2012). Since the mid-2000s, agency assessments have indicated the emergence of naturally produced lake trout recruitment in most management units based on collections of age-0 and untagged sub-adult and adult fish (Riley et al. 2007; He et al. 2012). The year-class strength or annual relative abundance of naturally produced lake trout has increased dramatically over time in three locations where long-term monitoring has occurred:

Drummond Island, American off-shore reefs, and Canadian jurisdictions of the central main basin (He et al. 2012; Johnson et al. 2015). Fifty percent of young (<7 yr) lake were naturally produced by 2013 (Johnson et al. 2015). Naturally produced year-class abundances in different areas of Lake Huron are positively correlated, suggesting that factors contributing to improvements in natural recruitment are occurring at lake-wide scales. Basin-wide adult catch per effort (CPE) has been found to be correlated with year class strength of naturally produced fish suggesting a connection between the abundance of spawning stocks and levels of natural reproduction (Fitzsimons et al. 2010; He et al. 2012). Natural recruitment further increased as thiamine concentration in lake trout eggs increased after the alewife population collapse in 2003 (Riley et al. 2012). Presently, naturally produced lake trout have been estimated to compose approximately 40% of the total lake trout spawning biomass in Lake Huron (Ji He, Michigan DNR, personal communication).

A critical need to support lake trout restoration efforts is improved understanding of recruitment variation over space and time (Zimmerman and Krueger 2009). Because naturally produced fish derived from hatchery parents are not phenotypically distinct, managers lack the ability to characterize strain-specific rates of recruitment and whether recruitment at specific locales can be attributed to either reproductive success of strains stocked at or near each sampling station, or to dispersal and immigration of individuals from strains stocked elsewhere (e.g., source-sink effects; Pullium 1988). Previous research (Page et al. 2003; Madenjian et al. 2004) has shown that strain abundance at spawning locales is a poor predictor of strain-specific reproductive success, but this work was not spatially extensive enough to address the general questions of the effects of dispersal and local stocking on the abundance and spatial distribution of naturally produced lake trout origination from different strains. Knowledge of recruitment levels of hatchery strains stocked in different management units, and how recruitment varies across time periods, locations, and as a function of local or regional (i.e., statistical reporting districts or combinations of districts in American and Canadian waters) stocking efforts, will provide information needed to improve future stocking efforts in Lake Huron and other Great Lakes.

The main goal of our research was to determine relative contributions of stocked hatchery strains to

emerging naturally produced lake trout populations in Lake Huron. Specific objectives were to: (1) determine whether strain contributions differed spatially among Lake Huron sub-basin regions, or combinations of these (e.g., American versus Ontario jurisdictional waters); (2) determine the degree of temporal consistency in hatchery strain contributions within management districts; (3) determine whether strain contributions deviated from what was expected given historical stocking levels and incorporating current knowledge about post-stocking dispersal and survival of lake trout; and (4) quantify levels of assortative mating and inferentially the proportion of inter-strain hybrids in naturally produced mixtures in different regions of Lake Huron during different years. Our specific hypotheses that we were interested in testing included the following:

(a) the strain composition of naturally produced lake trout varied among management districts within Lake Huron ;

(b) the strain composition of naturally produced lake trout differed between time periods within a management district.

(c) the strain composition of naturally produced lake trout within a management district was consistent with stocking levels of hatchery fish in that lake district which are sexually mature at the time naturally produced-caught lake trout were produced;

(d) dispersal by lake trout strains from regions of stocking contributed to recruitment of naturally produced lake trout in different regions.

We note that these are hypotheses to test and not statements of belief. For example if (d) were true this would mean that (c) would not be.

Methods

We utilized genotypes from microsatellite DNA loci from naturally produced lake trout and from 13 hatchery strains stocked into Lake Huron (Table 1). We used samples of naturally produced subadults and adults (age range 4-10 yrs) collected in assessment fisheries conducted by agency and tribal cooperators during March-December during two time periods that we designated as ‘early’ (2002-2004)

and ‘late’ (2009-2012) time periods, which generally correspond to periods of naturally produced lake trout emergence and establishment, respectively (see Figure 1 for details concerning timing and ages of samples from American and Canadian locations).

Sample collections - Samples of naturally produced lake trout were collected by cooperating agencies including the Michigan Department of Natural Resources (MDNR), Chippewa Ottawa Resource Authority (CORA), U.S. Fish and Wildlife Service (USFWS), Ontario Ministry of Natural Resources and Forestry (OMNRF) and U.S. Geological Survey (USGS) from 10 American (MH1, MH2, MH345) and Canadian (OH1, OH2&3, OH4&5, NC1&2, NC3, GB1-3, GB4) defined regions located throughout Lake Huron (Figure 2). These names refer to spatial regions corresponding to management districts or amalgamation of such districts (e.g., OH4&5 corresponds to the combined OH4 and OH5 statistical districts in Canadian waters of Lake Huron). Samples were collected using multifilament nylon gill nets or trap nets that were bottom set overnight across depth contours (see He et al. 2012 for details). Sample sizes by period and location are detailed in Table 1. Individuals were identified as “naturally produced” based on absence of coded wire tags (CWT) and absence of clipped fins. Tissue samples were preserved in 95% ethanol (fins, muscle) or were placed in scale envelopes and allowed to dry (fins, scales). The date of capture, total length, capture site, and age were recorded for each lake trout. Age determinations were made for all individuals using either pectoral fin rays, otoliths, or (for smaller fish) scales.

Genetic data - DNA was extracted from fin or scale tissue for both mixture and baseline individuals using QIAGEN DNeasy kits (QIAGEN Inc., Germantown, MD) using manufacturer’s protocols. A spectrophotometer was used to quantify DNA concentrations for all samples for use in PCR reactions.

Individuals were genotyped at 15 microsatellite loci: Ogo1a (Olsen et al. 1998); One9 (Scribner et al. 1996); Sco19 (Taylor et al. 2001); Sfo1, Sfo12 and Sfo18 (Angers et al. 1995); Sco202 (DeHaan and Ardren 2005); SfoC38 and SfoC88 (King et al. 2012); and SnaMSU01, SnaMSU03, SnaMSU05, SnaMSU08, SnaMSU10 and SnaMSU11 (Rollins et al. 2009). All loci were amplified by PCR in single

locus reactions. Samples of American lake trout hatchery strains and samples from American (Michigan) management districts in Lake Huron were genotyped at Michigan State University. Samples of Canadian lake trout hatchery strains and samples from Canadian (Ontario) management districts in Lake Huron were genotyped at the U.S. Geological Survey Great Lakes Science Center.

