The genetic composition of wild recruits in a recovering lake trout population in Lake Michigan

Wesley A. Larson[1,*], Matthew S. Kornis[2], Keith N. Turnquist[3], Charles R. Bronte[2], Mark E. Holey[2], S. Dale Hanson[2], Theodore J. Treska[2], Wendylee Stott[4], Brian L. Sloss[3]

[1]U.S. Geological Survey, Wisconsin Cooperative Fishery Research Unit, College of Natural Resources, University of Wisconsin-Stevens Point, 800 Reserve St., Stevens Point, WI, USA, 54481 (email: wes.larson@uwsp.edu)

[2]U. S. Fish and Wildlife Service, Green Bay Fish and Wildlife Conservation Office, 2661 Scott Tower Drive, New Franken, Wisconsin, USA, 54229 (email: matthew_kornis@fws.gov, charles_bronte@fws.gov, markeholey@gmail.com, dale_hanson@fws.gov, ted_treska@fws.gov)

[3]College of Natural Resources, University of Wisconsin-Stevens Point, 800 Reserve St., Stevens Point, WI, USA, 54481 (email: kturnqui@uwsp.edu, bsloss@uwsp.edu)

[4]Michigan State University CESU working for U.S. Geological Survey, Great Lakes Science Center, 1451 Green Road, Ann Arbor, MI 48105, United States of America (email: wstott@contractor.usgs.gov)

[*]**Corresponding author:** Wesley Larson (e-mail: wes.larson@uwsp.edu, phone: 1-715-346-3150)

Key words: lake trout, hybridization, recruitment, restoration, propagation, genetic strain

Abstract

Strain performance evaluations are vital for developing successful fishery management and restoration strategies. Here, we utilized genotypes from 36 microsatellites to investigate hatchery strain contribution to collections of naturally produced lake trout sampled across Lake Michigan. Strain composition varied by area, with recoveries of Seneca Lake strain exceeding expectations based on stocking records in northern Lake Michigan but performing similarly to other strains in southern Lake Michigan. Interstrain hybrids were present at moderate frequencies similar to expectations based on simulations suggesting that strains are interbreeding randomly. We hypothesize that the superior performance of the Seneca Lake strain in northern Lake Michigan is partially due to adaptive advantages that facilitate increased survival in areas with high mortality from sea lamprey predation such as northern Lake Michigan. However, when this selective pressures is lessened, the Seneca Lake strain performs similarly to other strains. Our study demonstrates that strain performance can vary across small spatial scales and illustrates the importance of conducting thorough strain evaluations to inform management and conservation.

Introduction

Attempting to rebuild collapsed or extirpated fish populations through stocking of hatchery-reared individuals is a common practice, but the short- and long-term success of this approach is highly variable (George et al. 2009). Genetic analysis represents a powerful tool for monitoring the success of stocking programs and can provide important information to inform future propagation practices (Fraser 2008; Miller and Kapuscinski 2003). For example, genetic data can be used to assess strain contributions in systems where more than one strain was stocked (Miller et al. 2009), estimate introgression between wild and hatchery strains (Seamons et al. 2012), and understand potential fitness consequences when multiple strains are crossed (Huff et al. 2011). Genetic studies have provided important insights on the success of propagation programs; however, these types of analyses are still relatively rare compared to the massive inputs of propagated organisms released into the wild (Halverson 2008). Here, we utilize genetic tools to investigate strain-specific recruitment dynamics and introgression in lake trout from Lake Michigan, a heavily propagated species that has shown signs of population recovery.

Lake trout (*Salvelinus namaycush*) recovery efforts in the Laurentian Great Lakes are a prime example of the use of hatchery propagation to rebuild collapsed fisheries and represent one of the largest re-stocking efforts in the world. In Lake Michigan, as in most other Great Lakes, lake trout historically were the apex predator and supported large commercial and recreational fisheries (Holey et al. 1995; Wells and McLain 1973). Predation by invasive sea lamprey (*Petromyzon marinus*) and overfishing largely led to lake trout extirpation in Lake Michigan by the mid-1950s (Hansen 1999; Smith 1968; Wells and McLain 1973). In the late 1950s and 60s, management steps were taken to rebuild populations by the expansion of sea lamprey control, increased fishery regulation, and repopulation from stocking hatchery-reared fish (Heinrich et al.

2003; Holey et al. 1995; Wells and McLain 1973). Nine genetic strains of lake trout have been stocked into Lake Michigan since 1959 (Fig. 1). Most of these strains were derived from the lean (i.e. shallow water) form of lake trout and were captured in Lake Superior (Marquette, Apostle Islands, and Isle Royale) or Lake Huron (Parry/Big Sound). However, two strains (Green Lake and Lewis Lake) were derived from lean lake trout that were captured in Lake Michigan and stocked in smaller inland lakes (Jenny Lake/Lewis Lake, WY; Green Lake, WI) before lake trout were extirpated from Lake Michigan. Descendants of these fish were then captured from these smaller inland lakes to produce hatchery brood stock from remnant Lake Michigan genetic strains (Krueger et al. 1983; Page et al. 2004). Supplementation of the Green Lake broodstock in the 1980s also involved capturing feral fish uniquely fin-clipped from southern Lake Michigan as a gamete source (Kincaid et al. 1993). Additionally, the Seneca Lake strain was derived from lean lake trout native to Seneca Lake, NY, a finger lake in the Lake Ontario drainage, and the Klondike Reef strain was derived from the humper (i.e. deep water) ecomorphotype found in Lake Superior (see Muir et al. 2014 for context on lake trout ecomorphotypes). The strain composition of lake trout stocked into Lake Michigan has varied substantially over the last half century. From 1962-1990, most lake trout stocked in Lake Michigan were derived from the Marquette strain. Stocking in the 1990s and early 2000s consisted of primarily Lewis Lake and Apostle Islands strains. More recently, stocking has consisted primarily of Seneca Lake, Lewis Lake, Parry Sound, and Klondike strains (Kornis et al. 2019b) (Fig. 1).

Early stocking efforts were able to rebuild populations, as hatchery-reared fish survived well and aggregated for spawning, but there was little evidence of natural reproduction (Holey et al. 1995). In Lake Michigan this lack of successful natural reproduction was likely due to a number of factors acting simultaneously but mostly centered around insufficient numbers of

93 stocked fish and resulting spawning stock, inappropriate stocking locations, and poor survival of
94 eggs and fry (Bronte et al. 2003; Hansen 1999). Modifications of stocking practices informed by
95 these early efforts were made in the 1980s and included stocking fish on historically important
96 offshore reefs instead of only nearshore sites, building up overall population densities, and
97 introducing additional strains to promote genetic and ecological diversity (Bronte et al. 2008;
98 Dexter et al. 2011; Holey et al. 1995). These modifications established spawning aggregations at
99 some key sites (Bronte et al. 2007). However, natural reproduction of stocked fish remained rare,
100 potentially due to predation on lake trout fry by invasive alewife (*Alosa pseudoharengus*)
101 (Krueger et al. 2014; Krueger et al. 1995; Madenjian et al. 2008) and/or from thiamine
102 deficiency attributed to alewife consumed by adult lake trout that results in poor survival of eggs
103 and larvae (Brown et al. 2005). High mortality of adult lake trout from sea lamprey and/or
104 fishing has also limited establishment of self-sustaining populations in some areas (Bronte et al.
105 2007; Kornis et al. 2019b). Since 2002, abundances of alewife in Lake Michigan have steadily
106 decreased (Madenjian et al. 2018), and Hanson et al. (2013) reported that by 2005 stocked lake
107 trout had begun to successfully reproduce in Lake Michigan, with wild lake trout comprising an
108 average of 30% of the sport fishery by 2018 driven by wild recruitment in southern Lake
109 Michigan (Kornis et al. 2019a), and similar proportions in fishery independent surveys
110 (LMLTWG 2019).

111 Understanding the relative contributions of different genetic strains to emerging natural
112 reproduction of lake trout in Lake Michigan is vital for shaping future management and
113 restoration efforts. Fortunately, strains from Lake Superior, Lake Huron, Lake Michigan, and
114 Seneca Lake are genetically distinct, making genetic assignment of wild fish of unknown genetic
115 origin feasible (Page et al. 2004). There is also evidence of differential survival among strains,

which may be related to adaptive differences (McKee et al. 2004; Rogers et al. 2019; Scribner et al. 2018). For example, the Seneca Lake strain, which co-evolved with sea lamprey, appears to exhibit lower rates of sea lamprey-induced mortality compared to other strains (Bergstedt et al. 2003; Bronte et al. 2007; Schneider et al. 1996), while strains of Lake Michigan origin (Lewis Lake and Green Lake) appeared to have higher survival relative to Seneca Lake strain in southern Lake Michigan with low abundance of sea lamprey (Kornis et al. 2019b).

The overall goal of our study was to use genetic techniques to investigate the strain composition of wild lake trout produced in Lake Michigan. Our specific objectives were to: 1) identify a subset of microsatellite markers that reliably delineate strain of origin for the lake trout stocked into Lake Michigan; 2) determine if the genetic markers and reference data are capable of diagnosing the genetic heritage of interstrain crosses and wild caught lake trout; (3) analyze naturally reproduced lake trout of unknown origin captured in Lake Michigan to determine the relative contribution of different genetic strains; and (4) compare observed strain contributions with a composite estimate of expected strain compositions reconstructed from stocking records, age, survival, movement, and fecundity data to investigate differential reproductive success among strains and (5) determine if interstrain mating was random.

Materials and methods

Sample collection

We obtained lake trout tissue samples from eight hatchery strains (hereafter referred to as reference strains) and from 847 naturally produced lake trout captured throughout Lake Michigan (Tables 1,2). Tissue samples were fin clips preserved in > 95% ethanol or dried in scale envelopes. Hatchery reference strain samples were collected by the U.S. Fish and Wildlife Service from 1999-2015 (Table 1). The genetic structure of seven of the eight reference strains

139 included in the current study was evaluated by Page et al. (2004); the only strain not analyzed in
140 Page et al. (2004) was Klondike Reef humper, which was not stocked until 2010. No samples
141 from the ninth strain stocked in Lake Michigan, Clearwater Lake, were obtained as this stock
142 was used infrequently and only before 1980. Naturally produced lake trout (Table 2) were
143 collected from surveys conducted by the U.S. Fish and Wildlife Service, Chippewa Ottawa
144 Resource Authority, Little Traverse Bay Band of Odawa Indians, Grand Traverse Band of
145 Ottawa and Chippewa Indians, and the Wisconsin, Michigan, Illinois, and Indiana Departments
146 of Natural Resources. Additional specimens were collected from the sport fishery by the U.S.
147 Fish and Wildlife Services' coded-wire tagging and recovery program (Bronte et al. 2012). Lake
148 trout caught in these surveys were classified as naturally produced if they did not have a coded
149 wire tag or visible fin clip, as all hatchery-reared lake trout were marked/tagged prior to
150 stocking. This classification is supported by work showing that the vast majority of Lake
151 Michigan lake trout that lack a fin clip or coded wire tag are indeed of wild origin, as opposed to
152 erroneously unclipped hatchery fish or migrants from Lake Huron (Landsman et al. 2017).
153 Metadata collected on naturally produced lake trout included date and location of capture, total
154 length (mm), weight (g), sex, maturity status, and age as estimated from otolith cross sections as
155 described by Campana et al. (2008). The vast majority of naturally produced fish had survived
156 early juvenile life-stages and > 95% of fish longer than 300 mm.
157 Naturally produced fish included in this study were sampled from across Lake Michigan,
158 with the most fish sampled in the north and southwest areas of the lake. We constructed seven
159 geographic strata based on available samples to facilitate spatial estimates of strain contributions
160 (Table 2). These strata were drawn based on management district boundaries, with the exception
161 of the Southern Refuge stratum, which was included as its own stratum due to high sampling

effort in the area and evidence that this area may be the source for many of the naturally reproduced lake trout in Lake Michigan (Kornis et al. 2019b; Patterson et al. 2016). It is important to note that numbers of fish captured in each stratum are a function of sampling effort and do not necessarily correlate with lake trout abundances in these strata. We did not conduct temporal analyses across years or seasons as sample sizes were insufficient for these comparisons.

