PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Yin X, Schraidt CE, Sparks MM, Euclide PT, Hoyt TJ, Ruetz III CR, Höök TO, Christie MR. 2025 Parallel genetic adaptation amid a background of changing effective population sizes in divergent yellow perch (*Perca flavescens*) populations. *Proc. R. Soc. B* **292**: 20242339. https://doi.org/10.1098/rspb.2024.2339

Received: 21 April 2024
Accepted: 18 November 2024

Subject Category:
Evolution

Subject Areas:
evolution, genetics, genomics

Keywords:
demographic history, effective population size, fisheries, genetic adaptation, genetic diversity, parallel evolution

Authors for correspondence:
Xiaoshen Yin
e-mail: yinxiaoshen@ouc.edu.cn
Mark R. Christie
e-mail: christ99@purdue.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.7597090.

THE ROYAL SOCIETY PUBLISHING

Parallel genetic adaptation amid a background of changing effective population sizes in divergent yellow perch (*Perca flavescens*) populations

Xiaoshen Yin[1,2], Claire E. Schraidt[3,4], Morgan M. Sparks[2,5], Peter T. Euclide[3,6], Tyler J. Hoyt[7], Carl R. Ruetz III[7], Tomas O. Höök[3,6] and Mark R. Christie[2,3]

[1]MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao 266003, People's Republic of China
[2]Department of Biological Sciences, and [3]Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907-2054, USA
[4]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06511, USA
[5]US Forest Service, Rocky Mountain Research Station, Boise, ID 83702, USA
[6]Illinois-Indiana Sea Grant, Purdue University, West Lafayette, ID 47907, USA
[7]Annis Water Resources Institute, Grand Valley State University, Muskegon, MI 49441, USA

XY, 0000-0002-2687-4473; CRRIII, 0000-0002-9336-5118; MRC, 0000-0001-7285-5364

Aquatic ecosystems are highly dynamic environments vulnerable to natural and anthropogenic disturbances. High-economic-value fisheries are one of many ecosystem services affected by these disturbances, and it is critical to accurately characterize the genetic diversity and effective population sizes of valuable fish stocks through time. We used genome-wide data to reconstruct the demographic histories of economically important yellow perch (*Perca flavescens*) populations. In two isolated and genetically divergent populations, we provide independent evidence for simultaneous increases in effective population sizes over both historic and contemporary time scales including negative genome-wide estimates of Tajima's D, 3.1 times more single nucleotide polymorphisms than adjacent populations, and contemporary effective population sizes that have increased 10- and 47-fold from their minimum, respectively. The excess of segregating sites and negative Tajima's D values probably arose from mutations accompanying historic population expansions with insufficient time for purifying selection, whereas linkage disequilibrium-based estimates of *Ne* also suggest contemporary increases that may have been driven by reduced fishing pressure or environmental remediation. We also identified parallel, genetic adaptation to reduced visual clarity in the same two habitats. These results suggest that the synchrony of key ecological and evolutionary processes can drive parallel demographic and evolutionary trajectories across independent populations.

1. Introduction

Accurate estimates of effective population sizes (*Ne*) are critical for the successful management and conservation of exploited, threatened and endangered species. The concept of *Ne* was first introduced in 1931 and often refers to the size of an idealized population with constant size and random mating that experiences an equivalent rate of genetic drift as the population under study [1]. Consequently, estimates of effective population sizes can often be smaller than estimates of census population sizes (*Ne* : *N* ratio ranging from 10^{-5} to approx. 0.5 [2–13]). From a conservation and management perspective, effective population size matters in several key aspects: (i)

Ne is a major determinant of nucleotide variation (i.e. $4Ne\mu$ under a neutral model); (ii) *Ne* measures the expected rate of genetic drift, a major stochastic evolutionary force acting on genetic diversity across the genome and through generations; and (iii) *Ne* provides a unified frame of reference for comparisons across species or populations [14,15]. Thus, changes in effective population sizes through time can provide information on how populations are responding to novel stressors, such as climate change [16–19], changes in food web dynamics [20], habitat alteration and destruction [17,20], and overexploitation [16–18,21,22].

Accurate estimates of effective population sizes are particularly needed in fisheries to guide sustainable harvest practices in economically important populations. *Ne* is a critical metric for the successful assessment and management of fisheries stocks as it not only provides a fisheries-independent approach, avoiding the biases inherent in some field- or catch-based methods [23] but also reflects the adaptative potential of different stocks under constantly changing environments [23,24]. Estimates of effective population sizes may forecast overharvest and potential collapses in commercial fisheries, as signalled by either substantial reductions in effective population sizes or increasing ratios of effective to census population sizes (i.e. overharvest reduces both effective and census population sizes but reduces census population size more quickly, so the ratio of effective to census population size goes up [4,25–27]). Nevertheless, several challenges remain for estimating effective population sizes in fish and fisheries. Estimates of *Ne* often include infinity when the signal for genetic variation arising from genetic drift cannot be estimated with high precision and is instead overwhelmed by noise arising from sampling error [28,29]. Also, fish populations are often highly connected due to extensive larval dispersal and juvenile or adult migration [30–33] and such migration and subsequent gene flow can bias, or at the very least complicate estimates of *Ne* [14,34,35]. Additional challenges include overlapping generations [34–36], varied genomic architectures [34], rare alleles [34], small sample sizes [34,35], sampling design [34] and sequencing or genotyping errors [34]. Despite these challenges, advances in high-throughput sequencing and the influx of genomic data have resulted in the development of many methods [37–41] that more accurately infer temporal changes in effective population sizes for various data types, including polymorphic sites, short sequence blocks and whole-genome data [14].