PCR conditions and data standardization protocols are described in the Supplemental Materials Methods. All genotypes were independently scored by two experienced lab personnel and 10% of the samples were randomly selected and re-genotyped at all 15 loci. Genotype scores were compared with the original scores to derive an empirical estimate of scoring error, which was estimated to be 0.4% as averaged across all 15 loci.

Data Analysis

Estimation of allele frequency and summary measures of genetic diversity-

Estimates of allele frequency and summary measures of genetic diversity including heterozygosity, number of alleles per locus, and Wright's inbreeding coefficient (F_{is}) for hatchery strains and naturally produced individuals collected during the early and late time periods and from different management districts were estimated using program FSTAT (Guodet 2001). F-statistics (Weir and Cockerham 1984) quantifying the variance in allele frequency among lake trout hatchery strains were also calculated using FSTAT.

As a demonstration that our lake trout hatchery baseline data would permit accurate assignment of individuals to their strain of origin, we simulated single-population samples (i.e., 100% mixture simulations). Simulations were conducted in program ONCOR (Kalinowski et al. 2008). The program implements simulations described by Anderson et al. (2007) involving the generation of multiple (1000 iterations) mixtures comprised of solely one hatchery strain. Bootstrapped mixture sample sizes were simulated for 200 fish, and the hatchery baseline sample sizes were set equal to the actual sample sizes for empirical American and Canadian hatchery strains. Our target accuracy level for the 100% mixture

simulations was 90% for each spawning populations, a target accuracy benchmark used previously in empirical fisheries literature (Seeb and Crane 1999; Beacham et al. 2012).

Estimation of hatchery strain contributions to first-generation lake trout mixtures in Lake Huron - The analysis to estimate hatchery strain contributions to emerging naturally produced lake trout populations in Lake Huron was based on a model described by Gaggiotti et al (2002; 2004) and expanded by Guo et al. (2008) for quantifying source population contributions to newly founded colonies. A basic assumption of the model is that all naturally produced fish are first-generational (F1) descendants of hatchery fish previously stocked by Great Lakes fishery management agencies (Gaggiotti et al. 2002; 2004). Potential violations of this assumption are addressed in the Discussion. A feature of the estimation model is that it allows for assortative mating of individuals from the same hatchery strain (Gagiotti et al. 2002; 2004), which could arise from various factors such as differences in spawning habitat or limited dispersal from stocking locations. In the original formulation of the model by Gaggiotti et al. (2002; 2004), source population contributions were modeled as functions of biotic or abiotic data, such as distance of source populations from the newly formed colony or some measure of reproductive productivity of the populations. We chose to not use this approach for the hatchery strain contributions and instead estimated the contributions as freely-varying parameters (albeit recognizing the unit-sum constraint of the contributions).

Hatchery strain contributions to the naturally produced lake trout populations were estimated using a Bayesian inferential approach, which better allows for characterization of uncertainty in parameter estimates. For this approach, the posterior probability distribution for the unknown parameters [i.e., hatchery strain contributions ($\boldsymbol{p}$), assortative mating coefficient (ω), and allele relative frequencies of the hatchery strains ($\boldsymbol{Q}$)], can be specified as

$$\pi\left(\boldsymbol{Q}, \boldsymbol{p}, \omega \mid \boldsymbol{Y}, \boldsymbol{X}\right) \propto \pi\left(\boldsymbol{Y} \mid \boldsymbol{Q}, \boldsymbol{p}, \omega\right) \pi\left(\boldsymbol{Q} \mid \boldsymbol{X}\right) \pi\left(\boldsymbol{p}\right) \pi\left(\omega\right) \quad (1)$$

where $\boldsymbol{Y}$ is the multi-locus genotypes observed in a sample of naturally produced lake trout, $\boldsymbol{X}$ is the genotypes observed in samples taken from the hatchery strains, $\pi(\boldsymbol{p})$ and $\pi(\omega)$ are the prior probability distributions assumed for the hatchery strain contributions and assortative mating coefficient, respectively, $\pi(\boldsymbol{Q} \mid \boldsymbol{X})$ is the prior probability distribution for allele relative frequencies of the hatchery strains given the collection and genotyping of individuals from the strains, and $\pi(\boldsymbol{Y} \mid \boldsymbol{Q}, \boldsymbol{p}, \omega)$ is the probability of observing the multi-locus genotypes observed in a sample of naturally produced lake trout for given values of $\boldsymbol{Q}$, $\boldsymbol{p}$, and ω (i.e., the model likelihood).

A Dirichlet probability density function was assumed for $\pi(\boldsymbol{p})$ with concentration parameters set equal to the inverse of the number of hatchery strains, meaning that prior to data collection each hatchery strain was assumed to contribute equally to the mixture. A uniform probability density function with lower and upper bounds of 0.0 and 1.0 was assumed for $\pi(\omega)$. Our specification of $\pi(\boldsymbol{Q} \mid \boldsymbol{X})$ followed that of Rannala and Mountain (1997). Specifically, the Rannala and Mountain (1997) approach is based on a separate Bayesian analysis, where the posterior distribution from that analysis provides the prior $\pi(\boldsymbol{Q} \mid \boldsymbol{X})$ for the mixture model. The separate Bayesian analysis is based on the assumption that before individuals are sampled from the hatchery strains, the alleles at a locus in each strain are considered equally likely. This prior belief is then updated based on the number of observed copies of an allele for a particular locus and hatchery strain once individuals from the strains have been collected and genotyped (Rannala and Mountain 1997). Thus, $\pi(\boldsymbol{Q} \mid \boldsymbol{X})$ is an informative prior based on samples from the hatchery strains. As in Rannala and Mountain (1997), a Dirichlet probability density function was assumed for $\pi(\boldsymbol{Q} \mid \boldsymbol{X})$. When fitting the mixture models, the parameters of $\pi(\boldsymbol{Q} \mid \boldsymbol{X})$ were fixed so that the distribution of $\boldsymbol{Q}$ was not updated as part of the model fitting process.