Laboratory analysis and quality control

Microsatellite genotyping of lake trout was conducted using methods similar to Ruzich et al. (2019). DNA was extracted with the Promega Wizard® Genomic DNA purification kit (Promega Corp., Madison, Wisconsin), following a 96-well configuration. Purified genomic DNA was quantified using a Nanodrop® ND-100 spectrophotometer (Nanodrop Technologies, Wilmington, Delaware), and normalized to a final concentration of 20 ng/µL. PCR amplification was conducted to amplify 49 microsatellite DNA loci developed for lake trout as well as other salmonids (see Table S1 for multiplex information). An ABI 3730 DNA Analyzer (Life Technologies, Carlsbad, California) was used to separate PCR amplicons and determine allele sizes. Allele sizes were visually verified using Genemapper® software V4 (Life Technologies, Carlsbad, California) and allele calls were collated into a collection of multi-locus data for each individual. Approximately 15% fish were double scored as a quality control check.

After genotyping, we removed loci with > 30% missing data, loci that produced peaks that were difficult to score due to stutter banding, loci that were monomorphic, and loci that deviated from Hardy-Weinberg or linkage disequilibrium in more than half of the eight reference strains. Deviations from Hardy-Weinberg or linkage disequilibrium were assessed with exact tests conducted in Genepop V4 (Rousset 2008) (alpha = 0.05). We also removed individuals with

185 > 50% missing data and individuals that showed evidence of contamination (> 2 alleles at > 2
186 loci).
187 *Statistical analysis of reference strains*
188 We calculated basic summary statistics for the reference strain dataset to investigate
189 variation in genetic diversity across loci and strains. Observed and expected heterozygosity (H_O,
190 H_E) were calculated in GenAlEx (Peakall and Smouse 2012), number of alleles (A) and allelic
191 richness (A_R) were calculated in FSTAT (Goudet 1995), and locus-specific F_{ST} and F_{IS} (Weir and
192 Cockerham 1984) were calculated in Genepop. We also used Genepop to calculate pairwise F_{ST}
193 values for each population pair. Finally, we constructed a neighbor-joining dendrogram based on
194 Nei's D_A distance (Nei et al. 1983) in POPTREE2 (Takezaki et al. 2010) to visualize genetic
195 relationships among populations. Support for each node was assessed with 1,000 bootstrap
196 replicates.
197 We used the Bayesian MCMC approach implemented in STRUCTURE 2.3.4 (Pritchard
198 et al. 2000) to infer the number of major genetic clusters in our data. This program groups
199 individuals into K genetic clusters by minimizing overall deviations from Hardy-Weinberg and
200 linkage equilibrium within clusters. STRUCTURE was run on K-values from 1-10, ten runs were
201 conducted for each K-value, and each run consisted of a burn-in period of 10,000 iterations
202 followed by 100,000 iterations. We selected the "admixture model" and "use location as prior"
203 options in STRUCTURE as suggested by Hubisz et al. (2009), who found that the "use location
204 as prior" option facilitated detection of population structure at lower levels of divergence. The
205 most likely value of K was evaluated based on likelihood values as well as with the ΔK method
206 (Evanno et al. 2005), and the results were summarized with Structure Harvester (Earl and

vonHoldt 2012). We also used the program CLUMPAK (Kopelman et al. 2015) to assess convergence among runs and generate consensus data for each K-value.

Assignment of unknown individuals to strain of origin

A primary objective of the current study was to determine if our reference dataset had sufficient power for assigning individuals of unknown origin to their natal genetic strain and classifying individuals as either pure or interstrain crosses (i.e., hybrids). To assess our assignment power, we simulated 1,000 individuals from all possible pure and F1 crosses of our reference strains with the program HYBRIDLAB (Nielsen et al. 2006). For this analysis, we combined the Isle Royale, Apostle Islands, and Marquette collections into a single Lake Superior strain because they are genetically similar (see Results). We then assessed our ability to assign simulated individuals to their strain of origin by calculating posterior probabilities of assignment in STRUCTURE. STRUCTURE was run at K=6 (see Results) with the reference strains classified as known individuals (POPFLAG=1) and the simulated individuals classified as unknowns (POPFLAG=0), and run parameters were the same as above. HYBRIDLAB does not simulate missing data, but < 3% of genotypes were missing from individuals on average indicating that simulated results should be comparable to results from empirical data.

After exploratory analysis using a range of cutoffs and comparing results to simulated (i.e. known proportion) samples, we determined that the highest assignment accuracies were achieved with a cutoff of 0.7. That is, if an individual had an assignment probability (i.e. q-value) > 0.7 for a given strain, that individual was classified as pure, and if an individual had an assignment probability < 0.7, it was classified as a hybrid. All individuals with assignment probabilities < 0.7 were classified as F1 hybrids between the top two contributing strains. A probability cutoff of 0.7 was shown to produce high assignment accuracy for F1 hybrids between

populations with similar F_{ST} values as those in our study and using a similar numbers of markers as we genotyped (Vaha and Primmer 2006) providing further evidence that our approach should be robust.

We then compared the mixture estimates for simulated crosses derived from STRUCTURE to mixture estimates derived from the program ONCOR (http://www.montana.edu/kalinowski/Software.htm). ONCOR outperformed STRUCTURE for pure crosses, but substantially underperformed STRUCTURE for hybrid crosses (see results). We proceeded with assignment using STRUCTURE since a high proportion of the fish in our study were classified as putative hybrids, and high error assigning hybrid crosses would lead to high error in our mixture estimates. This large discrepancy between STRUCTURE and ONCOR indicates that the mixture calculation algorithm employed in ONCOR and similar programs may not be well suited for estimating mixture proportions when hybrids are present. Unfortunately, the STRUCTURE approach does not facilitate the estimation of confidence intervals. However, we do discuss how strain-specific error rates may influence estimates in the results section.

After conducting assignment tests and choosing an approach for mixture analysis, we assigned naturally produced individuals to their strain of origin using the same methods. It is important to note that since natural reproduction of lake trout has likely been occurring for < 1 generation in Lake Michigan (Hanson et al. 2013), we assumed all hybrids in the population are F1 hybrids for this analysis. If advanced generation hybrids were present, they may be assigned as pure, but we believe this is highly unlikely as there were only four fish in our dataset over the age of 20 and these could have represented individuals where fin clips to denote hatchery origin had regrown. To determine the observed proportional contribution for each strain among the samples collected, we needed to account for both pure and hybrid contributions. We therefore

253 assigned hybrids 50% of the weight of pure individuals. For example, for a mixture containing
254 30 pure Seneca Lake fish, 20 Seneca Lake x Lewis Lake hybrids, and 10 pure Lewis Lake fish,
255 the proportion of Seneca ancestry is calculated as (30+(0.5*20))/60 = 66%.
256 We derived 95% confidence intervals for proportional estimates using the prop.test
257 function in R and the estimated assignment accuracy for each strain derived from simulated
258 individuals. The prop.test function produces a confidence interval that incorporates sample size
259 but does not take into account assignment accuracy for each strain. Therefore, we expanded these
260 confidence intervals based on assignment accuracy. For example, if the original confidence
261 interval calculated with prop.test was 0.5-0.6 and the assignment accuracy to a given strain was
262 0.9, the adjusted confidence interval would be 0.45-0.66. This approach is conservative and takes
263 into account both sampling error and assignment accuracy. For estimates that included both pure
264 and hybrid crosses, assignment accuracy was the average assignment accuracy of all crosses
265 included in the estimate.
266 *Calculating expected strain and cross contributions*
267 Deriving the expected strain proportions based on stocking levels is vital for investigating
268 strain-specific differences in survival and reproductive success. We sought to compare expected
269 proportions with empirical observations to determine strain contributions to natural reproduction
270 relative to stocking levels. Our general approach for deriving strata-specific expected strain
271 proportions was to use a composite method to reconstruct the strain proportions of the spawners
272 that produced the lake trout that we sampled. This analysis required data on stocking rates, age
273 composition, fecundity, and movement patterns. In general, the expected proportions of each
274 strain were relatively similar to the stocking rates of that strain during the brood year the parents
275 of most naturally produced fish were produced (generally the early 2000s). However, other

276 factors such as movement and age-specific fecundity make it inappropriate to simply use the
277 stocking and age data alone. It is also important to note that these expected proportions do not
278 incorporate any direct measures of survival of stocked strains or recruitment across strains or
279 strata, as data on these metrics is sparse. As such, they do not reflect the actual strain
280 composition of spawners, but rather our best estimate of what the strain composition should have
281 been given stocking records and other metrics. Differences in survival or recruitment among
282 strains and strata are not incorporated into these estimates thus any differences in observed and
283 expected strain proportions of naturally produced fish may be a result of these two metrics,
284 among others (see discussion).

285 To calculate estimates of expected strain proportions, we first derived the cohort (i.e.,
286 year class) of all naturally produced lake trout in our study using the year of capture minus the
287 otolith age, and then calculated cohort proportions for each stratum. We then used spawner
288 survey data collected according to Schneeberger et al. (1998) to calculate the age composition of
289 spawners for each cohort and stratum. Only ages 6-20 were included in this analysis because six
290 is the age when most lake trout reach sexual maturity in Lake Michigan (Madenjian et al. 1998)
291 and very few spawners were older than 20. Spawner age compositions were corrected with a
292 fecundity multiplier (Table S2) because older fish (at least up to 20 years old) are expected to
293 produce more offspring than younger fish (e.g., Peck 1988). Fecundity multipliers were
294 calculated from predicted mean total length of lake trout at each age in each strata using the
295 relationship between fecundity and total length presented by Fitzsimons and O'Gorman (1996)
296 for lake trout from Lake Ontario. Predicted mean lengths-at-age were determined from von
297 Bertalanffy growth models fit to age/length data from spawner surveys, and fecundity multipliers
298 were expressed relative to expected fecundity at the minimum age of maturity (age 8 for

299 Southern Refuge, age 6 for all other locations) (See Supplementary File 1 for additional details).
300 Next, we obtained stocking proportions by year and stratum from available stocking data
301 (USFWS and GLFC 2017). We then weighted these stocking proportions based on the cohort
302 composition of the naturally produced fish and the age composition and estimated relative
303 fecundity of spawners in each stratum to obtain a stratum-specific expected stocking proportion.
304 Finally, we corrected for movement of lake trout spawners among strata by multiplying stocking
305 proportions by a movement matrix (Table S3). Movement data were derived from coded-wire
306 tagged lake trout, which provided data on both stocking and recovery location (M. Kornis and C.
307 Bronte, unpublished data). Movement was expressed as the percentage of lake trout spawners
308 recovered in each stratum that were expected to originate from other strata. See Supplementary
309 File 2 for details.
310 We also calculated the expected proportion of each cross type (i.e. pure versus hybrids)
311 under random mating to investigate whether hybrids or certain cross types were over or
312 underrepresented in samples of naturally produced fish. Expected proportions of each cross type
313 in each strata were calculated using either observed strain proportions based on genetic data or
314 expected strain proportions based on stocking data. The hypothetical spawning population used
315 to estimate these proportions was assumed to contain only pure strain individuals, a reasonable
316 assumption given that meaningful natural reproduction of lake trout in Lake Michigan is a
317 relatively recent development. Proportions were calculated as follows: the proportion of pure
318 strain individuals was calculated as the strain of those individuals in the current generation
319 squared and the proportion of interstrain hybrids was calculated as two times the proportion of
320 strain one times the proportion of strain two. For example, if the proportion of the Seneca strain
321 in a given strata was 0.7 and the proportion of the Green Lake strain was 0.3 the expected

322 proportion of pure Seneca strain fish is $0.7^2 = 0.49$ and the proportion of Seneca x Green Lake
323 hybrids is 2 * 0.7 * 0.3 = 0.42.