Empirical assessments of *Ne* over contemporary time scales (e.g. decades to centuries) have largely demonstrated reductions in effective population sizes [16–22,42–45]. Although rarely documented [6,45,46], contemporary increases in effective population sizes exist and can occur via one of three mechanisms. First, a reduction in the variance in reproductive success, for example, by equalizing mating success throughout a single generation, can result in larger effective population sizes in the next generation [47,48], as demonstrated in captive breeding programs where equalization of full-sib family sizes is often recommended to increase effective population sizes [49–52]. Second, introgression from neighbouring, genetically divergent populations may generate biases in estimating *Ne*, the direction of which is determined by the magnitude and continuity of gene flow [34,35]. Third, large, consistent increases in adult abundance within a population, perhaps due to improved environmental conditions and management policies, can also cause increases in effective population sizes [46]. Over longer time scales, these last two mechanisms may also be accompanied by several other genetic indicators including negative values of Tajima's D, shifts in site frequency spectra and larger numbers of polymorphic loci with low minor allele frequency (MAF) [14].

In this study, we explored changes in the effective population sizes and patterns of adaptive divergence in yellow perch (*Perca flavescens*) populations found throughout the main basin of Lake Michigan and two connected water bodies, Green Bay and Muskegon Lake (figure 1). Lake Michigan, with a drainage basin of about 118 000 km^2, is the fourth largest freshwater lake in the world [53,54]. Green Bay, an inlet of northwestern Lake Michigan, drains approximately 40 468 km^2 while Muskegon Lake, a drowned river mouth (DRM) lake of 16.8 km^2, receives water from the Muskegon River and drains directly into Lake Michigan. DRM lakes formed as the mouths of rivers were cut off from flowing into the Great Lakes following the retreat of glaciers. The melting river water was trapped behind land formations, carving out basins like Muskegon Lake and eventually reconnecting after the buildup of water pressure created channels allowing them to once again flow into the Great Lakes [55,56]. The main basin of Lake Michigan, Green Bay and Muskegon Lake are all home to yellow perch but are vastly different in both abiotic (e.g. water clarity) and biotic (e.g. community composition) characteristics and have distinct environmental histories. Yellow perch migrate between nearshore Lake Michigan and Muskegon Lake seasonally, but they do not appear to interbreed with resident populations [57,58]. Previous work in this system using a RAD-Seq dataset revealed that yellow perch in the main basin of Lake Michigan are genetically divergent from Green Bay [59,60]; however, the demographic and evolutionary histories of yellow perch from these three highly differentiated habitats have not yet been investigated [57–60].

Yellow perch remain one of the most ecologically and economically important species throughout the Great Lakes. They are an abundant nearshore fish species, serve as important predators of small fishes and invertebrates and are themselves important prey for larger fishes (see 'Yellow perch life history' in electronic supplementary material for relevant life-history information) [61]. Historically, yellow perch supported commercially important fisheries throughout the Great Lakes region; in Lake Michigan alone, peak annual commercial harvest would now represent close to US$16 million in dockside value and much more at retail. However, Lake Michigan yellow perch populations began declining during the late 1980s and early 1990s, which led to closures of most yellow perch commercial fisheries. Here, we sequenced the complete genomes and reconstructed the demographic histories for 210 yellow perch collected from seven sample sites encompassing the main basin of Lake Michigan, Green Bay and Muskegon Lake (figure 1a). Using genome-wide data, we ask the following questions. (i) What are the differences between genetic differentiation and genetic diversity among sample sites? (ii) what are the differences in *Ne* among sample sites using multiple approaches that can estimate both historical and more recent effective population sizes? (iii) What genes show signals of responding to selection in the three environmentally distinct regions (i.e. main basin, Green Bay and Muskegon Lake)? We identify three highly discrete populations, provide evidence for historic and contemporary increases in *Ne*, and find signatures of parallel, genetic adaptation.