The probability of observing the multi-locus genotypes observed in a sample of naturally produced lake trout for given values of $\boldsymbol{Q}$, $\boldsymbol{p}$, and ω followed directly from Gaggiotti et al. (2002; 2004) as well as from Guo et al. (2008)

$$\pi\left(\boldsymbol{Y} \mid \boldsymbol{Q}, \boldsymbol{p}, \omega\right) = \prod_{m=1}^{M}\left(\omega\sum_{i=1}^{I} p_i f\left(\boldsymbol{y}_m \mid ii\right) + \left(1-\omega\right)\left[\sum_{i=1}^{I} p_i^2 f\left(\boldsymbol{y}_m \mid ii\right) + \sum_{i=1}^{I}\sum_{j\neq i}^{J} p_i p_j f\left(\boldsymbol{y}_m \mid ij\right) \right]\right), \quad (2)$$

where M is the total number of sampled naturally produced lake trout, p_i and p_j are the proportional contributions of the i-th and j-th hatchery strains to the naturally produced lake trout population (elements of $\boldsymbol{p}$), and $f(\boldsymbol{y}_m|ij)$ is a genetic model describing the probability a lake trout having the genotype of the m-th individual given that one parent is of the i-th strain and another parent is from the j-th strain (potentially $i = j$ when both parents are from the same strain). For individuals with both parents from the same strain (i.e., $i = j$), the probability of a lake trout having the genotype of the m-th individual is

$$f\left(\boldsymbol{y}_m \mid ii\right) = \prod_{l=1}^{L} \delta_{lm} q_{a_{1lm;li}} q_{a_{2lm;li}} \quad (3)$$

where $q_{a_{1lm;li}}$ is the allele frequency of the l-th locus in the i-th hatchery strain corresponding to the first allele observed in the m-th individual at the l-th locus, $q_{a_{2lm;li}}$ is the allele frequency of the l-th locus in the i-th hatchery strain corresponding to the second allele observed in the m-th individual at the l-th locus, and δ_{lm} is an indicator variable defined as

$$\delta_{lm} = \begin{cases} 1 & \text{if } a_{1lm} = a_{2lm} \\ 2 & \text{if } a_{1lm} \neq a_{2lm} \end{cases}. \quad (4)$$

For individuals with parents from two different strains (i.e., $i \neq j$), the probability of a lake trout having the genotype of the m-th individual is

$$f\left(\boldsymbol{y}_m \mid ij\right) = \prod_{l=1}^{L}\left(q_{a_{1lm;li}} q_{a_{2lm;lj}} + \gamma_{lm} q_{a_{2lm;li}} q_{a_{1lm;lj}} \right) \quad (5)$$

where

$$\gamma_{lm} = \begin{cases} 0 & \text{if } a_{1lm} = a_{2lm} \\ 1 & \text{if } a_{1lm} \neq a_{2lm} \end{cases}. \quad (6)$$

Given that ω corresponds to the probability of a naturally produced lake trout arising from assortative mating among hatchery strains, 1-ω corresponds to the probability of a naturally produced lake trout arising from random mating among the strains (Gaggiotti et al. 2002; 2004).

To assess the degree of spatial and temporal consistency in the hatchery strain contributions, we fit a series of models to different groupings of naturally produced lake trout data. The groupings consisted of two categories: 1) spatial/temporal, and 2) spatial. The spatial/temporal grouping involved naturally produced lake trout collected from MH-1, MH-2, MH-345, and OH-1 where samples were available from both early and late time periods. For most other Lake Huron statistical districts, sample sizes of naturally produced lake trout during early periods were low (in many cases 0; Table 1), which is why we limited analyses to just these four management districts. The spatial grouping involved naturally produced lake trout collected from the spatial management districts GB-1234, MH-1, MH-2, MH-345, OH-1, OH-23, OH-45, NC-12, and NC-3 during the late period only. For each grouping, four models were fit. In the case of the spatial/temporal grouping, the four models that were fit were pooled (common strain contributions and assortative mating coefficients for all spatial regions and time periods), separate (unique strain contributions and assortative mating coefficients for each spatial region and time period), spatial (unique strain contributions and assortative mating coefficients for each spatial region but pooled over time periods), and temporal (unique strain contributions and assortative mating coefficients for each time period but pooled over spatial region). In the case of the spatial grouping, the four models that were fit were pooled (common strain contributions and assortative mating coefficients for all spatial regions), separate (unique strain contributions and assortative mating coefficients for each spatial region), Michigan vs. Ontario (unique strain contributions and assortative mating coefficients by jurisdictional authority), and basin [unique strain contributions and assortative mating coefficients by Lake Huron basin (i.e., main basin, Georgian Bay, North Channel)].

The estimation procedure was programmed in AD Model Builder, which includes a Metropolis-Hasting algorithm for conducting Markov chain Monte Carlo (MCMC) approximation to posterior probability distributions (Fournier et al. 2012). The objective function used to estimate the models equaled the sum of the negative $\log_e$ likelihood and negative $\log_e$ priors specified above. For most models, MCMC chains were run for 1 million steps, sampling every 100th step, and discarding the initial 3,000 saved steps as a burn-in period. For the separate model under the spatial/temporal grouping MCMC chains were run for 1.5 million steps, sampling every 100th step and discarding the initial 5,000 saved steps as a burn-in period. For the separate model under the spatial grouping, MCMC chains were run for 3.0 million steps, sampling every 100th step and discarding the initial 20,000 saved steps as a burn-in period. Convergence of the MCMC chain for each model was evaluated by constructing trace plots for the model negative $\log_e$ likelihood as a visual check to ensure the chain was well-mixed and using Z-score tests to evaluate differences between the means of the first 10% and last 50% of the saved chain (Geweke 1992). Additionally, we calculated effective sample size of the saved MCMC chain (for $\log_e$ likelihood) to evaluate whether the saved chain contained enough information to make inferences about the posterior distribution. With an effective sample size in the thousands we believe such inferences are easily supported. Convergence diagnostics were conducted on the model negative $\log_e$ likelihood as an omnibus test of convergence given that some estimated models had in excess of 100 parameters being estimated. Means of the posterior probability distributions for the model parameters were used as parameter point estimates. Ninety-five percent highest posterior density intervals were used to characterize the uncertainty associated with each parameter. All MCMC diagnostic measures and highest posterior density interval calculations were conducted in R (R Core Team, 2012) using the "coda" package (Plummer et al. 2006).