324 # Results

325 *Laboratory analysis and quality control*

326 We attempted to genotype 49 microsatellite DNA loci and retained 36 that produced high
327 quality data (Table S4). Of the 49 original loci, we removed four that were monomorphic in our
328 dataset, three that had > 30% missing data, two that were difficult to score (i.e. produced
329 ambiguous allele peaks), two that were out of Hardy-Weinberg equilibrium, and two that were in
330 linkage disequilibrium with another locus (Table S4). The loci that we retained had an average of
331 13 alleles in our dataset (range: 2 to 38), an average H_O of 0.56 (range 0.00 to 0.93), and an
332 average F_{ST} of 0.04 across strain comparisons (range 0.00 to 0.18). We genotyped 1,485
333 individuals at the 36 loci that we retained and removed 34 individuals with > 50% missing data,
334 42 putatively contaminated individuals, and one individual that was likely a migrant from Lake
335 Huron (see below), resulting in a final dataset of 1,409 individuals (genotype data in Table S5).
336 *Reference strain analysis*

337 Genetic structure among reference strains was relatively high, with pairwise F_{ST} values
338 averaging 0.04 (Table 3). In general, population structure was partitioned by lake of origin, with
339 the largest differentiation in the dataset observed between Seneca Lake and all other populations
340 (Fig. 2, Table 3). Three of the four populations in Lake Superior (Marquette, Apostle Islands,
341 Isle Royale) were genetically similar ($F_{ST} < 0.01$), and the Klondike Reef strain also grouped
342 with these populations according to the neighbor-joining dendrogram but was more divergent
343 (Fig. 2). Green Lake and Lewis Lake, which were both derived from Lake Michigan, displayed
344 an F_{ST} of 0.04, indicating that they are diverged even though they are derived from a similar

345 geographic area. Genetic diversity was relatively similar among strains, with H_O ranging from
346 0.55 to 0.57. However, the three lean populations from Lake Superior did contain slightly more
347 alleles (average A = 9.79 and effective A = 5.29 for Superior leans, 7.61 and 4.27 respectively for
348 the rest of the dataset).

349 Results from STRUCTURE analysis were largely congruent with patterns of population
350 structure suggested by F_{ST} values and the neighbor-joining dendrogram (Fig. 3). However, we
351 were able to observe some potential population admixture that was not apparent with the other
352 analyses. Analysis of multiple K-values revealed that the largest ΔK occurred at K=2 followed
353 by K=4 and K=6 (Table S6). However, the likelihood of each K increased substantially until
354 K=6, where it began to plateau (Table S6). K=6 displayed low variance among runs and was also
355 the first K where the Parry Sound population formed its own cluster. For these reasons, we chose
356 to visualize our data at K=6 and use this K for assignment of unknown individuals (see below).
357 The genetic clusters at K=6 generally correspond to Seneca Lake, Parry Sound, Lewis Lake,
358 Green Lake, Klondike Reef and Lake Superior lean (includes Marquette, Apostle Islands, and
359 Isle Royale) (Fig. 2). The Seneca Lake, Parry Sound, Lewis Lake, Green Lake, Apostle Islands
360 and Isle Royale reference strains displayed little admixture, with the vast majority of individuals
361 in each strain appearing pure. However, we did observe substantial admixture in the Klondike
362 Reef strain, where ~20% of individuals were genetically similar to the Lake Superior lean cluster
363 and 80% formed a genetically distinct cluster. We also observed apparent admixture from the
364 Lewis Lake strain in the Marquette sample, which suggests that these strains may have
365 introgressed in the hatchery at some point.

366 *Assignment of accuracy of simulated individuals*

We constructed five reporting groups for assignment based on STRUCTURE analysis and preliminary assignment tests: Seneca Lake, Parry Sound, Lewis Lake, Green Lake, and Superior (Table 1). We originally included a separate reporting group for Klondike Reef, but low assignment accuracy (63% correct for pure simulations) prompted us to combine Klondike Reef with the rest of the Lake Superior populations to form a single reporting group. Correct assignment of pure crosses was 91% on average and correct assignment of F1 interstrain hybrids was 83% on average (Table 4). Misclassification of cross type (i.e. pure or hybrid) was rare, with zero simulated pure individuals identified as putative hybrids, and < 5% of simulated hybrid individuals identified as pure. Pure and hybrid crosses from Seneca Lake and Parry Sound had the highest assignment accuracies (> 95% for pure crosses, 91% for Seneca x Parry Sound hybrid cross), while strains from Lake Superior had lower accuracy (75% for pure cross, ~80% for hybrid crosses between Superior and other strains). The lower assignment accuracy for the Lake Superior strain was largely caused by misassignment of simulated individuals to the Lewis Lake reporting group, an expected result given the admixture observed in the STRUCTURE analysis. Fortunately, most naturally reproduced lake trout with Lake Superior ancestry found in Lake Michigan are likely to be derived from the genetically pure Apostle Island strain, as the Marquette strain was not stocked at high numbers after 1990. This means that our realized assignment accuracy for Lake Superior strains is likely higher than the 75% derived with simulations.

Assignment accuracy of mixture proportions was higher for pure strains using ONCOR but much higher for hybrid crosses with STRUCTURE (Table 4). ONCOR performed poorly for all hybrid crosses except Seneca x Parry, with substantial mis-assignment to the Superior group across many different crosses. Accuracy of mixture proportions was generally high with

390 STRUCTURE, with the lowest accuracy to the Superior pure strain (87%). The lowest accuracy
391 to hybrid classes with STRUCTURE was 44% (expected 50%) for the Green Lake strain in two
392 crosses. Although we were unable to construct confidence intervals from the STRUCTURE
393 analysis, data from mixtures of simulated individuals does provide information on potential
394 sources of error. Superior strain fish are misclassified most frequently, with pure Superior
395 individuals being classified as Lewis Lake 7% of the time and Green Lake 3%. Mis-assignment
396 in the hybrid crosses often consisted of individuals from Green Lake or Lewis Lake strains being
397 assigned to Lake Superior. These results indicate that estimates for the Lake Superior strain
398 could be slightly higher or lower than reported depending on the number of hybrids found in the
399 mixture. However, we are highly confident that our estimates are within a few percent of the
400 correct proportions given the relatively low error rates in general with the STRUCTURE
401 analysis.
402 *Strain composition of naturally produced lake trout*
403 All but one of the 848 naturally produced lake trout that we analyzed were assigned to a
404 strain that has been stocked into Lake Michigan. The single exception was a five-year-old fish
405 which assigned as a Parry Sound x Superior hybrid and was likely a migrant from Lake Huron
406 based on stocking history of the Parry Sound strain in Lake Huron. We subsequently removed
407 this sample from the dataset. Metadata for all naturally produced lake trout is in Table S7.
408 The Seneca Lake strain was the most frequent genetic origin of wild fish captured in
409 Lake Michigan and was found at proportions > 50% for six of seven spatial strata (Fig 4, Table
410 5). The sole exception was the Grand Traverse Bay stratum, where the Seneca Lake strain was
411 encountered at 26% compared to 52% for the Lewis Lake strain. Although there were few clear
412 spatial differences in strain composition in strata outside of Grand Traverse Bay, the proportion

of the Green Lake strain did appear to increase from North to South. Additionally, the proportion of the Superior strain appeared to be higher on the west side of the lake. The proportion of pure and hybrid crosses was extremely similar on average (52% pure, 48% hybrid), with six of seven strata displaying proportions of pure crosses between 39% and 60% (Table S8). Interestingly, the north stratum displayed the highest percentage of pure crosses (75%), with most of these (60% of total sample) represented by the Seneca Lake strain.

Comparison between observed and expected strain contributions

The expected stocking proportions that we derived largely reflected stocking proportions in the early 2000s, as this is the period when the fish that produced the natural recruits we sampled were likely stocked. In this time period, the four major genetic strains that we investigated (Seneca Lake, Lewis Lake, Green Lake, and Lake Superior), were stocked at similar rates ranging from 19% overall for Seneca Lake to 32% overall for Lake Superior (Table 5). However, these proportions varied among spatial strata, with Lewis Lake and Superior stocked at higher numbers in the northern part of Lake Michigan and Green Lake and Seneca Lake stocked more in the South.

Comparisons of strain proportions derived from genetic data to expected proportions calculated using the composite method described above revealed that the Seneca Lake strain was found at higher proportions than expected, but that this pattern varied among strata (Fig. 5, Table 5). Seneca Lake outperformed expectations based on stocking for all strata and was overrepresented by 38% on average (range 12% to 66%, Table 5). Contrastingly, the Superior and Green Lake strains performed the poorest on average, and underperformed expectations by ~15% (range underperform by 46% to overperform by 6%). The Lewis Lake strain was

435 intermediate, underperforming expectations by 9% on average (range underperform by 21% to
436 overperform by 3%.)

437 Strain performance was highly variable by area, with the largest differences in
438 performance among strains observed in the North stratum and the smallest differences observed
439 in the Southern Refuge and Grand Traverse Bay strata (Fig. 5, Table 5). For example, in the
440 North stratum, the only highly successful strain appeared to be Seneca Lake, whereas in the
441 Southern Refuge and Grand Traverse Bay strata, most strains performed similarly to
442 expectations based on stocking. This spatial variation is well illustrated by the Lake Superior
443 strain, which performed poorly in the North stratum but exceeded expectations by 5% on average
444 in the Southwest and Southern Refuge strata. Of the four strains, only Green Lake failed to
445 exceed stocking expectations in two or more strata, indicating that this strain may perform poorly
446 regardless of where it is stocked.

447 When composite estimate of expected strain compositions reconstructed from stocking
448 records (i.e. expected proportions in Table 5) were used to estimate the proportion of each cross
449 type that should be observed based on random mating, these estimates differed substantially from
450 what we observed in our genetic data (Fig. 6, Table S8). However, when random mating
451 simulations were conducted based on observed genetic proportions (i.e. observed proportions in
452 Table 5), differences in the frequency of each cross type between observed and expected data
453 were minimal. The large differences between stocking proportions and observed genetic
454 proportions for each cross type (Fig. 6a) is likely a result of differential survival among strains.
455 However, the fact that there are few differences between observed and expected proportions of
456 each cross type based on genetic data (Fig. 6b) suggests that there is no substantial fitness

457 difference between pure and hybrid cross types, that is, there is no strong evidence of either
458 hybrid inferiority or hybrid vigor.

459 ## Discussion

460 *Genetic differentiation of reference strains and accuracy of strain assignment*

461 The patterns of genetic differentiation among strains that we documented were similar to
462 those observed in past studies using allozymes and microsatellites (Marsden et al. 1989; Page et
463 al. 2004; Page et al. 2003). For example, Page et al. (2004) genotyped seven of the eight strains
464 that were analyzed in the current study with a panel of nine microsatellites and also found that
465 Seneca Lake and Parry Sound strains were highly diverged, with less divergence among strains
466 derived from Lake Superior and remnant Lake Michigan populations. Additionally, our results
467 and those of Page et al. (2004) indicated that Lewis Lake and Green Lake broodstock, which
468 were both derived from Lake Michigan lake trout, have diverged substantially during captivity
469 and release into other systems. It is possible that some of this differentiation may be due to
470 genetic drift caused by moderate population bottlenecks as postulated by Page et al. (2004).
471 However, the complex histories of these strains make it difficult to determine why they are
472 differentiated from each other. While genetic diversity of the strains analyzed here appears to be
473 adequate, with no major differences in genetic diversity among strains, some of these strains
474 appear to have been substantially modified through domestication and hatchery practices,
475 highlighting the importance of propagation strategies that maintain genetic integrity (e.g. Waters
476 et al. 2015).