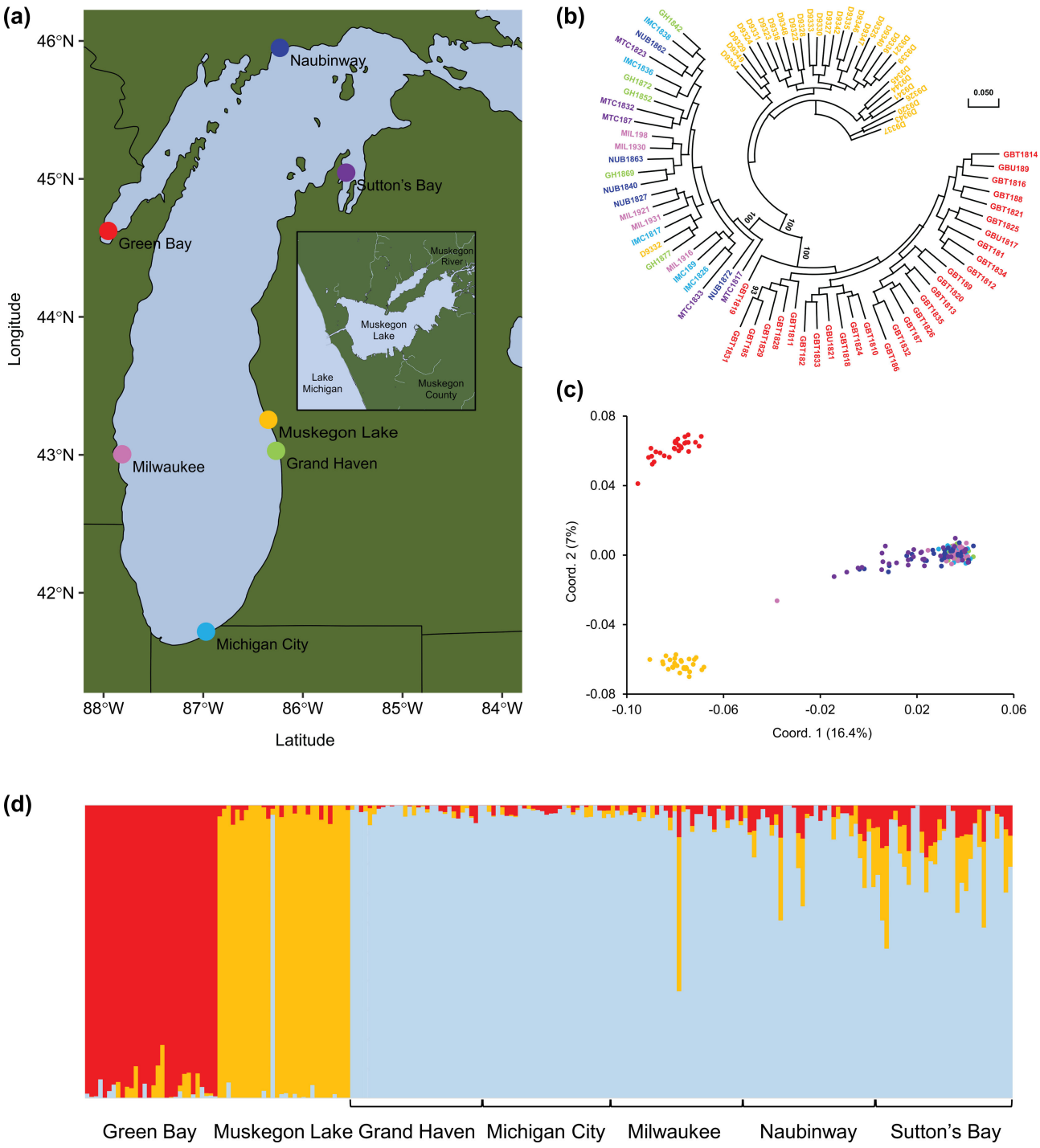

Figure 1. Population structure of Lake Michigan yellow perch populations. (a) Yellow perch collected from seven Lake Michigan sites, including the comparatively isolated Muskegon Lake, which is connected to the main basin by a narrow channel (see inset). (b) A phylogenetic tree illustrates three distinct groups: Green Bay, Muskegon Lake and the five main basin sample sites. Twenty-five main basin individuals were randomly sampled for visual clarity; see electronic supplementary material, figure S1 for the full phylogenetic tree with all 210 individuals. (c,d) Both principal coordinate analysis of 210 individuals (c) and admixture analyses (d) further support the delineation of samples into three discrete populations. Note that one individual sampled in Muskegon Lake probably swam in from the main basin (electronic supplementary material, figure S2; see 'Genetic divergence among yellow perch sample sites' in electronic supplementary material).

2. Methods

(a) Data collection

We sampled 210 adult yellow perch from seven locations (Green Bay, Muskegon Lake, Grand Haven, Michigan City, Milwaukee, Naubinway and Sutton's Bay; figure 1a; electronic supplementary material, table S1) in 2016, 2018 or 2019 using fyke netting, night-time boat electrofishing, bottom trawls, gill netting and creel surveys and conducted whole-genome resequencing to a depth of 12.50× after aligning to the yellow perch genome [62] (ranging from 5.93× to 18.58×; see 'Whole-genome sequencing' in electronic supplementary material). We called single nucleotide polymorphisms (SNPs) using Genome Analysis Toolkit (GATK) 4.1.9.0 [63] (see 'Variant calling and filtering' in electronic supplementary material). We filtered SNPs by MAF in two ways to create two datasets of SNPs suitable for different analyses (electronic supplementary material, table S2): a study-wide

dataset and within-group dataset (see 'Variant calling and filtering' in electronic supplementary material) [64]. For population structure analyses, we used the study-wide dataset, whereas we used the within-group datasets for all subsequent analyses (electronic supplementary material, table S2).

(b) Population structure analysis

To infer the phylogenetic relationship among the seven sample sites, we constructed phylogenetic trees with 5000 randomly selected SNPs using RAxML 8.2.12 [65] and viewed the phylogenetic tree using MEGA 11 [66] (see 'Population structure analysis' in electronic supplementary material). To further investigate the population structure, we first conducted principal coordinate analysis (PCA) on all individuals with 50 000 randomly selected SNPs in *hierfstat* in R 4.2.1 [67]. Next, we conducted PCA with pairwise F_{ST} using GenAlEx 6.51b2 [68]. Weir & Cockerham's F_{ST} was calculated for each pair of sample sites using VCFtools 0.1.16 [69]. Last, we estimated the ancestry for Green Bay, Muskegon Lake and the five main basin sample sites, separately and jointly, with 10 000 randomly selected SNPs using Admixture 1.3 [70] by setting K from 1 to 10.