Performance of the four models fit to each grouping of the lake trout data was assessed using deviance information criteria (DIC) (Spiegelhalter et al. 2012). DIC was calculated as

$$\mathrm{DIC} = \bar{D} + p_D \tag{7}$$

where $\bar{D}$ is the average deviance for a model measuring fit and p_D and is the effective number of parameters. The average deviance for a model was calculated as

$$\bar{D} = \frac{1}{C}\sum_{c=1}^{C} -2\log_e\left(\pi\left(\boldsymbol{Y} \mid \boldsymbol{Q}, \boldsymbol{p}, \omega\right)\right) \quad (8)$$

with C equal to the number of MCMC steps saved minus the burn-in, whereas p_D was calculated as

$$p_D = \bar{D} - D\left(\bar{\theta}\right) \quad (9)$$

where $D\left(\bar{\theta}\right)$ is the deviance evaluated at the posterior mean parameter estimates.

Expected hatchery strain contributions

Expected contributions of hatchery strains to naturally produced lake trout populations were calculated by combined numbers of hatchery strains stocked in various management districts in Lake Huron (Supplemental Table 1) with mortality estimates for Lake Huron lake trout from statistical catch-at-age assessment (SCAA) models fit to different regions of the lake (Supplemental Table 2) and an assumed movement matrix (Supplemental Table 3) that was used to allocate stocked lake trout to different regions of the lake and based on analyses of coded-wire tagging data from lake trout (Adlerstein et al. 2007). We used numbers of hatchery strains stocked in various management units in Lake Huron from American agencies (data available at http://www.glfc.org/fishstocking/) and the OMNRF (Adam Cottrill, personal communication). When calculating expected contributions of hatchery strains, we assumed that all early lake trout data were obtained in 2003 and all late lake trout data were obtained in 2011. Assuming lake trout of ages 7–14 were mature, the fish that contributed to the spawning event in 1997 that resulted in the fish collected in 2003 (which for simplicity we assumed to be all 6 year olds) were stocked from 1983-1990. Similarly, the fish that contributed to the spawning event in 2005 that resulted in the fish collected in 2011 were stocked from 1991-1998. The contributions to these spawning events of stocked fish stocked in a given year were assumed to vary depending on their age (or stocking year), calculated using the mortality estimates derived from catch-at-age models.

Results

Characteristics of American and Canadian hatchery strains used in mixture analyses

In total, 1675 naturally produced caught and 1143 hatchery lake trout were genotyped (Table 1). Estimates of allele frequency and summary measures of genetic diversity for all strains of hatchery lake trout are provided in Supplemental Tables 4. American hatchery strains generally were characterized by higher levels of genetic diversity than Canadian hatchery strains (Supplemental Table 4) including allelic richness (A_R range 6.67 to 9.04 across American strains vs 5.68 to 8.21 across Canadian strains), multi-locus expected heterozygosity (H_E ranged from 0.502 to 0.678 for American strains vs 0.410 to 0.592 for Canadian strains) and Wright's inbreeding coefficient (F_{is} range -0.037 to 0.037 for American strains vs -0.017 to 0.093 for Canadian strains). With the exception of Canadian Big/Parry Sound and Michipicoten strains, genotype frequencies were in approximate Hardy Weinberg equilibrium (Supplemental Table 4) and there was no evidence of significant gametic disequilibrium between loci for any strain ($P>0.05$ after Bonferroni correction for multiple testing).

We documented high levels of hatchery inter-strain variance in allele frequency (mean F_{st}=0.061, $P<0.001$; Supplemental Table 5). Supplemental Table S6 shows pair-wise estimates of variance (F_{st}). Estimates of variance in allele frequency (inter-strain F_{st}) ranged from 0.0091 to 0.091 ($P<0.001$ for all inter-strain comparisons). Pairwise inter-strain estimates of F_{st} were highest for the American and Canadian Seneca strain lake trout (Supplemental Table 6), which were the only strains that originated outside the Great Lakes. Large inter-strain variance in allele frequency was further reflected in high strain allocation accuracy estimated based on 100% simulations (Table 2). Slightly lower levels of hatchery strain allocation accuracy for the American Marquette and Traverse Island strains resulted from the 100% simulations (Table 2), with misallocations primarily occurring between these two strains. We attribute this to the American Marquette and Traverse Island strains having both originated from Marquette Bay in Lake Superior. A modest level of misallocation was also observed between the U.S. and Canadian

Seneca strains (Table 2). Collectively, these simulation results showed that strain could be identified with high accuracy based on the genetic data from the potential sources.

Characteristics of open-water mixtures during the early and late sampling periods

Estimates of allele frequency and measures of genetic diversity for the 10 location/period sampling groups are presented in Supplemental Table 7. Generally, expected heterozygosity across the American management districts in both early and late time periods were comparable (H_E ranges 0.591 to 0.620). Greater variability was observed among Canadian management districts (H_E range 0.561 to 0.633). Allelic diversity was generally higher in American management districts (range 7.53 to 11.33) than in Canadian management districts (7.60 to 10.60), generally reflecting the greater genetic diversities of American hatchery strains (Supplemental Table 4) stocked into American waters of Lake Huron. Estimates of Wright's inbreeding coefficient F_{is}, revealed higher positive F_{is} values from mixtures in American waters during the early relative to late period (Supplemental Table 7), indicating less inter-strain mixing or assortative mating in the early relative to late period. Management districts MH1 and MH345 were characterized by significant positive F_{is} indicating heterozygote deficiency (0.056 and 0.054, $P<0.05$, respectively; Supplemental Table 7). Estimates of F_{is} were generally higher in Canadian relative to American sampling units. The magnitude of heterozygote deficiencies likely are a result of the sample mixture composition and lack of interbreeding among hatchery strains. High positive F_{is} values were particularly notable in the Canadian North Channel (NC3) and units in Georgian Bay (GB123 and GB4; Supplemental Table 7).