477 The only strain that we analyzed that was not analyzed by Page et al. (2004) was the
478 Klondike Reef strain; we found that approximately 20% of fish in this strain were genetically
479 similar to Lake Superior lean collections and 80% represented a distinct genetic cluster, probably

480 composed of Klondike Reef humpers. This pattern was also observed by Salvesen (2015), who
481 conducted assignment tests and found approximately 20% misassignment of this strain to Lake
482 Superior lean populations. It is possible that the Klondike Reef broodstock may contain
483 approximately 20% lean lake trout and 80% humper lake trout, but genetic data are not
484 diagnostic for ecotype (Perreault-Payette et al. 2017). Finally, we documented a small amount of
485 potential admixture between the Marquette and Lewis Lake strains, which was also observed
486 using assignment tests by Page et al. (2003). This admixture is likely the result of complex
487 genetic history, as the Marquette strain is the oldest lake trout broodstock used in the Great
488 Lakes, was derived from several sources in Lake Superior, and has intermittently been mixed
489 with others (Page et al. 2004). For example, in the 1960s lake trout from the Apostle Islands and
490 Green Lake hatchery strain were added to the Marquette broodstock (Krueger et al. 1983).
491 We were able to assign fish of unknown origin to reference strains with relatively high
492 accuracy despite the small amounts of genetic admixture described above. Strain assignment of
493 lake trout in the Great Lakes has been conducted for over 30 years, first with allozymes
494 (Marsden et al. 1989) and more recently with panels of 4 to 15 microsatellites (DeKoning et al.
495 2006; Page et al. 2003; Roseman et al. 2009; Scribner et al. 2018). These previous studies have
496 generally been able to assign individuals to pure crosses with relatively high accuracy, but
497 assignment of interstrain crosses has been difficult. Recently, Scribner et al. (2018) used
498 genotypes from 15 microsatellites and a modification of the Rannala and Mountain (1997)
499 assignment algorithm developed by Gaggiotti et al. (2004) to estimate assortative mating
500 probabilities among strains. While this approach provides important information on assortative
501 mating, it is still not possible to calculate the frequency of a given cross type to determine if
502 certain cross types are found more or less frequently than expected. Fortunately, the increase in

503 assignment power facilitated by genotyping 36 microsatellites compared to 15 allowed us to
504 assign individuals to cross type with relatively high accuracy, representing a significant advance
505 for studies of lake trout strain composition. However, assignment accuracy for certain crosses
506 was still somewhat low (< 80% for three comparisons), and accurate assignment past F1 crosses
507 is likely impossible with the current methods; these assignment accuracies could be increased
508 using genomic tools (Allendorf et al. 2010), which are currently being developed by K. Scribner
509 at Michigan State University.

510 *Relative performance of reference strains*

511 Strain assignment of naturally produced lake trout captured throughout Lake Michigan
512 indicated that the Seneca Lake strain was overrepresented in most geographic strata, whereas the
513 other strains were underrepresented on average with some variation among strata. Here, we focus
514 on the Seneca strain as it was most consistently overrepresented (see below), but it is also
515 important to note that the Green Lake strain performed poorly in every strata where it was
516 stocked in at least moderate numbers. It is possible that the long history of domestication in this
517 strain led to inbreeding depression or domestication selection, making it maladapted to the wild
518 compared to other strains (reviewed in Christie et al. 2014). Performance of the Lewis Lake and
519 Lake Superior strains was more variable, with the Lewis Lake strain performing similar to or
520 exceeding expectations in the Traverse Bay and Southern Refuge strata and the Superior strain
521 performing similar to or exceeding expectations in the Southwest and Southern Refuge strata.
522 This spatial variation in strain performance provides important information for managers that can
523 potentially be used to inform stocking with the goal of maximizing post-release survival and
524 increase the probability of producing wild recruits.

525 Although we generally interpreted our results based on point estimates for stock
526 compositions, understanding and interpreting uncertainty in our data is also important. The two
527 major sources of uncertainty in our data are uncertainty due to finite and sometimes relatively
528 small sample sizes (e.g. as low as 77 fish in the East strata) and uncertainty due to variation in
529 assignment accuracy across strains (as low as 75% accuracy for the pure Superior strain). Our
530 approach for constructing confidence intervals integrated both of these sources of uncertainty,
531 and uncertainty due to sample sizes was much more influential than uncertainty due to
532 assignment accuracy. Specifically, weighting for assignment accuracy increased the size of
533 confidence intervals by a maximum of about 0.15, while variation in sample sizes changed the
534 size of confidence intervals much more substantially. Fortunately, differences between observed
535 and expected strain proportions are generally very large across our dataset, making our
536 inferences highly robust despite relatively wide confidence intervals.
537 The Seneca Lake strain was overrepresented based on expectations from stocking in the
538 northern and, to a lesser degree, southern areas of the lake, but the proportional contribution of
539 the Seneca Lake strain was similar to expectations in the Southern Refuge and Grand Traverse
540 Bay. Previous analyses using genetic assignment to estimate strain composition of naturally
541 produced lake trout in Lakes Ontario, Huron, and Michigan have generally found that the
542 observed contributions of the Seneca Lake strain are higher than expected (DeKoning et al.
543 2006; Marsden et al. 1989; Page et al. 2003; Roseman et al. 2009; Scribner et al. 2018). In Lake
544 Michigan, both Page et al. (2003), who conducted genetic analysis on young-of-the-year samples
545 from Little Traverse Bay (northeastern Lake Michigan), and DeKoning et al. (2006), who
546 analyzed fry and egg samples from southern Lake Michigan, also documented higher than
547 expected proportions of Seneca Lake individuals among naturally produced recruits. However,

548 relative survival results based on coded wire tag (CWT) recoveries of adults in Lake Michigan
549 do not necessarily reflect the same trends (Bronte et al. 2007; Kornis et al. 2019b; McKee et al.
550 2004). Earlier CWT studies in Lake Michigan (Bronte et al. 2007) and Lake Huron (Eshenroder
551 et al. 1995) found greater survival of Seneca Lake strain compared to Lewis Lake strain. By
552 contrast, a more recent analysis of CWT returns from 1993-2003 year-classes found that the
553 Seneca Lake strain had lower relative survival compared to Lewis Lake and Green Lake strains
554 when stocked in southern Lake Michigan (Kornis et al. 2019b). Findings from our study appear
555 more similar to recent results from CWT data than from genetic analysis of eggs and fry.
556 However, it is difficult to decipher whether differences between studies are the result of samples
557 being collected in different time periods or at different life stages. Nevertheless, our study
558 illustrates the importance of conducting spatially representative sampling when evaluating strain
559 performance, as our results suggest that dynamics that lead to differences in strain representation
560 are multifaceted and spatially variable (see discussion of Seneca strain dynamics below).
561 Overrepresentation of Seneca Lake strain may stem from a combination of movement of
562 wild fish after recruitment and characteristics of the strain that facilitate survival in challenging
563 environments. The Seneca Lake strain was most overrepresented in northern Lake Michigan,
564 where lake trout experience high mortality from predation by sea lamprey as well as exploitation
565 by commercial and, to a lesser extent, recreational fisheries (Kornis et al. 2019b). These sources
566 of mortality have resulted in a truncated age distribution in northern Lake Michigan, with fewer
567 sexually mature fish compared to other areas (Bronte et al. 2007; Kornis et al. 2019b) and, as a
568 result, little evidence of natural reproduction has been observed (Kornis et al. 2019a; LMLTWG
569 2019). In contrast, lake trout in southern Lake Michigan, in particular the Southern Refuge, are
570 not as exposed to sea lamprey predation or fishing mortality and hence have higher densities of

571 older, mature fish that substantially contribute to natural reproduction in Lake Michigan (Bronte
572 et al. 2007; Kornis et al. 2019b). The Seneca Lake strain is stocked heavily in southern Lake
573 Michigan and especially on the Southern Refuge, therefore, many of the sexually mature fish in
574 this area are Seneca Lake strain. We hypothesize that a partial explanation for the high
575 proportion of naturally produced Seneca Lake strain fish in many parts of Lake Michigan,
576 especially the north, is that these fish were produced in southern Lake Michigan and then
577 migrated to other parts of the lake.

578 Our estimates of expected proportions accounted for movement of spawners (stocked
579 fish) among strata but were not weighted for potential reproductive output. That is, estimated
580 proportions assumed of equal potential for reproduction in all strata. Although wild reproduction
581 has been observed on other reefs in Illinois (Patterson et al. 2016) and is likely occurring at
582 several locations around the lake, it is likely there are recruitment hotspots given the habitat
583 requirements for lake trout reproduction (e.g., Marsden et al. 1995) and the fact that wild lake
584 trout reproduction in Lake Michigan is a relatively recent development. Our results, although not
585 conclusive, are consistent with what would be expected if disproportionately high levels of
586 recruitment occurred on the Southern Refuge with dispersal to other strata thereafter. The
587 Southern Refuge is rich in spawning habitat, protected from exploitation, has a robust age
588 structure of parental stock, and a high abundance of spawners (Dawson et al. 1997; Kornis et al.
589 2019b). The same can be said for Julian's Reef and other reefs in Illinois (Patterson et al. 2016),
590 but spawner populations are likely largest on the Southern Refuge due to more habitat, higher
591 stocking levels, and lower total mortality. Movement of stocked fish from the Southern Refuge
592 to all strata was high, with the exception of northern Lake Michigan and Grand Traverse Bay
593 (Supplementary File 2). However, even a low rate of immigration of wild fish from southern

Lake Michigan to the north could still explain overrepresentation of Seneca strain in this area as most stocked fish recovered in the north stratum originated there, and recruitment in this area is likely near zero as suggested by survey data (Kornis et al. 2019b).

Another possible explanation for the higher frequency of the Seneca Lake strain in northern Lake Michigan is unique and heritable characteristics that may make them more successful in the northern part of the lake (Bronte et al. 2007). The Seneca Lake strain appears to occupy cooler and deeper water than other strains, likely reducing their encounter rate with sea lamprey (Schneider et al. 1996). Others have noted that Seneca Lake strain have a relatively fast growth rate compared to other strains and linked this to bathythermal distribution, a heritable trait (Elrod et al. 1996; Royce 1951), although this did not appear to translate into differences of age-at-maturity (Elrod et al. 1996). We hypothesize that the high frequency of the Seneca Lake strain in northern Lake Michigan is likely a result of both the movement patterns described above as well as the life history characteristics of this strain that facilitate increased survival.

Our results suggest that strain performance can vary substantially across relatively small spatial scales potentially due to heritable genetic differences among strains. Heritable genetic differences in traits such as run timing (Prince et al. 2017), temperature tolerance (Philipp and Whitt 1991), and spawning habitat preferences (Jennings et al. 1996) have been frequently documented in fish. Additionally, management agencies often stock multiple genetic strains of fish with the goal of providing robust fisheries and increased fishing opportunities (Bartron et al. 2004; Miller et al. 2009). However, while strain evaluations often find substantial differences in behavior or survival among strains (Van Offelen et al. 1993), these studies are relatively rare and are often isolated to small geographic areas. Our study illustrates the importance of conducting thorough strain evaluations to gain a more complete picture of strain performance; we suggest

617 that management agencies undertake these types of studies more frequently to improve stocking
618 efficiency and promote robust fisheries.
619 *Frequency of interstrain hybrids*
620 The expected proportions of each cross type based on stocking data were substantially
621 different than observed proportions but observed and expected proportions were generally
622 similar when expected proportions were calculated based on genetic data. We hypothesize that
623 the large differences in expected cross type frequencies calculated from stocking data are largely
624 a function of differences in strain survival and movement patterns (discussed above).
625 Additionally, we hypothesize that the relatively small differences between observed and
626 expected proportions based on genetic data indicate that the strains are generally breeding
627 randomly, and that there is no substantial fitness advantage or disadvantage for F1 hybrids. It is
628 notable that the frequency of crosses including Seneca Lake are also similar to expectations,
629 especially in the southern area of the lake. Seneca Lake fish likely occupy different depths (see
630 above), yet this does not seem to facilitate any reproductive isolation from the other strains,
631 which suggests similar spawning habitat and behavior.
632 Our results differ slightly from those of Scribner et al. (2018), who documented
633 departures from random mating in early generations of lake trout natural reproduction in Lake
634 Huron and hypothesized that F1 crosses could have lower fitness. However, in Scribner et al.
635 (2018) mating appeared to become more random in subsequent generations suggesting that
636 interstrain crosses are common and relatively fit in Lake Huron. Both our study and Scribner et
637 al. (2018) indicate that interstrain hybridization is common in lake trout inhabiting the Great
638 Lakes; if random mating in these populations continues, the populations will likely resemble a

hybrid swarm (i.e. a population composed mostly of hybrid individuals and multi-generation backcrosses) in only a few generations.