(c) Minor allele frequency, Tajima's D and genetic diversity

Using the within-group datasets, we calculated minor allele frequencies, Tajima's D and genetic diversity: (i) we calculated minor allele frequencies for seven sample sites and viewed its distribution by plotting density curves with unit bins of 0.015 with ggplot2 [71] and ggridges in R 4.2.1 [67], (ii) we calculated Tajima's D by setting the bin size to 5 MB with the within-group dataset filtered by MAF ≥ 0.02 (electronic supplementary material, table S2) using VCFtools 0.1.16, and (iii) we estimated genetic diversity by calculating observed heterozygosity for every SNP (not including invariant sites) and averaged observed heterozygosity across overlapping windows of 5 MB using a step size of 2.5 MB across each chromosome. We also calculated nucleotide diversity (π) for each population with biallelic sites that are covered by at least three reads in each sample (including both polymorphic and invariant sites) using PIXY [72].

(d) Effective population size and demographic history

Using the within-group datasets (electronic supplementary material, table S2), we estimated effective population size with 30 000 randomly selected SNPs (replicated three times per population) for each population using NeEstimator v2 [28] (see 'Effective population size and demographic history' in electronic supplementary material). To further validate the *Ne* estimates, we estimated effective population size with 3000 randomly selected SNPs from a previously obtained RAD-Seq dataset [59], again replicated three times per sample site for each of 12 sample sites (see 'Effective population size and demographic history' in electronic supplementary material), using the same options described above. Returning to our whole-genome datasets, we reconstructed demographic histories by running GONE [39] with all SNPs in the within-group datasets 100 times for each sample site. Specifically, we reconstructed the demographic history for Green Bay and Muskegon Lake by randomly selecting 30 000 SNPs from the within-group dataset for each sample site (replicated 1000 times per sample site; electronic supplementary material, table S2) and running GONE [39] for each of 1000 randomly selected sets of 30 000 SNPs for each sample site using the default parameters. To examine changes in *Ne* over longer time scales, we also ran Stairway Plot 2 [38], for each sample site, with all biallelic SNPs at which genotype information is available for all samples in each of the seven sample sites, where generation time was estimated as 4 ± 1 yr (see 'Yellow perch life history' in electronic supplementary material).

(e) Candidate genes and gene ontology hierarchy networks underlying rapid adaptation

To identify genomic regions with overlapping windows of high F_{ST}, we first calculated pairwise Weir and Cockerham's F_{ST} with separate VCF files for 10 pairs of populations using VCFtools (0.1.16) [69] and identified outlier windows from a minimum of eight out of the ten pairwise population comparisons sharing at least one high F_{ST} SNP. We mapped SNPs on outlier windows to the yellow perch genome [62] to identify candidate outlier genes potentially explaining parallel genetic adaptation in Green Bay and Muskegon Lake in comparison with main basin populations. We annotated candidate outlier genes using UniProtKB ID via UniProt (https://www.uniprot.org/), with which we retrieved gene ontology (GO) terms for these outlier genes from Gene Ontology and GO Annotations (https://www.ebi.ac.uk/QuickGO/). Last, we constructed GO hierarchy networks of GO terms associated with at least two outlier genes in the domain of 'biological process' for each pairwise comparison (i.e. Green Bay or Muskegon Lake versus main basin populations) using *metacoder* [73].

(f) Water clarity, water quality and phenotypic measurements

Because many of the outlier genes identified from the previous analysis were related to vision, we investigated differences in water clarity, water quality and yellow perch eye size between the main basin, Green Bay and Muskegon Lake sample sites. Differences in eye size among those sample sites was not an *a priori* hypothesis but was rather an *a posteriori* hypothesis guided by our candidate genes (see Discussion). Water clarity was estimated using Secchi depth readings collected between 1 June and 30 August from Green Bay (320 readings), Muskegon Lake (72 readings) and the southern basin of Lake Michigan (180 measurements) between 2010 and 2022 (see 'Water clarity, water quality and phenotypic measurements' in electronic

supplementary material). We conducted an analysis of variance (ANOVA) to test differences in Secchi depths among Green Bay, Muskegon Lake and the southern basin of Lake Michigan.

To determine whether eye diameter differed among Green Bay, Muskegon Lake and Lake Michigan, we collected yellow perch from the three locations, measured their eye diameter and analysed yellow perch with total lengths of 150–250 mm to ensure comparisons of yellow perch with similar sizes among the three locations (Green Bay: n = 31, Muskegon Lake: n = 39, Grand Haven: n = 27; see 'Water clarity, water quality and phenotypic measurements' in electronic supplementary material). With this subset of fish, we used analysis of covariance (ANCOVA) to test whether eye diameter differed among the three locations while accounting for total length as a covariate (including interaction between location and covariate). We compared least-squares means with Tukey's method using the 'emmeans' function in the *emmeans* package in R 4.2.1 [67].