We observed significant differences in allele frequency among the 10 management districts (mean $F_{st} \pm$ SE= 0.019 $\pm$ 0.005; Supplemental Table 8). Pair-wise estimates of variance in allele frequency (F_{st}) among management districts (Supplemental Table 9) revealed that 35 of 45 pair-wise comparisons (78%) were statistically different from one another in allele frequency after Bonferroni correction for multiple tests.

Differentiation between samples from the early and late periods from the same spatial region

Samples were available in both early and lake periods for the same units in main-basin districts of Lake Huron (MH1, MH2, MH345 for American and OH1 for Canada; Table 4). We observed significant differences in allele frequency between periods for three of the four regions (MH1, F_{st}=0.025, P<0.001; MH3-5, F_{st}=0.005, P<0.001; OH1, F_{st}=0.058, P<0.001; Supplemental Table 9) that was consistent with differences in estimated strain contributions. Differences were most likely attributed to differences in the hatchery strains used to stock in each region and to estimates of hatchery strain contributions across regions (Tables 4 and 5).

Differentiation among samples from spatial regions

We observed greater differences in allele frequency among Canadian regions than among American regions (mean F_{st} among American and Canadian regions were 0.03 and 0.007, respectively; Tables 4 and 5). Levels of genetic differentiation between American and Canadian regions were also consistently high (mean F_{st}=0.024, P<0.001; Supplemental Table 9). Significant differences in allele frequency were documented even for adjacent regions (e.g. NC12 vs NC3, F_{st}=0.012, P<0.001; OH23 vs GB123, F_{st}=0.019, P<0.001)

Hatchery strain contributions to the naturally produced Lake Huron lake trout population

MCMC chains for the negative $\log_e$ likelihoods for each model fit to the spatial/temporal and spatial grouping of the naturally produced lake trout data were found to have converged on stationary distributions. Examination of trace plots (not shown) indicated that chains were well mixed with no apparent stickiness. Geweke (1992) Z-scores for testing convergence of the models ranged from -1.216 to 0.770, suggesting the chains had effectively converged (Table 3). Effective sample sizes of the saved MCMC chains for the $\log_e$ likelihood for the various models ranged from 5011.4 to 7000, indicating both a relatively low level of autocorrelation in the saved chains and adequate information to make inferences about the posterior probability distributions (Table 3). Based on DIC, the separate models had the best

performance for both the spatial/temporal (MH-1, MH-2, MH-345, OH-1 management districts early and late period) and spatial (GB-1234, MH-1, MH-2, MH-345, OH-1, OH-23, OH-45, NC-12, and NC-3 late period only) groupings (Table 3). The DIC weights indicated that the strength of evidence for the separate models for both grouping was overwhelming and there essentially was no empirical support for the other models (Table 3).

For American spatial management districts of Lake Huron, Lewis Lake, American Seneca, and Marquette hatchery strains were in general the greatest contributors to naturally produced lake trout (Table 4). For the MH-1 management district, the Apostle Island and Green Lake strains were estimated to have fairly large (13 to 29%) contributions during the early time period, but the contributions of these strains declined to zero during the late time period. The Green Lake strain was also estimated to have had around a 7% contribution during the early time period for the MH-2 region, but similar to MH-1 the contribution declined to near zero during the late time period (Table 4). The relative contributions of the American Seneca strain increased dramatically from the early to the late period in all American management districts (Table 4), most notably in MH-1 and MH-345. The Canadian Seneca strain was the only strain from Canada to have a notable contribution to naturally produced lake trout production in American management districts. For MH-345 during the early time period, the Canadian Seneca strain contributed around 17% to naturally produced lake trout. During the late time period, the Canadian Seneca strain contributed around 13, 11, and 14% to the MH-1, MH-2, and MH-345 regions, respectively. Contributions from the American Seneca, Lewis Lake and Apostle Island strains exceeded expectations based on historical stocking and historical stocking considering survival and dispersal (Table 4). Representation of the Marquette strain was consistently below expectations based on the same criteria (Table 4). Allele frequency differences between American management districts were not significant in the early sampling period (F_{st} for comparisons between MH1, MH2, MH345 not statistically significant, $P>0.05$; Supplemental Table 9) likely reflecting similarities in relative proportions of strains stocked into all units (Table 4) and high movements expected between units (Supplemental Table 3).

Unlike American districts, there was not a single hatchery strain that composed a majority of contributions to naturally produced lake trout production in Canadian districts. The Lake Manitou strain had the largest contribution in the OH-1 region during the early period and in GB-1234 and NC-3 during the late period (Table 5). Canadian Seneca strain had the largest contribution in OH-23, OH-45, and NC-12 (Table 5). Lake Manitou lake trout were present in high frequency in NC3 during the late period and in all GB regions and in OH1. The American Seneca strain was a large contributor to naturally produced lake trout in several of the Canadian management districts. The American Seneca strain was the largest contributor to OH-1 during the late sampling period, and contributed from 22 to 43% in the GB-1234, NC-12, OH-23, and OH-45 management districts during the late sampling period. The Michipicoten strain was estimated to have contributed 6% and 17% for the GB1234, OH-23, and OH-45 management districts during the late sampling period. The Big/Parry Sound strain was estimated to have contributed around 6% to both the GB-1234 and NC-12 management districts during the late sampling period. The Iroquois Bay strain was estimated to have contributed around 17% in the NC-3 management district, which was in close alignment to the stocking rate in this district (Table 5). The Slate Island strain was not estimated to have made any contribution to any of the districts (Table 5).

Estimates of the assortative mating coefficient (ω in equation 1) were relatively high in the early sampling period in American waters (0.519, 0.558, and 0.270 in statistical districts MH-1, MH-2, and MH-345, respectively; Table 4). However, for the late sampling period, estimates of the assortative mating coefficient for the same units ranged from 0.016 to 0.121 (Table 4). For Canadian waters, the only statistical district for which samples were available during the early period was OH-1, where the estimated assortative mating coefficient for this district and period was 0.01, which was considerably lower than the corresponding values for American waters. During the late sampling period, estimates of the assortative mating coefficient for the OH-1, NC-12, NC-3, OH-45 management districts were comparable to those observed in American waters, ranging from 0.010 to 0.114 (Table 5). However, for OH-23 and GB-1234, the estimates of the assortative mating coefficient ranged from 0.230 to 0.452 (Table 5).