Although the term hybrid swarm often has a negative connotation, the fitness impacts of hybridization are still unclear. For example, Johansen-Morris and Latta (2006) examined the fitness consequences of hybridization in a plant species and found evidence of both hybrid vigor (improved fitness of hybrids due to dominance effects) and hybrid breakdown (disruption of coadapted gene complexes causing reductions in fitness). Hybridization can reverse adaptive differentiation by facilitating gene flow between locally adapted populations (Abbott et al. 2013; Baillie et al. 2016), for example locally adapted Seneca Lake and remnant Lake Michigan strains. However, hybridization can also create unique genetic combinations that are more fit for their current environment than either of the two parental populations (Comeault and Matute 2018). It is important to note that all hybrids in this study are assumed to be F1 hybrids, but many F2+ generation hybrids will likely be present in the future. The fitness of advanced generation backcrosses has been shown to decrease quickly when fish of different species hybridize (e.g., Muhlfeld et al. 2009), but the fitness consequences for moderately diverged populations such as the lake trout strains in our study are less clear. This uncertainty highlights the importance of conducting future monitoring to ensure that interstrain hybrids (especially advanced-generation backcrosses) are not leading to a reduction in population viability of lake trout in Lake Michigan.

Management implications and conclusions

We suggest that future stocking efforts for lean lake trout in Lake Michigan primarily utilize the Seneca Lake and Lewis Lake strains. Mortality from sea lamprey predation is relatively high in the northern part of Lake Michigan, and the Seneca Lake strain appears to

662 possess unique characteristics that allow it to be more successful than other strains in this
663 environment. However, the Seneca Lake strain's performance was similar to other strains in
664 areas where sea lamprey predation is lower, which suggests that the performance advantage of
665 the Seneca Lake strain is decreased in the absence of sea lamprey. Movement of wild fish post-
666 recruitment may also contribute to the high amount of wild fish with Seneca Lake heritage,
667 especially if the Southern Refuge is a nexus for recruitment, in which case the observed
668 differences in performance may not be solely related to fitness. The Lewis Lake strain is the last
669 available strain derived from remnant Lake Michigan lake trout (the Green Lake strain has been
670 discontinued) and likely still contains locally adapted alleles despite its history of feralization in
671 Lewis Lake, WY and domestication in the hatchery system. Additionally, the Lewis Lake strain
672 generally performed well outside of the northern area of Lake Michigan including the Southern
673 Refuge, where it performed similarly to expectations based on stocking. Our stocking
674 suggestions, which promote genetic diversity without requiring the expensive maintenance of
675 many potentially redundant hatchery strains, are generally consistent with other recent
676 recommendations (e.g., Kornis et al. 2019b) and with current management plans (Dexter et al.
677 2011).

678 There is also discussion about continuing to stock the Klondike Reef humper strain in
679 Lake Michigan and introducing it into Lake Ontario (S. Lapan, N.Y. Department of
680 Environmental Conservation, personal communication). This strain was introduced to the Lake
681 Michigan propagation program in 2011 and is now being recruited into assessment fisheries and
682 will potentially be recruiting into the parental stock soon. Thus, we were not able to evaluate its
683 potential to contribute to natural lake trout recruitment. Humper lake trout naturally occur in
684 Lake Superior where they are typically restricted to deep offshore shoals or reefs (e.g., "humps";

685 Rahrer 1965), which suggests they may have restricted movement relative to lean morphs.
686 Humpers also consume more Mysis, benthic fish (Peck 1975; Rogers et al. 2019; Sierszen et al.
687 2014; Vinson et al. in press; Zimmerman et al. 2007) and terrestrial invertebrates (Sitar et al.
688 2020) than lean lake trout. Changes in the Great Lakes food web, such as the decline in the
689 pelagic alewife (*Alosa pseudoharengus*) and the increase in the benthic round goby (*Neogobius*
690 *melanostomus*) (Madenjian et al. 2018), are already reflected in diets of lake trout (e.g., Happel
691 et al. 2017; Kornis et al. 2020; Luo et al. 2019) and some other salmonines (Leonhardt et al.
692 2020), and may further support the Klondike Reef strain as an appropriate choice for restoration
693 efforts. However, it will be important to assess the ultimate risk of Klondike Reef humpers
694 interbreeding with lean lake trout, which could have fitness consequences if hybrid offspring are
695 maladapted to both shallow and deepwater environments. We speculate that the genetic risks of
696 stocking the Klondike strain are relatively low given evidence that lake trout ecotypes often
697 naturally coexist in other systems. We suggest routine genetic monitoring confirming the genetic
698 purity of humper strain broodstock and continued genetic monitoring of natural recruits to
699 confirm that humper-lean hybrids are rare.

700 In conclusion, we found that strain performance of lake trout varied substantially across
701 Lake Michigan and hypothesize that these differences are largely due to adaptive differences
702 among strains and post-recruitment movement. Additionally, we found no evidence for
703 reproductive isolation among strains and no evidence for lower or higher fitness of interstrain
704 crosses, which suggests that mating among lake trout strains stocked into Lake Michigan is
705 essentially random. Our study provided a more nuanced understanding of strain performance
706 across the whole lake and suggested adequate performance of several strains, compared to
707 previous genetic studies in Lake Michigan, which focused on smaller geographic scales and

suggested that the Seneca Lake strain was superior to all other strains. Additionally, the increased power of our genetic marker panel compared to panels used in previous genetic studies improved strain assignment accuracy and facilitated assignment of interstrain crosses for the first time in this system. Our study demonstrates the utility of thorough strain evaluations for informing conservation and management and provides a roadmap for future researchers conducting strain evaluations in other taxa.

Acknowledgements

We thank members of the U.S. Fish and Wildlife Service, Chippewa Ottawa Resource Authority, Little Traverse Bay Band of Odawa Indians, Grand Traverse Band of Ottawa and Chippewa Indians, Wisconsin, Michigan, Illinois, and Indiana Department of Natural Resources for assisting with sample collection. We also thank Brad Erdman and Jenna Ruzich from UW-Stevens Point for assisting with laboratory work Kim Scriber (Michigan State University) for sharing samples from reference strains and Meredith Bartron (U.S. Fish and Wildlife Service) for sharing samples from reference strains and reviewing a draft of this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B.,

731 Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., and Zinner, D.
732 2013. Hybridization and speciation. J. Evol. Biol. **26**(2): 229-246. doi:10.1111/j.1420-
733 9101.2012.02599.x.
734 Allendorf, F.W., Hohenlohe, P.A., and Luikart, G. 2010. Genomics and the future of
735 conservation genetics. Nat. Rev. Genet. **11**(10): 697-709. doi:10.1038/nrg2844.
736 Baillie, S.M., Muir, A.M., Scribner, K., Bentzen, P., and Krueger, C.C. 2016. Loss of genetic
737 diversity and reduction of genetic distance among lake trout *Salvelinus namaycush*
738 ecomorphs, Lake Superior 1959 to 2013. J. Great. Lakes. Res. **42**(2): 204-216.
739 doi:https://doi.org/10.1016/j.jglr.2016.02.001.
740 Bartron, M.L., Swank, D.R., Rutherford, E.S., and Scribner, K.T. 2004. Methodological bias in
741 estimates of strain composition and straying of hatchery-produced steelhead in Lake
742 Michigan tributaries. N. Am. J. Fish. Manage. **24**(4): 1288-1299. doi:10.1577/M03-
743 083.1.
744 Bergstedt, R.A., Argyle, R.L., Seelye, J.G., Scribner, K.T., and Curtis, G.L. 2003. In situ
745 determination of the annual thermal habitat use by lake trout (*Salvelinus namaycush*) in
746 Lake Huron. J. Great. Lakes. Res. **29**: 347-361. doi:https://doi.org/10.1016/S0380-
747 1330(03)70499-7.
748 Bronte, C.R., Krueger, C.C., Holey, M.E., Toneys, M.L., Eshenroder, R.L., and Jonas, J.L. 2008.
749 A guide for the rehabilitation of lake trout in Lake Michigan. Great Lakes Fishery Comm.
750 Misc. Publ. 2008-01. Available from http://www.glfc.org/pubs/pub.htm#misc.
751 Bronte, C.R., Walch, K.A., Dettmers, J.M., Gaden, M., Connerton, M.J., Daniels, M.E., and
752 Newcomb, T.J. 2012. A Coordinated Mass Marking Program for Salmonines Stocked

753 into the Laurentian Great Lakes. In Advances in Fish Tagging and Marking Technology,
754 American Fisheries Society, Bethesda, Maryland, USA.
755 Bronte, C.R., Jonas, J.L., Holey, M.E., Eshenroder, R.L., Toneys, M.L., McKee, P., Breidert, B.,
756 Claramunt, R.M., Ebener, M.P., Krueger, C.C., Wright, G., and Hess, R. 2003. Possible
757 impediments to Lake Trout restoration in Lake Michigan. Great Lakes Fishery
758 Commission, Ann Arbor, Michigan.
759 Bronte, C.R., Holey, M.E., Madenjian, C.P., Jonas, J.L., Claramunt, R.M., McKee, P.C., Toneys,
760 M.L., Ebener, M.P., Breidert, B., Fleischer, G.W., Hess, R., Martell, A.W., and Olsen,
761 E.J. 2007. Relative abundance, site fidelity, and survival of adult lake trout in Lake
762 Michigan from 1999 to 2001: implications for future restoration strategies. N. Am. J.
763 Fish. Manage. **27**(1): 137-155. doi:10.1577/M05-214.2.
764 Brown, S.B., Fitzsimons, J.D., Honeyfield, D.C., and Tillitt, D.E. 2005. Implications of thiamine
765 deficiency in Great Lakes salmonines. J. Aquat. Anim. Health **17**(1): 113-124.
766 doi:10.1577/H04-015.1.
767 Campana, S.E., Casselman, J.M., and Jones, C.M. 2008. Bomb radiocarbon chronologies in the
768 Arctic, with implications for the age validation of lake trout (*Salvelinus namaycush*) and
769 other Arctic species. Can. J. Fish. Aquat. Sci. **65**(4): 733-743. doi:10.1139/f08-012.
770 Christie, M.R., Ford, M.J., and Blouin, M.S. 2014. On the reproductive success of early-
771 generation hatchery fish in the wild. Evol. Appl. **7**(8): 883-896. doi:10.1111/eva.12183.
772 Comeault, A.A., and Matute, D.R. 2018. Genetic divergence and the number of hybridizing
773 species affect the path to homoploid hybrid speciation. Proc. Natl Acad. Sci. USA
774 **115**(39): 9761-9766. doi:10.1073/pnas.1809685115.