3. Results

(a) Population structure

We mapped an average of 99.6% of 15.8 billion, 150 bp paired-end reads to the yellow perch genome [62] and, on average, over 90% and 83% of reads had a mapping quality equal to or greater than 20 or 60, respectively (electronic supplementary material, table S3). The study-wide dataset, which was used for population genetic and stock identification analyses, resulted in 714 666 SNPs per sample site while the within-group dataset, which was used for demographic analyses, resulted in an average of 1 160 504 SNPs with less than 0.9% of SNPs out of Hardy–Weinberg equilibrium in each sample site (electronic supplementary material, tables S2 and S4). After inferring phylogenetic relationships among populations, we found that Green Bay, Muskegon Lake and the main basin sample sites (i.e. Grand Haven, Michigan City, Milwaukee, Naubinway and Sutton's Bay) form three distinct lineages (figure 1b; electronic supplementary material, figure S1). By conducting PCA with all individuals, we found further support that Green Bay, Muskegon Lake and the main basin sample sites are clearly separated from each other (figure 1c; electronic supplementary material, figure S2a–c), except for one Muskegon Lake sample clustering with main basin populations and one Milwaukee main basin sample falling comparatively further away from the main basin populations (figure 1c; electronic supplementary material, figure S2d–f, and 'Yellow perch life history' and 'Genetic divergence among yellow perch sample sites'). This clear delineation of three populations was further supported by PCA on the mean pairwise estimates of F_{ST} where we detected high genetic differentiation between Green Bay and Muskegon Lake (F_{ST} = 0.0520), both of which are genetically divergent from the five main basin sample sites (F_{ST} = 0.0550–0.0667, 0.0460–0.0592, respectively; electronic supplementary material, table S5 and figure S3a). In contrast, the five main basin sample sites are comparatively similar to each other, with mean pairwise F_{ST} ranging from −0.0002 to 0.0041 (electronic supplementary material, table S5 and figure S3b). The optimal number of clusters/populations identified with admixture analysis [70] for all populations was 3: Green Bay, Muskegon Lake and the main basin sample sites (figure 1d; electronic supplementary material, figure S4).

(b) Numbers of SNPs, minor allele frequency and Tajima's D

When using the within-group dataset, we detected 2.8–3.4 times more polymorphic SNPs in Green Bay and Muskegon Lake sample sites, respectively, in comparison with the main basin sample sites (2 186 768 and 2 275 935 versus 670 578 to 776 088 SNPs in main basin populations; figure 2a). The median MAF of SNPs was 0.086 and 0.083 in Green Bay and Muskegon Lake, respectively, which was 0.46 times lower than that in main basin populations, which ranged from 0.183 to 0.185 (figure 2b). These differences in the number of SNPs and median minor allele frequencies were not driven by differences in sample sizes (electronic supplementary material, table S1), sequencing quality (see 'Whole-genome sequencing' in electronic supplementary material), sequencing depth (electronic supplementary material, table S6), read depth of aligned reads (electronic supplementary material, table S7) or missing data (see 'Variant calling and filtering' in electronic supplementary material). Using the within-group dataset, we observed consistently positive Tajima's D values in the main basin populations, varying from 0.332 to 0.428 on average, which is suggestive of slight population declines in the main basin (figure 2c). In contrast, Green Bay and Muskegon Lake populations had a genome-wide average Tajima's D of −0.632 and −0.603, respectively (figure 2c), which is suggestive of a recent (over evolutionary time scales) demographic expansion.

(c) Effective population size, genetic diversity and demographic history

In comparison with the main basin sample sites, Green Bay and Muskegon Lake have significantly lower effective population sizes; the point estimates using whole-genome sequencing data for Green Bay and Muskegon Lake were 23 and 90, respectively, by NeEstimator v2 and 46 and 159, respectively, by GONE (figure 3a; electronic supplementary material, table S8). Effective population size estimates by NeEstimator v2 using our RAD-Seq dataset with larger numbers of samples [59], but smaller numbers of loci, suggest that southern and middle Green Bay (SGB, MEN) have smaller effective population sizes than northern Green Bay (BDN, LBD), with high levels of concurrence between both datasets, suggesting that main basin sample sites have considerably larger effective population sizes (figure 3b). Genetic diversity of main basin sample sites, with a mean observed heterozygosity ranging from 0.312 to 0.323, was considerably higher than that of Green Bay and Muskegon Lake populations, with a mean observed heterozygosity of 0.208 and 0.206, respectively (figure 3c; electronic supplementary material, figure S5). It should be noted that we calculated genetic diversity (i.e. observed heterozygosity) using only polymorphic sites found in each within-group dataset, meaning that the sets of SNPs for calculation differ among populations. Therefore, despite Green Bay