The numbers of fish of each hatchery strain stocked into waters of each management district were not predictive of strain contributions to mixtures sampled (% stocked columns in Tables 4 and 5). Likewise, stocking combined with age-specific mortality (from statistical catch at age models) and movements based on observations of aged stocked fish (Alderstein et al. 2007; Tables 4 and 5) also were not generally reflective of mixture composition (Tables 4 and 5). The high proportional representation of Seneca Lake lake trout suggests that members of this strain have higher survival, and/or higher fecundity than lake trout of other strains. In contrast, several American and Canadian strains contributed little to naturally produced fish sampled (Tables 4 and 5). Given limitations of hatchery space and pending recommendations to alter stocking prescriptions in Lake Huron, the higher success of Seneca strain lake trout should be considered if stocking continues.

Discussion

Sustainability of economically, ecologically, and culturally important natural populations of Great Lakes fishes requires greater understanding of relationships between recruitment from natural and hatchery sources and dispersal and habitat occupancy by naturally produced recruits. Our research applied methodology for quantifying temporal and strain-specific contributions to mixed stocks of naturally produced offspring produced from hatchery strains occupying open-water areas in the Great Lakes. We identified lake trout hatchery strains that contributed disproportionately to the open-water assessments in American and Canadian management districts and how individual strain contribution varied spatially and temporally. Identification of management districts used by individuals of different ages and strains (populations in general) originating from different management jurisdictions or stocking locations is a fundamental requisite for lake trout management and recovery in Lake Huron and elsewhere.

Populations of many fish species are spatially genetically structured as a function of rates of straying and due to genetic drift associated with small effective population size (Taylor et al. 2001; Allendorf et al. 2013). In the case of lake trout in Lake Huron, native populations were extirpated in all locations with the exception of Parry Sound and Iroquois Bay (Berst and Spangler 1972; Reid et al. 2001). Therefore,

spatial genetic structure observed in the form of significant differences in allele frequency among management units are due to compositional differences in hatchery strain relative abundance and their respective reproductive contributions, as these strains themselves are all genetically differentiated.

The relative abundance of American strains of lake trout in multiple Canadian management districts that were not stocked with American strains suggests that fish from American strains are straying from waters in which they are stocked and are reproducing in Canadian waters. For example, allele frequencies and estimates of strain contributions to the Canadian the N12 and American MH1 regions were similar. An alternative explanation would be that naturally produced lake trout stray to a greater degree than hatchery lake trout, which is not likely, as straying of hatchery fish can be widespread (Quinn 1993).

Large non-zero estimates of assortative mating coefficients in the early sampling period in American waters and in some of the Canadian management units in the later period suggest that mating among different strains is not always random. Values of the assortative mating coefficient in later periods were generally low and in many cases essentially zero suggesting high levels of strain mixing. Formally a zero value indicates that the genotypes reflect the combinations expected given the proportional contributions of the different strains, if they were combined at random. The model cannot produce more strain mixing than is expected based on random combinations, but a zero estimated value may occur when this is the case. Although assortative mating coefficients generally were characterized by large confidence intervals, the results show general spatial and temporal trends that may have implications for management. One potential explanation for the decrease in the assortative mating coefficient between the early and later period is that a substantial proportion of naturally produced individuals sampled in the late period may not be F1 (first generation naturally produced fish) but rather represent offspring from matings among naturally produced parents. These fish would tend to reflect more mixing of strains. An additional contributing factor would be if the more mixed genotypes of fish of higher filial generation survived better. Given that hatchery strains have been under domestication for a number of generations (e.g., Page et al. 2003), some level of inbreeding depression may exist. Matings among members of the

same strain (or even between two pure strains) may have not been as successful (in terms of relative reproductive success) than individuals produced from outbreed matings among members of different strains. Heterosis or hybrid vigor has been commonly observed in situations where populations (or domestic stocks) have existed in low numbers and have experienced some levels of inbreeding depression (Lynch 1991). Indeed, naturally produced lake trout across most spatial regions were more genetically variable than were the progenitor hatchery strains (comparisons for H_E, A_R, and F_{is} in Supplemental Tables 4 and 6).

Non-zero and varying levels of assortative mating that indicate the interbreeding of individuals of different strains was not surprising and was previously incorporated in genetic stock identification analyses for lake trout in the Great Lakes (Marsden et al. 1989). Marsden et al. (1989) constructed artificial baseline hatchery strains that were hybrids of existing strains to use as unique 'baselines'. Here we take a model-based approach (equation 3) and estimate the proportion of individuals mating assortatively and randomly as parameters in the overall Bayesian mixture model. The levels of mixing among strains we observed based on the estimated assortative mating coefficient indicate that caution should be used when estimating proportional contributions of strains to mixtures using traditional likelihood approaches (e.g., Pella and Milner 1987; Pella and Masuda 2001). Likewise, considerable attention has been focused on use of individual assignment tests for purposes of mixture analysis (e.g., Manel et al. 2005). For the same reason, caution is advised when assigning individuals to strain of origin given evidence for inter-breeding between strains (Guinand et al. 2004).

While the emergence of naturally produced lake trout in Lake Huron is promising, restoration is in an early stage. In comparison with Lake Superior where the lake experienced a successful transition from a hatchery stocked population to a naturally produced fish dominated population, the current spawning biomass in Lake Huron is still low (12-15 versus 70 adult per km gillnet per night in Lakes Huron and Lake Superior, respectively; He et al. 2012). To maintain sufficient top-down influence on a dynamically changing food web, and to ensure the success of natural reproduction and recruitment, adult density of the top predator like lake trout should be higher than a minimum level (Walters and Kitchell 2001).