775 Dawson, K.A., Eshenroder, R.L., Holey, M.E., and Ward, C. 1997. Quantification of historic
776 lake trout (*Salvelinus namaycush*) spawning aggregations in Lake Michigan. Can. J. Fish.
777 Aquat. Sci. **54**(10): 2290-2302. doi:10.1139/f97-136.
778 DeKoning, J., Keatley, K., Phillips, R., Rhydderch, J., Janssen, J., and Noakes, M. 2006. Genetic
779 analysis of wild lake trout embryos recovered from Lake Michigan. Trans. Am. Fish.
780 Soc. **135**(2): 399-407. doi:10.1577/t05-044.1.
781 Dexter, J.L., Eggold, B.T., Gorenflow, T.K., Horns, W.H., Robillard, S.R., and Shipman, S.T.
782 2011. A fisheries management implementation strategy for the rehabilitation of Lake
783 Trout in Lake Michigan. Great Lakes Fishery Commission, Lake Michigan Committee,
784 Ann Arbor, Michigan.
785 Earl, D.A., and vonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for
786 visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet.
787 Resour. **4**(2): 359-361. doi:10.1007/s12686-011-9548-7.
788 Elrod, J.H., Ostergaard, D.E., and Schneider, C.P. 1988. Comparison of Hatchery-Reared Lake
789 Trout Stocked as Fall Fingerlings and as Spring Yearlings in Lake Ontario. N. Am. J.
790 Fish. Manage. **8**(4): 455-462. doi:10.1577/1548-8675(1988)008<0455:cohrlt>2.3.co;2.
791 Elrod, J.H., O'Gorman, R., and Schneider, C.P. 1996. Bathythermal distribution, maturity, and
792 growth of lake trout strains stocked in U.S. Waters of Lake Ontario, 1978–1993. J. Great.
793 Lakes. Res. **22**(3): 722-743. doi:https://doi.org/10.1016/S0380-1330(96)70992-9.
794 Eshenroder, R.L., Robert Payne, N., Johnson, J.E., Bowen, C., and Ebener, M.P. 1995. Lake
795 trout rehabilitation in Lake Huron. J. Great. Lakes. Res. **21**: 108-127.
796 doi:https://doi.org/10.1016/S0380-1330(95)71086-3.

797 Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals
798 using the software STRUCTURE: a simulation study. Mol. Ecol. **14**(8): 2611-2620.
799 doi:10.1111/j.1365-294X.2005.02553.x.
800 Fitzsimons, J.D., and O'Gorman, R. 1996. Fecundity of hatchery lake trout in Lake Ontario. J.
801 Great. Lakes. Res. **22**(2): 304-309. doi:https://doi.org/10.1016/S0380-1330(96)70957-7.
802 Fraser, D.J. 2008. How well can captive breeding programs conserve biodiversity? A review of
803 salmonids. Evol. Appl. **1**(4): 535-586. doi:10.1111/j.1752-4571.2008.00036.x.
804 Gaggiotti, O.E., Brooks, S.P., Amos, W., and Harwood, J. 2004. Combining demographic,
805 environmental and genetic data to test hypotheses about colonization events in
806 metapopulations. Mol. Ecol. **13**(4): 811-825. doi:10.1046/j.1365-294X.2003.02028.x.
807 George, A.L., Kuhajda, B.R., Williams, J.D., Cantrell, M.A., Rakes, P.L., and Shute, J.R. 2009.
808 Guidelines for propagation and translocation for freshwater fish conservation. Fisheries
809 **34**(11): 529-545. doi:10.1577/1548-8446-34.11.529.
810 Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate *F*-statistics. J. Hered.
811 **86**(6): 485-486. doi:https://doi.org/10.1093/oxfordjournals.jhered.a111627.
812 Halverson, M.A. 2008. Stocking trends: A quantitative review of governmental fish stocking in
813 the United States, 1931 to 2004. Fisheries **33**(2): 69-75. doi:10.1577/1548-8446-33.2.69.
814 Hansen, M.J. 1999. Lake trout in the Great Lakes: basin wide stock collapse and binational
815 restoration. In: Taylor WW, Ferreri CP, editors. Great Lakes fisheries and policy
816 management: a binational perspective. East Lansing: Michigan State University Press. p.
817 417–453.

Hanson, S.D., Holey, M.E., Treska, T.J., Bronte, C.R., and Eggebraaten, T.H. 2013. Evidence of wild juvenile Lake Trout recruitment in western Lake Michigan. N. Am. J. Fish. Manage. **33**(1): 186-191. doi:10.1080/02755947.2012.754804.

Happel, A., Jonas, J.L., McKenna, P.R., Rinchard, J., He, J.X., and Czesny, S.J. 2017. Spatial variability of lake trout diets in Lakes Huron and Michigan revealed by stomach content and fatty acid profiles. Can. J. Fish. Aquat. Sci. **75**(1): 95-105. doi:10.1139/cjfas-2016-0202.

Heinrich, J.W., Mullett, K.M., Hansen, M.J., Adams, J.V., Klar, G.T., Johnson, D.A., Christie, G.C., and Young, R.J. 2003. Sea lamprey abundance and management in Lake Superior, 1957 to 1999. J. Great. Lakes. Res. **29**: 566-583. doi:https://doi.org/10.1016/S0380-1330(03)70517-6.

Holey, M.E., Rybicki, R.W., Eck, G.W., Brown, E.H., Marsden, J.E., Lavis, D.S., Toneys, M.L., Trudeau, T.N., and Horrall, R.M. 1995. Progress toward lake trout restoration in Lake Michigan. J. Great. Lakes. Res. **21**: 128-151. doi:https://doi.org/10.1016/S0380-1330(95)71087-5.

Hubisz, M.J., Falush, D., Stephens, M., and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. **9**(5): 1322-1332. doi:10.1111/j.1755-0998.2009.02591.x.

Huff, D.D., Miller, L.M., Chizinski, C.J., and Vondracek, B. 2011. Mixed-source reintroductions lead to outbreeding depression in second-generation descendents of a native North American fish. Mol. Ecol. **20**(20): 4246-4258. doi:10.1111/j.1365-294X.2011.05271.x.

839 Jennings, M.J., Claussen, J.E., and Philipp, D.P. 1996. Evidence for heritable preferences for
840 spawning habitat between two walleye populations. Trans. Am. Fish. Soc. **125**(6): 978-
841 982. doi:10.1577/1548-8659(1996)125<0978:EFHPFS>2.3.CO;2.
842 Johansen-Morris, A.D., and Latta, R.G. 2006. Fitness consequences of hybridization between
843 ecotypes of *Avena barbata*: hybrid breakdown, hybrid vigor, and transgressive
844 segregation. Evolution **60**(8): 1585-1595. doi:10.1111/j.0014-3820.2006.tb00503.x.
845 Kincaid, H.L., Krueger, C.C., and May, B. 1993. Preservation of genetic variation in the Green
846 Lake strain Lake Trout derived from remnant domestic and feral populations. N. Am. J.
847 Fish. Manage. **13**(2): 318-325. doi:10.1577/1548-
848 8675(1993)013<0318:POGVIT>2.3.CO;2.
849 Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., and Mayrose, I. 2015. Clumpak:
850 a program for identifying clustering modes and packaging population structure inferences
851 across K. Mol. Ecol. Resour. **15**(5): 1179-1191. doi:10.1111/1755-0998.12387.
852 Kornis, M.S., Bunnell, D.B., Swanson, H.K., and Bronte, C.R. 2020. Spatiotemporal patterns in
853 trophic niche overlap among five salmonines in Lake Michigan, USA. Can. J. Fish.
854 Aquat. Sci. **77**(6): 1059-1075. doi:10.1139/cjfas-2019-0288.
855 Kornis, M.S., Pankow, S.R., Cressman, J.L., Webster, A., Lane, A., and Bronte, C.R. 2019a.
856 Summary statistics of tag and data field recovery program on Lakes Michigan and Huron.
857 Report #2019-03, USFWS-Green Bay Fish and Wildlife Conservation Office, New
858 Franken, WI.
859 Kornis, M.S., Bronte, C.R., Holey, M.E., Hanson, S.D., Treska, T.J., Jonas, J.L., Madenjian,
860 C.P., Claramunt, R.M., Robillard, S.R., Breidert, B., Donner, K.C., Lenart, S.J., Martell
861 Jr., A.W., McKee, P.C., and Olsen, E.J. 2019b. Factors affecting postrelease survival of

862 coded-wire-tagged lake trout in Lake Michigan at four historical spawning locations. N.
863 Am. J. Fish. Manage. **39**(5): 868-895. doi:10.1002/nafm.10338.
864 Krueger, C.C., Horrall, R.M., and Gruenthal, H. 1983. Strategy for the use of lake trout strains in
865 Lake Michigan. Wisconsin Department of Natural Resources, Administrative Report 17,
866 Madison.
867 Krueger, C.C., Shepherd, W.C., and Muir, A.M. 2014. Predation by alewife on lake trout fry
868 emerging from laboratory reefs: Estimation of fry survival and assessment of predation
869 potential. J. Great. Lakes. Res. **40**(2): 429-434.
870 doi:https://doi.org/10.1016/j.jglr.2014.01.009.
871 Krueger, C.C., Perkins, D.L., Mills, E.L., and Ellen Marsden, J. 1995. Predation by alewives on
872 lake trout fry in Lake Ontario: Role of an exotic species in preventing restoration of a
873 native species. J. Great. Lakes. Res. **21**: 458-469. doi:https://doi.org/10.1016/S0380-
874 1330(95)71117-0.
875 Landsman, S., Stein, J.A., Whitledge, G., and Robillard, S.R. 2017. Stable oxygen isotope
876 analysis confirms natural recruitment of Lake Michigan-origin lake trout (*Salvelinus*
877 *namaycush*) to the adult life stage. Fish Res. **190**: 15-23.
878 doi:https://doi.org/10.1016/j.fishres.2017.01.013.
879 Leonhardt, B.S., Happel, A., Bootsma, H., Bronte, C.R., Czesny, S., Feiner, Z., Kornis, M.S.,
880 Rinchard, J., Turschak, B., and Höök, T. 2020. Diet complexity of Lake Michigan
881 Salmonines: 2015–2016. J. Great. Lakes. Res.
882 doi:https://doi.org/10.1016/j.jglr.2020.05.010.

LMLTWG. 2019. Lake Michigan Lake Trout Working Group. 2019 Lake Michigan Lake Trout Working Group report. Great Lakes Fishery Commission, Lake Michigan Committee, Lake Michigan Technical Committee, Ann Arbor, Michigan.

Luo, M.K., Madenjian, C.P., Diana, J.S., Kornis, M.S., and Bronte, C.R. 2019. Shifting Diets of Lake Trout in Northeastern Lake Michigan. N. Am. J. Fish. Manage. **39**(4): 793-806. doi:10.1002/nafm.10318.

Madenjian, C.P., DeSorcie, T.J., and Stedman, R.M. 1998. Maturity schedules of lake trout in Lake Michigan. J. Great. Lakes. Res. **24**(2): 404-410. doi:https://doi.org/10.1016/S0380-1330(98)70831-7.

Madenjian, C.P., Bunnell, D.B., DeSorcie, T.J., Armenio, P., and Adams, J.V. 2018. Status and trends of prey fish populations in Lake Michigan, 2017. U.S. Geological Survey, Report to the Great Lakes Fishery Commission, Lake Michigan Committee, Sault Ste. Marie, Ontario.

Madenjian, C.P., O'Gorman, R., Bunnell, D.B., Argyle, R.L., Roseman, E.F., Warner, D.M., Stockwell, J.D., and Stapanian, M.A. 2008. Adverse effects of alewives on Laurentian Great Lakes fish communities. N. Am. J. Fish. Manage. **28**(1): 263-282. doi:10.1577/M07-012.1.

Marsden, E., Krueger, C.C., and 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. Am. J. Fish. Manage. **9**(3): 257-268. doi:10.1577/1548-8675(1989)009<0257:IOPOON>2.3.CO;2.