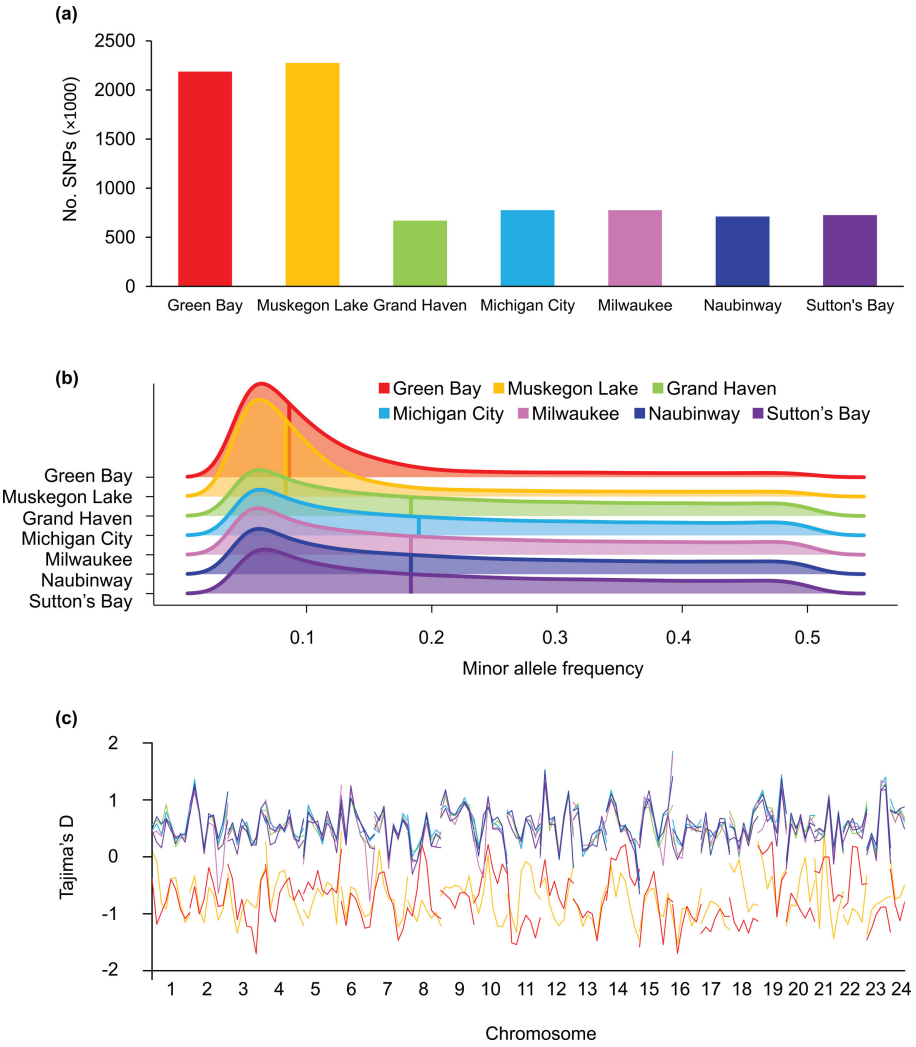

Figure 2. Numbers of SNPs, MAF and Tajima's D. (a) Green Bay and Muskegon Lake populations have 2.8–3.4 times more SNPs than the five main basin sample sites, despite all sample sites having equal sample sizes, sequencing quality, read depth and missing data. (b) Both Green Bay and Muskegon Lake also have a lower median MAF compared with the five main basin populations, here illustrated with a density plot where the median MAF is depicted with a solid vertical line. (c) The negative Tajima's D for both Green Bay and Muskegon Lake populations suggests recent population expansion. Collectively, these results suggest that the three populations (Green Bay, Muskegon Lake and the main basin sample sites) have different demographic histories.

and Muskegon Lake having a greater number of SNPs, the lower observed heterozygosity is predicted from theory given the much lower average minor allele frequencies. In contrast, estimates of nucleotide diversity (π), which is calculated using both polymorphic and invariant sites, were nearly three times higher in Green Bay and Muskegon Lake in comparison with the Lake Michigan main basin samples (electronic supplementary material, figure S6). Using GONE, we reconstructed contemporary demographic histories and found that Green Bay and Muskegon Lake experienced rapid growth of *Ne* over the last 10–30 generations (figure 3d,e; electronic supplementary material, figure S7). Stairway Plot 2 [38], based on site frequency spectra, suggests that increases in *Ne* in Green Bay and Muskegon Lake also occurred around 2000 years ago and even longer for main basin populations (electronic supplementary material, figures S8–S10).

(d) Candidate genes underlying genetic adaptation

Given the high levels of genetic divergence between Green Bay, Muskegon Lake and the main basin sites, we used a genome-wide scan to determine if any regions of the genome are suggestive of genetic adaptation to the different environments [74,75]. We retained a total of 13 outlier windows from nine chromosomes that were shared in at least 80% of comparisons between Green Bay and the main basin collection sites as well as Muskegon Lake and the main basin populations (see Methods; electronic supplementary material, table S9). The 13 outlier windows contained an average of 1427 SNPs (range: 363–3203) and an average of 225 high F_{ST} SNPs ($F_{ST} \geq 0.6$; range: 82–360 SNPs per window; figure 4a). SNPs within these 13 outlier windows mapped to 36 genes, among which three were located within a single outlier window on chromosome 4 (chromosome 4: 34762415–34852528). The three genes, *opn1sw*, *opn1lw* and *opn1lw2*, encode visual pigments, light-absorbing molecules mediating vision. Out of the three visual pigment-encoding genes, *opn1lw2* has two outlier SNPs with $F_{ST} \geq 0.5$ in comparisons between Green Bay and the main basin sample sites, and the exact same two outlier SNPs, one with $F_{ST} \geq 0.76$ and the other with $F_{ST} \geq 0.46$, were found in comparisons between Muskegon Lake and the main basin sample sites (figure 4a). The observation that three visual pigment-encoding genes were found on the same outlier window in all 10 pairwise comparisons suggests that this region has responded to recent positive and parallel selection. Across all outlier windows, seven