Discovering a relationship between strain type and successful reproduction can enhance managers' understanding of adaptive mechanisms and contribute to the development of more efficient and effective rehabilitation strategies across the Great Lakes. We utilized a general model developed by Gaggiotti et al. (2002; 2004) and Guo et al. (2008) and a Bayesian approach for estimating the proportional contribution of source populations or strains to newly founded colonies, as a form of genetic stock identification (GSI). Further, we expanded on the Guo et al. (2008) approach to evaluate model fit to investigate whether there was evidence for temporal and spatial variation in strain contributions (DIC analyses reported in Table 3). We also accounted for inter-strain mixing in our models. This method could be combined with previously developed modeling approaches (Tsehaye et al. 2016; Brenden et al. 2018) to characterize age (or cohort) specific differences in recruitment by strain and to more rigorously model effects of strain-specific stocking numbers, survival, and movements to estimated strain contributions. Although we applied the approach to data for Lake Huron lake trout, the methodology could be readily adaptable to any other species for which appropriate data exist or can be obtained. Improved knowledge of stock contribution and recruitment will allow for more effective management of other native fishes.

Acknowledgments

Funding for this project was provided by the Great Lakes Fishery Commission. Additional funding was provided by supporting partners of the Quantitative Fisheries Center, which includes Michigan State University, Michigan Department of Natural Resources, Great Lakes Fishery Commission, and Council of Lake Committee fishery agencies. Samples of naturally produced lake trout and aging data was provided by the Michigan Department of Natural Resources (MDNR), Chippewa Ottawa Resource Authority (CORA), the U.S. fish and Wildlife Service (USFWS), Ontario Ministry of Natural Resources and Forestry (OMNRF) and the U.S. Geological Survey (USGS). Use of trade names does not signify endorsement by the US government. This work was authorized by the Director of the U.S. Geological

Survey. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. All sampling and handling of fish during research are carried out in accordance with guidelines for the care and use of fishes by the American Fisheries Society (http://fisheries.org/docs/wp/Guidelines-for-Use of-Fishes.pdf). This is contribution 20XX-XX of the Quantitative Fisheries Center at Michigan State University.

Data Availability

Standardized data from this project will be established at (https://www.sciencebase.gov/catalog/item/5540f811e4b0a658d793a535).

References

Adlerstein SA, Rutherford ES, Clevenger, Johnson JE, Clapp DF, Wodt P. 2007. Lake trout movements in U.S. waters of Lake Huron interpreted from coded wire tag recoveries in recreational fisheries. *J Great Lakes Res*. 33:186-201

Allendorf FW, Luikart G, Aitken SN. 2013. *Conservation and the Genetics of Populations*. 2nd Ed. Wiley-Blackwell, Oxford, UK.

Anderson EC, Waples RS, Kalinowski ST. 2007. An improved method for predicting the accuracy of genetic stock identification. *Can J Fish Aquat Sci.* 65:1475-1486.

Angers B, Bernatchez L, Angers A, Desgroseillers L. 1995. Specific microsatellite loci for brook charr reveal strong population subdivision on a microgeographic scale. *J Fish Biol*. 47A:177-185.

Beacham TD, Wallace CG, Le CD,Beere M. 2012. Population structure and run timing of Steelhead in the Skeena River, British Columbia. *N Amer J Fish Manage.* 32:262-275.

Berst AH, Spangler GR. 1972. Lake Huron: effects of exploitation, introductions, and eutrophication on the salmonid community. *J Fish Res Board Can.* 29:877-887.

Brenden TO, Tsehaye I, Bence JR, Liu W, Scribner KT, Kanefsky J. 2018. Indexing recruitment fluctuations for populations contributing to mixture fisheries by simultaneous analysis of age and genetic information. *Can J Fish Aquat Sci*. doi.org/10.1139/cjfas-2016-0525 .

DeHaan PW, Ardren WR. 2005. Characterization of 20 highly variable tetranucleotide microsatellite loci for bull trout (*Salvelinus confluentus*) and cross-amplification in other *Salvelinus* species. *Mol Ecol Notes.* 5:582-585.

Hansen, M. J. 1999. Lake trout in the Great Lakes: basin wide stock collapse and binational restoration. Pages 417–453 in W. W. Taylor and C. P. Ferreri, editors. *Great Lakes fisheries and policy management: a binational perspective*. Michigan State University Press, East Lansing.

Fitzsimons JD, Brown S, Brown L, Honeyfield D, He J, Johnson JE. 2010. Increase in lake trout reproduction in Lake Huron following the collapse of alewife: Relief from thiamine deficiency or larval predation? *Aquat Ecosyst Health Manage.* 13:73–84.

Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder MN, Nielsen A, Sibert J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimiz Meth Software.* 27:233-249.

Gaggiotti OE, Jones F, Lee WM, Amos W, Harwood J, Nichols RA. 2002. Patterns of colonization in a metapopulation of grey seals. *Nature.* 416: 424-427.

Gaggiotti OE, Brooks SP, Amos W, Harwood J. 2004. Combining demographic, environmental, and genetic data to test hypotheses about colonization events in metapopulations. *Mol Ecol.* 13:811-825.

Geweke J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: JM, Berger JO, Dawid AP, Smith AFM, editors. *Bayesian Statistics 4*. Oxford University Press, Oxford, UK. pp. 169–193.

Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). <http://www.unil.ch/izea/softwares/fstat.html> (May 2 2005).

Guinand, B, Scribner KT, Topchy A, Page KS, Punch W, Burnham-Curtis MK. 2004. Sampling issues affecting accuracy of likelihood-based classification. *Envir Biol Fish*. 69:245-259.

Guo F, Dey DK,Holsinger KE. 2008. A hierarchical Bayesian approach for estimating the origin of a mixed population. In: B. Clarke and S. Ghosal, editors. *Pushing the limits of contemporary statistics: contributions in honor of Jayanta*. K. Ghosh, Institute of Mathematical Statistics, Beachwood, OH. p 237-250.

Hansen MJ. 1999. Lake trout in the Great Lakes: basinwide stock collapse and binational restoration. In: Taylor WW, Ferreri CP editors. *Great Lakes Fishery Policy and Management: a Binational Perspective*. Michigan State University Press, East Lansing, Michigan. p 417–453.

He J, Ebener M, Riley S, Cottrill A, Kowalski A, Poprosku S, Mohr L, Johnson J. 2012. Lake trout status in the main basin of Lake Huron, 1973-1010. *N Amer J Fish Manage.* 32:402-412.