Marsden, J.E., Casselman, J.M., Edsall, T.A., Elliott, R.F., Fitzsimons, J.D., Horns, W.H., Manny, B.A., McAughey, S.C., Sly, P.G., and Swanson, B.L. 1995. Lake trout spawning

habitat in the Great Lakes — a review of current knowledge. J. Great. Lakes. Res. **21**: 487-497. doi:https://doi.org/10.1016/S0380-1330(95)71120-0.

McKee, P.C., Toneys, M.L., Hansen, M.J., and Holey, M.E. 2004. Performance of two strains of lake trout stocked in the Midlake Refuge of Lake Michigan. N. Am. J. Fish. Manage. **24**(4): 1101-1111. doi:10.1577/M03-142.1.

Miller, L.M., and Kapuscinski, A.R. 2003. Genetic guidelines for hatchery supplementation programmes. In E. M. Hallerman, ed. Population Genetics: Principles and Applications for Fisheries Scientists, pp. 329–355. American Fisheries Society, Bethesda, MD.

Miller, L.M., Mero, S.W., and Younk, J.A. 2009. The genetic legacy of stocking Muskellunge in a northern Minnesota lake. Trans. Am. Fish. Soc. **138**(3): 602-615. doi:10.1577/t08-114.1.

Muhlfeld, C.C., Kalinowski, S.T., McMahon, T.E., Taper, M.L., Painter, S., Leary, R.F., and Allendorf, F.W. 2009. Hybridization rapidly reduces fitness of a native trout in the wild. Biol. Lett. **5**(3): 328-331. doi:10.1098/rsbl.2009.0033.

Muir, A.M., Bronte, C.R., Zimmerman, M.S., Quinlan, H.R., Glase, J.D., and Krueger, C.C. 2014. Ecomorphological diversity of Lake Trout at Isle Royale, Lake Superior. Trans. Am. Fish. Soc. **143**(4): 972-987. doi:10.1080/00028487.2014.900823.

Nei, M., Tajima, F., and Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. J. Mol. Evol. **19**(2): 153-170. doi:10.1007/bf02300753.

Nielsen, E.E., Bach, L.A., and Kotlicki, P. 2006. HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. Mol. Ecol. Notes. **6**(4): 971-973. doi:10.1111/j.1471-8286.2006.01433.x.

928 Page, K.S., Scribner, K.T., and Burnham-Curtis, M. 2004. Genetic diversity of wild and hatchery
929 lake trout populations: Relevance for management and restoration in the Great Lakes.
930 Trans. Am. Fish. Soc. **133**(3): 674-691. doi:10.1577/T03-007.1.
931 Page, K.S., Scribner, K.T., Bennett, K.R., Garzel, L.M., and Burnham-Curtis, M.K. 2003.
932 Genetic assessment of strain-specific sources of lake trout recruitment in the Great Lakes.
933 Trans. Am. Fish. Soc. **132**(5): 877-894. doi:10.1577/T02-092.
934 Patterson, K.A., Stein, J.A., and Robillard, S.R. 2016. Progress toward Lake Trout rehabilitation
935 at a stocked and unstocked reef in southern Lake Michigan. N. Am. J. Fish. Manage.
936 **36**(6): 1405-1415. doi:10.1080/02755947.2016.1221000.
937 Peakall, R., and Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic
938 software for teaching and research-an update. Bioinformatics **28**(19): 2537-2539.
939 doi:10.1093/bioinformatics/bts460.
940 Peck, J.W. 1975. Brief life history accounts of five commercial salmonid fishes in Lake
941 Superior. Fisheries Research Report 1821. Michigan Department of Natural Resources,
942 Ann Arbor, Michigan.
943 Peck, J.W. 1988. Fecundity of hatchery and wild lake trout in Lake Superior. J. Great. Lakes.
944 Res. **14**(1): 9-13. doi:https://doi.org/10.1016/S0380-1330(88)71527-0.
945 Perreault-Payette, A., Muir, A.M., Goetz, F., Perrier, C., Normandeau, E., Sirois, P., and
946 Bernatchez, L. 2017. Investigating the extent of parallelism in morphological and
947 genomic divergence among lake trout ecotypes in Lake Superior. Mol. Ecol. **26**(6): 1477-
948 1497. doi:10.1111/mec.14018.

949 Philipp, D.P., and Whitt, G.S. 1991. Survival and growth of northern, Florida, and reciprocal Fl
950 hybrid largemouth bass in central Illinois. Trans. Am. Fish. Soc. **120**(1): 58-64.
951 doi:10.1577/1548-8659(1991)120<0058:SAGONF>2.3.CO;2.
952 Prince, D.J., O'Rourke, S.M., Thompson, T.Q., Ali, O.A., Lyman, H.S., Saglam, I.K., Hotaling,
953 T.J., Spidle, A.P., and Miller, M.R. 2017. The evolutionary basis of premature migration
954 in Pacific salmon highlights the utility of genomics for informing conservation. Science
955 Advances **3**(8). doi:10.1126/sciadv.1603198.
956 Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using
957 multilocus genotype data. Genetics **155**(2): 945-959.
958 Rahrer, J.F. 1965. Age, Growth, Maturity, and Fecundity of Humper Lake Trout, Isle Royale,
959 Lake Superior. Trans. Am. Fish. Soc. **94**(1): 75-83. doi:10.1577/1548-
960 8659(1965)94[75:AGMAFO]2.0.CO;2.
961 Rannala, B., and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes.
962 Proc. Natl. Acad. Sci. **94**(17): 9197-9201. doi:10.1073/pnas.94.17.9197.
963 Rogers, M.W., Markham, J.L., MacDougall, T., Murray, C., and Vandergoot, C.S. 2019. Life
964 history and ecological characteristics of humper and lean ecotypes of lake trout stocked
965 in Lake Erie. Hydrobiologia **840**(1): 363-377. doi:10.1007/s10750-019-03986-4.
966 Roseman, E.F., Stott, W., O'Brien, T.P., Riley, S.C., and Schaeffer, J.S. 2009. Heritage strain and
967 diet of wild young of year and yearling lake trout in the main basin of Lake Huron. J.
968 Great. Lakes. Res. **35**(4): 620-626. doi:https://doi.org/10.1016/j.jglr.2009.08.014.
969 Rousset, F. 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for
970 Windows and Linux. Mol. Ecol. Resour. **8**(1): 103-106. doi:10.1111/j.1471-
971 8286.2007.01931.x.

Royce, W.F. 1951. Breeding habits of lake trout in New York. U.S. Fish Wildl. Serv. Fish. Bull. 52: 59-76.

Ruzich, J., Turnquist, K., Nye, N., Rowe, D., and Larson, W.A. 2019. Isolation by a hydroelectric dam induces minimal impacts on genetic diversity and population structure in six fish species. Conserv. Genet. **20**(6): 1421-1436. doi:10.1007/s10592-019-01220-1.

Salvesen, K.E. 2015. Lake trout restoration in the Great Lakes: from hatchery to natural repoduction. Mater's Thesis: Pennsylvania State University.

Schneeberger, P., Toneys, M., Elliott, R., Jonas, J., Clapp, D., Hess, R., and Passino-Reader, D. 1998. Lakewide assessment plan for Lake Michigan fish communities. Ann Arbor, Michigan: Great Lakes Fishery Commission, Lake Michigan Technical Committee. Available: www. glfc. org/pubs/SpecialPubs/lwasses01. pdf. (Accessed February 2020).

Schneider, C.P., Owens, R.W., Bergstedt, R.A., and O'Gorman, R. 1996. Predation by sea lamprey (*Petromyzon marinus*) on lake trout (*Salvelinus namaycush*) in southern Lake Ontario, 1982-1992. Can. J. Fish. Aquat. Sci. **53**(9): 1921-1932. doi:10.1139/f96-129.

Scribner, K., Tsehaye, I., Brenden, T., Stott, W., Kanefsky, J., and Bence, J. 2018. Hatchery strain contributions to emerging wild lake trout populations in Lake Huron. J. Hered. **109**(6): 675-688. doi:10.1093/jhered/esy029.

Seamons, T.R., Hauser, L., Naish, K.A., and Quinn, T.P. 2012. Can interbreeding of wild and artificially propagated animals be prevented by using broodstock selected for a divergent life history? Evol. Appl. **5**(7): 705-719. doi:10.1111/j.1752-4571.2012.00247.x.

Sierszen, M.E., Hrabik, T.R., Stockwell, J.D., Cotter, A.M., Hoffman, J.C., and Yule, D.L. 2014. Depth gradients in food-web processes linking habitats in large lakes: Lake Superior as an exemplar ecosystem. Freshwater Biology **59**(10): 2122-2136. doi:10.1111/fwb.12415.

995 Sitar, S., Goetz, F., Jasonowicz, A., and Seider, M. 2020. Lipid levels and diet compositions in
996 lake charr ecotypes at Isle Royale in northern Lake Superior. J. Great. Lakes. Res. **46**(3):
997 569-577. doi:https://doi.org/10.1016/j.jglr.2020.03.001.
998 Smith, S.H. 1968. Species succession and fishery exploitation in the Great Lakes. Journal of the
999 Fisheries Research Board of Canada 25:667-693.
1000 Takezaki, N., Nei, M., and Tamura, K. 2010. POPTREE2: software for constructing population
1001 trees from allele frequency data and computing other population statistics with Windows
1002 interface. Mol. Biol. Evol. **27**(4): 747-752. doi:10.1093/molbev/msp312.
1003 USFWS, and GLFC. 2017. USFWS/GLFC 2017. Great Lakes Fish Stocking Database. U.S. Fish
1004 and Wildlife Service, Region 3 Fisheries Program and Great Lakes Fishery Commission,
1005 Ann Arbor, Michigan. Available at http://www.glfc.org/fishstocking/.
1006 Vaha, J.-P., and Primmer, C.R. 2006. Efficiency of model-based Bayesian methods for detecting
1007 hybrid individuals under different hybridization scenarios and with different numbers of
1008 loci. Mol. Ecol. **15**(1): 63-72. doi:10.1111/j.1365-294X.2005.02773.x.
1009 Van Offelen, H.K., Krueger, C.C., and Schofield, C.L. 1993. Survival, growth, movement, and
1010 distribution of two brook trout strains stocked into small Adirondack streams. N. Am. J.
1011 Fish. Manage. **13**(1): 86-95. doi:10.1577/1548-8675(1993)013<0086:Sgmado>2.3.Co;2.
1012 Vinson, M.R., Hoffmann, J.M., Muir, A.M., Rosinski, C.L., Krueger, C.C., Bronte, C.R.,
1013 Hansen, M.J., Sitar, S.P., Allen, E.W., Baker, L.F., and Swanson, H.K. in press. Gut
1014 contents from multiple morphs of lake trout (*Salvelinus namaycush*) at two offshore
1015 shoals in Lake Superior. J. Great. Lakes. Res.
1016 Waters, C.D., Hard, J.J., Brieuc, M.S.O., Fast, D.E., Warheit, K.I., Waples, R.S., Knudsen, C.M.,
1017 Bosch, W.J., and Naish, K.A. 2015. Effectiveness of managed gene flow in reducing

genetic divergence associated with captive breeding. Evol. Appl. **8**(10): 956-971. doi:10.1111/eva.12331.

Weir, B.S., and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population-structure. Evolution **38**(6): 1358-1370. doi:10.2307/2408641.

Wells, L., and McLain, A.L. 1973. Lake Michigan: man's effects on native fish stocks and other biota. Great Lakes Fisheries Commission, Technical Report No. 20.

Zimmerman, M.S., Krueger, C.C., and Eshenroder, R.L. 2007. Morphological and Ecological Differences Between Shallow-and Deep-water Lake Trout in Lake Mistassini, Quebec. J. Great. Lakes. Res. **33**(1): 156-169. doi:https://doi.org/10.3394/0380-1330(2007)33[156:MAEDBS]2.0.CO;2.