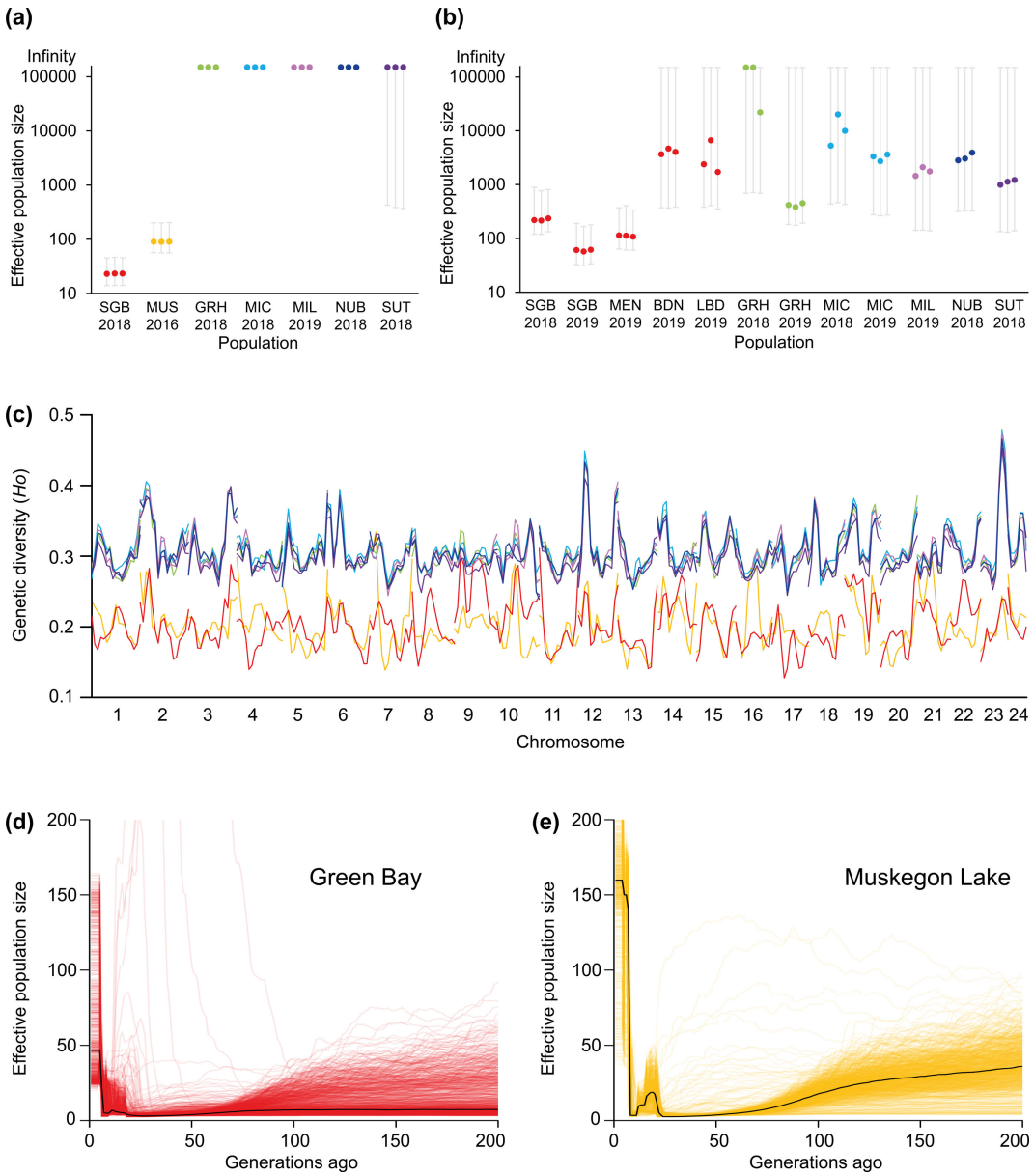

Figure 3. Demographic history of Lake Michigan yellow perch populations**.** Green Bay and Muskegon Lake sample sites have substantially lower effective population sizes, which average 23 and 90 estimated by NeEstimator v2, respectively, in comparison with main basin populations according to (a) estimates based on whole-genome sequencing data obtained in this study and (b) estimates based on RAD-Seq data obtained from Schraidt *et al.* [59] (note that *y*-axis is on a base 10 logarithmic scale). (c) Green Bay and Muskegon Lake have lower genetic diversity than main basin populations, measured here as observed heterozygosity averaged across overlapping windows of 5 MB using a step size of 2.5 MB across each chromosome. Notice that the main basin heterozygosity estimates are highly correlated, confirming higher gene flow among main basin sample sites. (d,e) The demographic history as reconstructed by GONE demonstrates a rapid increase in effective population size in (d) Green Bay and (e) Muskegon Lake since 10–30 generations ago. SGB, South Green Bay; MUS, Muskegon Lake; GRH, Grand Haven; MIC, Michigan City; MIL, Milwaukee; NUB, Naubinway; SUT, Sutton's Bay; MEN, Menominee; BDN, Big Bay de Noc; LBD, Little Bay de Noc.

major categories of biological processes were detected between Green Bay and the main basin sample sites and Muskegon Lake and the main basin sample sites: multicellular organismal process, response to stimulus, biological regulation, cellular process, developmental process, localization and metabolic process (figure 4b). Multiple visual perception-related biological processes, including detection of visible light, absorption of visible light, visual perception, cellular response to light stimulus and phototransduction, provide further support for parallel adaptive differences in vision in Green Bay and Muskegon Lake yellow perch populations (figure 4b). Using these candidate loci, we hypothesized that water clarity and eye size would differ between the two sites (Green Bay and Muskegon Lake) and the main basin. We found significant differences in water clarity between the main basin of Lake Michigan (low turbidity, high visual clarity; mean Secchi depth ± s.d.: 6.10 ± 3.23 m) and both Green Bay (high turbidity, low visual clarity; mean Secchi depth ± s.d.: 1.59 ± 1.01 m) and Muskegon Lake (high turbidity, low visual clarity; mean Secchi depth ± s.d.: 2.48 ± 1.23 m; ANOVA, $p < 0.001$; figure 4c). We also found that yellow perch eye diameter differed among locations (ANCOVA, $F_{2,91} = 12.03$; $p < 0.001$) and significantly increased with total length (ANCOVA, $F_{1,91} = 45.67$; $p < 0.001$), but the interaction was not significant (ANCOVA, $F_{2,91} = 0.11$; $p = 0.779$; electronic supplementary material, figure S11). Yellow perch eye diameter was significantly larger for Muskegon Lake yellow perch in comparison with fish collected from the main basin (Grand Haven; $p < 0.001$) and Green Bay (near Menominee; $p < 0.001$). Eye diameter for yellow perch from Green Bay was also significantly larger than fish from the main basin (Grand Haven; $p < 0.001$; figure 4d).