Johnson JE, He J, Fielder DG. 2015. Rehabilitation of stocking of walleyes and lake trout: restoration of reproducing stocks in Michigan waters of Lake Huron. *Trans Amer Fish Soc.* 77:396-408.

Kalinowski ST, Manlove KR, Taper ML. 2007. ONCOR: a computer program for genetic stock identification. Montana State University, Bozeman.

King TL, Lubinski BA, Burnham-Curtis MK, Stott W. Morgan II RP. 2012. Tools for the management and conservation of genetic diversity in brook trout (*Salvelinus fontinalis*): tri-and tetranucleotide microsatellite markers for the assessment of genetic diversity, phylogeography, and historical demographics. *Con Gen Res*. 4:539-543.

Lynch M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution.* 45:622-629.

Madenjian CP, Desourcie TJ, McClain JR, Woldt AP, Holuszko JD, Bowen CA II. 2004. Status of lake trout rehabilitation on six fathom bank and Yankee reef in Lake Huron. *N Amer J Fish Manage.* 24:1003-1016.

Manel S, Gaggiotti OJ, Waples RS. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trend Ecol Evol.* 20:136-142.

Marsden JE, Krueger CC, May B. 1989. Identification of parental origins of naturally produced lake trout in Lake Ontario: application of mixed stock analysis to a second generation. N Amer J Fish Manage. 9:257-268.

Muir AM, Krueger CC, Hansen J. 2013. Re-establishing lake trout in the Laurentian Great Lakes: past, present, and future. :In Taylor WW, Lynch AJ, Leonard NJ, editors. *Great Lakes Fishery Policy and Management: a Binational Perspective, 2nd edition*. Michigan State University Press, East Lansing, Michigan. p. 533–588.

Olsen JB, Bentzen P, Seeb JE. 1998. Characterization of seven microsatellite loci derived from pink salmon. *Mol Ecol.* 7:1087-1089.

O'Reilly PT, Hamilton LC, McConnell SK, Wright JW. 1996. Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Can J Fish Aquat Sci*. 53:2292-2298.

Pella JJ, Milner GB. 1987. Use of genetic marks in stock composition analysis. In: Ryman N, Utter F, editors. *Population genetics and fishery management*. Washington Sea Grant Program, Seattle. p 247-276.

Pella J, Masuda M. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. *Fish Bull.* 99:151-167.

Plummer M, Best N, Cowles K, Vines K. 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6:7-11.

Pulliam H.R. 1988. Sources, sinks, and population regulation. *Amer Nat*. 132:652-661.

Quinn TP. 1993. A review of homing and straying of wild and hatchery-produced salmon. *Fish Res*. 18:29-44.

R Core Team 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceed Nat Acad Sci*. 94:9197–9201.

Reid DM, Anderson DM, Henderson BA. 2001. Restoration of lake trout in Parry Sound, Lake Huron. *N Amer J Fish Manage.* 21:156-169.

Riley SC, Adams JV. 2010. Long-term trends in habitat use of offshore demersal fishes in western Lake Huron suggest large-scale ecosystem change. Trans Amer Fish Soc. 139:1322-1334.

Riley SC, He JX, Johnson JE, O'Brien TP, Schaeffer JS. 2007. Evidence of widespread natural reproduction by lake trout *Salvelinus namaycush* in the Michigan waters of Lake Huron. *J Great Lakes Res*. 33:917-921.

Riley SC, Rinchard J, Honeyfield DC, Evans AN, Begnoche L. 2012. Increasing thiamine concentrations in lake trout eggs from lakes Huron and Michigan coincide with low alewife abundance. *N Amer J Fish Manage*. 32:1052-1064.

Rollins MF, Vu NV, Spies IB, Kalinowski ST. 2009. Twelve microsatellite loci for lake trout (*Salvelinus namaycush*). *Mol Ecol Res.* 9:871-873.

Schram ST, Selgeb JH, Bronte CR, Swanson BL. 1995. Population recovery and natural recruitment of lake trout at Gull Island shoal, Lake Superior, 1964–1992. *J Great Lakes Res.* 21*(*Suppl. 1*):*225–232.

Scribner KT, Gust JR, Fields RL. 1996. Isolation and characterization of novel salmon microsatellite loci: cross-species amplification and population genetics applications. *Can J Fish Aquat Sci*. 53:833-841.

Seeb, LW, Crane PA. 1999. Allozymes and mitochondrial DNA discriminate Asian and North American populations of Chum Salmon in mixed stock fisheries along the south coast of the Alaska peninsula. *Trans Amer Fish Soc.* 128:88-103.

Spiegelhalter DJ, Sherlaw-Johnson C, Bardsley M, Blunt I, Wood C, Grigg O. 2012. Statistical methods for healthcare regulation: rating, screening and surveillance (with discussion). *J Roy Stat Soc Series A*. 175:1-25.

Taylor EB, Redenbach Z, Costello AB, Pollard SM, Pacas CJ. 2001. Nested analysis of genetic diversity in northwestern North American char, Dolly Varden (*Salvelinus malma*), and bull trout (*Salvelinus confluentus*). *Can J Fish Aquat Sci.* 53:406-420.

Tsehaye, I, Brenden TO,Bence JR, Liu W, Scribner KT, Kanefsky J, Bott K,Elliott RF. 2016. Combining genetics with age/length data to estimate consistent changes in year-class strength of spawning populations contributing to admixtures, with application to Lake Michigan lake sturgeon. Fish Res. 173:236-249.

Walters C, Kitchell JF. 2001. Cultivation/depensation effects on juvenile survival and recruitment: implications for the theory of fishing. *Can J Fish Aquat Sci.* 58:39-50.

Weir BS, Cockerham CC. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution.* 38:1358–1370.

Zimmerman M, Krueger C. 2009. An ecosystem perspective on re-establishing native deepwater fishes in the Laurentian Great Lakes. *N Amer J Fish Manage.* 29:1352-1371.

Figure Legend

Figure 1. Relative frequency of age classes of wild lake trout represented in USA and Canadian management districts.

Figure 2. Map showing the Lake Huron lake trout management districts.