Figure legends

Figure 1. Stocking history of lake trout strains in Lake Michigan. The Apostle Island, Isle Royale, and Marquette strains are derived from the Lake Superior lean (i.e. shallow water) ecomorphotype, the Klondike strain is derived from the Lake Superior humper (i.e. deeper water) ecomorphotype and is the only strain not derived from the lean ecomorphotype, Clearwater Lake is derived from Clearwater Lake, Manitoba, Canada (Latitude: 54.0438, Longitude: -101.1262), Green Lake and Lewis Lake are derived from remnant Lake Michigan lake trout, Parry Sound is derived from Lake Huron, and Seneca Lake is derived from Seneca Lake, NY. Lake trout stocked at different sizes were standardized to yearling equivalents (i.e. the size of average yearling lake trout) to facilitate comparisons see (Elrod et al. 1988). Data were available from (USFWS and GLFC 2017).

Figure 2. Neighbor-joining dendrogram of reference strains based on Nei's D_A genetic distance. Bootstrap support for each node is shown. See Table 1 for collection information.

Figure 3. Results from STRUCTURE clustering analysis of reference strains and unknown (i.e. mixed-origin) samples at K = 6. Reference strains were used to assign unknown individuals to their strain of origin. Black vertical lines denote reference strains and geographic sampling strata for unknown samples (see Tables 1,4). Strain abbreviations are Seneca (SN), Parry Sound (PR), Lewis Lake (LW), Green Lake (GN), Klondike Reef (KN), Marquette (MR), Apostle Island (AP), Isle Royale (IR). Strata abbreviations for mixed-origin samples are North (NO), Traverse Bay (TR), Northwest (NW), East (EA), Southwest (SW), Southern Refuge (SR), and Illinois/Indiana (II).

Figure 4. Estimates of contributions of reporting groups (Table 1) for hatchery strains to naturally produced lake trout for seven geographic strata in Lake Michigan. Sample sizes are below each pie. Colors represent reference strains. See Table 5 for proportion data and confidence intervals. See the Methods section for information on the composition of each reporting group. Statistical districts used to form strata are visualized in Fig. 1 of Kornis et al. (2019b).

Figure 5. Heatmap of observed versus expected strain proportions for seven geographic strata in Lake Michigan (see Fig. 4 for map of strata). Observed proportions were calculated from genetic data and expected proportions were calculated from stocking data (see Table 5). Positive values indicate a given strain is overrepresented in the genetic data.

Figure 6. Heatmaps of observed versus expected proportions of pure and hybrid crosses assuming random mating for seven geographic strata in Lake Michigan. Observed proportions for both panels (a) and (b) were the proportions of each cross type observed in the genetic data. See Table S8 for proportions and confidence intervals. Expected proportions for panel (a) were calculated based on stocking data (i.e. composite estimate of expected strain compositions reconstructed from stocking records) and expected proportions for panel (b) were calculated based on overall genetic proportions observed in each stratum (see Table 5). Positive values indicate a given cross type is overrepresented in the genetic data. See Table S6 for proportion data.

Figure S1. Age composition of naturally produced lake trout analyzed. Ages were estimated with otoliths.

Fig. S2. Year classes of naturally produced lake trout analyzed.

Tables

Table 1. Information on eight lake trout reference strains. Seneca Lake is a finger lake that drains into Lake Ontario through the Oswego River, and Lewis Lake and Green Lake contain lake trout that are descended from wild fish captured from Lake Michigan before lake trout were extirpated from this system. The other strains were all collected from within the Great Lakes (either lakes Huron or Superior. Reporting groups were determined using STRUCTURE analysis (Fig. 2) and assignment tests with simulated individuals (Table 3). N is the number of individuals successfully genotyped. Other abbreviations are observed heterozygosity (H_O), expected heterozygosity (H_E), number of alleles (A), and allelic richness (A_R). Latitude and longitude represent the source of each strain.

Strain	Lake/Drainage	Reporting Group	Latitude	Longitude	N	H_O	H_E	A	A_R
Seneca Lake	Ontario	Seneca Lake	42.680223	-76.915079	81	0.57	0.57	7.44	6.62
Parry Sound	Huron	Parry Sound	45.350004	-80.055935	41	0.55	0.54	6.94	6.86
Lewis Lake	Michigan	Lewis Lake	44.304252	-110.63071	97	0.56	0.56	8.03	7.16
Green Lake	Michigan	Green Lake	43.830280	-88.966445	71	0.56	0.56	7.72	7.05
Klondike Reef	Superior	Superior	47.340482	-85.801163	61	0.55	0.56	7.94	7.33
Marquette	Superior	Superior	46.747705	-87.226454	72	0.55	0.58	9.83	8.74
Apostle Islands	Superior	Superior	46.898999	-90.663642	77	0.56	0.57	10.36	8.79
Isle Royale	Superior	Superior	47.821319	-89.162644	62	0.57	0.59	9.19	8.34

Table 2. Sample information for naturally produced lake trout genotyped in this study. Districts are statistical harvest districts for Lake Michigan and are visualized in Fig. 1 of Kornis et al. (2019b). See Fig. 4 for a map of strata and Table S7 for metadata on each fish including date and location of capture, length, age, weight (when available), and sex (when available). The single fish assigned to a strain not stocked in Lake Michigan (Parry Sound x Superior hybrid) is not included in this table (see Results). Traverse Bay refers Grand Traverse Bay.

Strata	Districts included	N total	2009-2011	2012	2013	2014	2015
North	MM-1, MM-2, MM-3	216	9	20	66	54	67
Traverse Bay	MM-4	85	0	46	37	2	0
Northwest	WM-3, WM-4	102	0	27	14	32	29
East	MM-5, MM-6, MM-7, MM-8	77	0	0	2	0	75
Southwest	WM-5, WM-6	91	0	5	9	21	56
Southern Refuge	WM-5, MM-6, MM-7	116	0	0	0	0	116
Illinois/Indiana	ILL, IND	160	0	0	0	0	160
Total		847	9	98	94	109	503

Table 3. Pairwise F_{ST} values for reference strains calculated using 36 microsatellites. Bold values are significantly different from zero ($P < 0.01$).

Population	Seneca Lake	Parry Sound	Lewis Lake	Green Lake	Klondike Reef	Marquette	Apostle Islands
Parry Sound	**0.055**						
Lewis Lake	**0.058**	**0.046**					
Green Lake	**0.059**	**0.058**	**0.040**				
Klondike Reef	**0.075**	**0.083**	**0.049**	**0.063**			
Marquette	**0.048**	**0.043**	**0.018**	**0.026**	**0.033**		
Apostle Islands	**0.064**	**0.055**	**0.029**	**0.035**	**0.034**	0.006	
Isle Royale	**0.059**	**0.050**	**0.028**	**0.033**	**0.038**	**0.008**	**0.009**

Table 4. Assignment accuracy for five pure and ten hybrid cross types representing all possible combinations among reporting groups. For each cross type, 1,000 individuals were simulated and assigned to reference strains using STRUCTURE (see methods). N correct is the number of individuals correctly assigned out of 1,000. Classified pure is the number of individuals classified as a pure cross (i.e., not a hybrid) out of 1,000. Values in bold are cross types with correct assignment < 0.8. We also calculated mixture proportions ("mixture" columns) using STRUCTURE with the approach outlined in the methods section and ONCOR. The expected values are 100% for pure cross types and 50% for each strain in hybrid cross types. The Klondike strain was removed from this analysis because it produced low assignment accuracy (63% correct with STRUCTURE analysis), with high misassignment to the Superior reporting group. The Klondike strain was combined with other Superior strains for the analysis of unknown individuals. See Table 1 for information on strains.

Cross type	N correct	% correct	Classified pure	Mixture STRUCTURE	Mixture ONCOR
Pure					
Seneca	977	98%	977	99%	100%
Parry	969	97%	969	98%	100%
Lewis	942	94%	942	97%	100%
Green	912	91%	912	96%	100%
Superior	749	**75%**	749	87%	100%
Hybrid					
Seneca x Parry	912	91%	30	49%, 48%	49%, 48%
Seneca x Lewis	882	88%	26	49%, 47%	44%, 41%
Seneca x Green	854	85%	29	49%, 45%	48%, 31%
Seneca x Superior	833	83%	31	48%, 45%	29%, 71%
Parry x Lewis	834	83%	38	47%, 46%	38%, 51%
Parry x Green	806	81%	14	46%, 44%	40%, 29%
Parry x Superior	803	80%	43	46%, 45%	32%, 68%
Lewis x Green	779	**78%**	36	46%, 44%	38%, 35%
Lewis x Superior	790	**79%**	143	50%, 46%	24%, 76%
Green x Superior	808	81%	78	45%, 49%	28%, 72%

1128 Table 5. Observed (obs) and expected (exp) proportions of four genetic strains in seven
1129 geographic strata across Lake Michigan. Observed proportions are the genetic proportions of
1130 each stock calculated from STRUCTURE analysis and expected proportions were calculated
1131 from stocking data. 95% confidence intervals for genetic (i.e. observed) estimates are below
1132 point estimates. See Table 2 and Fig. 3 for more information on strata and Table 1 for more
1133 information on strains.

Strata	Seneca obs	Seneca exp	Lewis obs	Lewis exp	Green obs	Green exp	Superior obs	Superior exp
North	0.69 0.55-0.83	0.03	0.14 0.08-0.22	0.34	0.01 0.00-0.04	0.01	0.16 0.09-0.26	0.62
Traverse Bay	0.26 0.15-0.41	0.14	0.52 0.35-0.72	0.51	0.04 0.01-0.12	0.02	0.18 0.08-0.33	0.33
Northwest	0.53 0.38-0.70	0.19	0.19 0.10-0.32	0.22	0.04 0.01-0.12	0.22	0.24 0.13-0.4	0.38
East	0.69 0.51-0.87	0.20	0.12 0.05-0.25	0.23	0.06 0.02-0.18	0.20	0.13 0.05-0.28	0.36
Southwest	0.55 0.39-0.72	0.26	0.07 0.02-0.16	0.17	0.10 0.04-0.21	0.36	0.27 0.15-0.46	0.21
Southern Refuge	0.58 0.43-0.74	0.37	0.09 0.04-0.18	0.06	0.16 0.09-0.29	0.43	0.17 0.09-0.31	0.14
Illinois/Indiana	0.72 0.57-0.87	0.17	0.06 0.03-0.13	0.27	0.12 0.06-0.21	0.35	0.10 0.05-0.19	0.20

1134

1135 Table S1. Multiplexing and primer information for microsatellite loci genotyped in this study.

1136

1137 Table S2. Fecundity relative to the minimum age of maturity (Age 8 for Southern Refuge, Age 6
1138 for all others).

1139

1140 Table S3. Movement matrix used in expected strain proportion calculations.

1141

1142 Table S4. Summary statistics for microsatellite loci genotyped in this study.

1143

Table S5. Genotype data from all individuals in the study in GenAlEx format. Individuals denoted with "UNK" were naturally reproduced fish of unknown origin. Metadata for these fish are found in Table S4.

Table S6. Structure Harvester results. The K used for all analyses (K=6) is highlighted.

Table S7. Metadata for each naturally reproduced fish analyzed in the study. Fish with two strains separated with an "_" are putative hybrids. Year class is catch year – otolith age.

Table S8. Observed and expected proportions of cross types (i.e. pure and hybrid crosses) for seven geographic strata in Lake Michigan. Observed genetic data is the proportion of each type inferred directly from the genetic data, expected stocking is the proportion of each cross type expected under random mating given the stocking proportions calculated in Table 5, and expected is the proportion of each cross type expected under random mating given the genetic proportions calculated in Table 5. 95% confidences intervals are included in parentheses for observed genetic data.