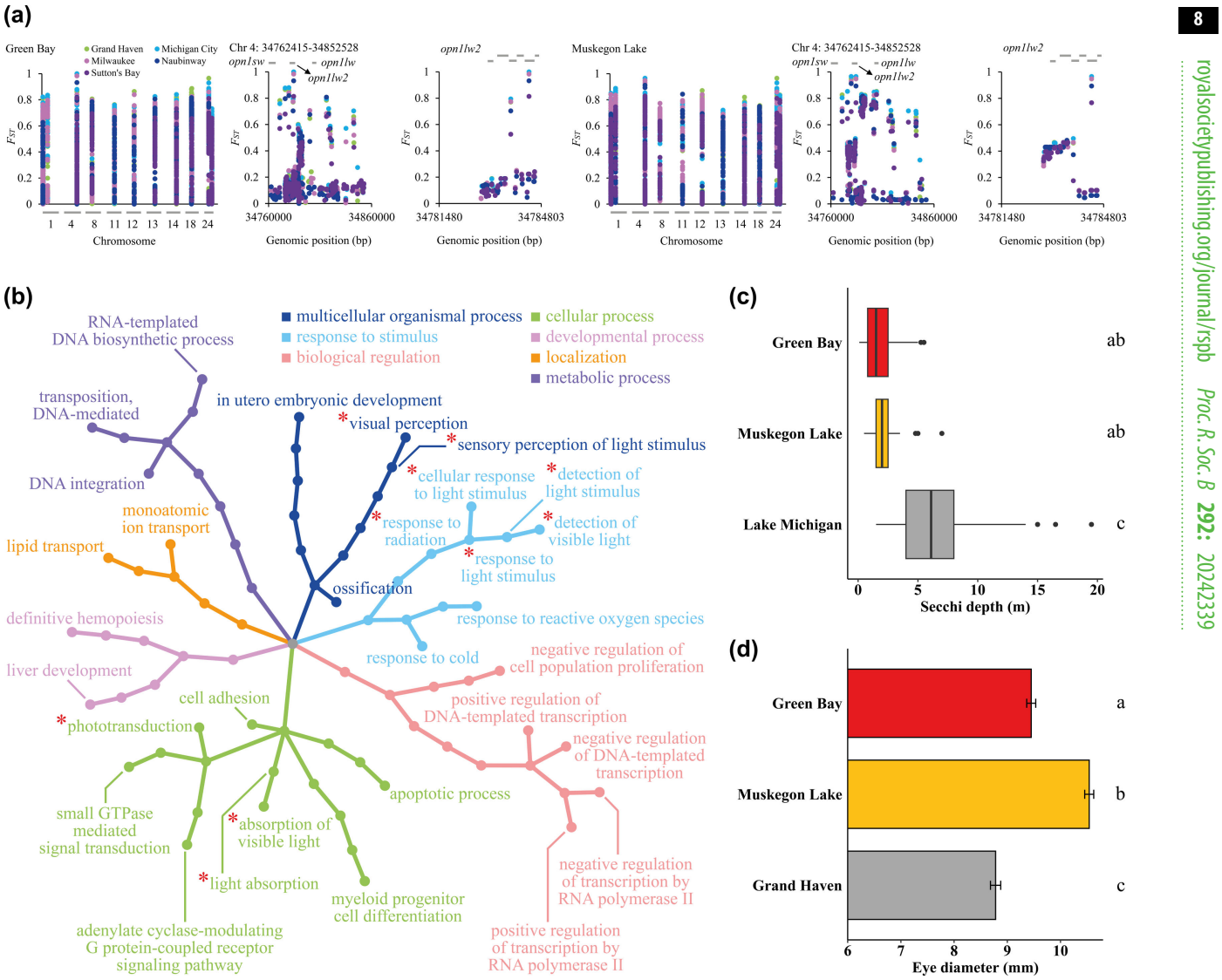

Figure 4. Candidate genes, biological processes and phenotypes involved in parallel genetic adaptation. (a) A total of 13 outlier windows were detected in comparisons between both Green Bay and the main basin sample sites and Muskegon Lake and the main basin sample sites. The outlier window on chromosome 4 contains visual pigment encoding genes, such as *opn1lw2*. The coding regions, as indicated by horizontal grey bars, of these outlier genes, contain SNPs that are highly differentiated. (b) Seven major categories of biological processes are potentially involved in adaptive differentiation in Green Bay and Muskegon Lake in comparison with main basin sample sites, among which visual perception-related processes (highlighted with red asterisks) further support parallel adaptive differences in vision in Green Bay and Muskegon Lake. (c,d) The genetic differentiation at *opn1lw2* is probably associated with water clarity, as evidenced by significant differences in (c) Secchi depths and (d) eye sizes between Green Bay or Muskegon Lake and main basin populations (i.e. Lake Michigan and Grand Haven). The Secchi depth in Green Bay and Muskegon Lake is significantly smaller than that in the southern basin of Lake Michigan ($p < 0.001$; c), and least-squares means of yellow perch eye diameter in Green Bay and Muskegon Lake is significantly larger than that in Grand Haven ($p < 0.001$; d). We accounted for differences in total length when comparing eye diameter of yellow perch collected from Green Bay, Muskegon Lake and Grand Haven (electronic supplementary material, figure S11; see Methods).

4. Discussion

Using whole-genome sequence data from 210 yellow perch, we observed a clear delineation of three populations and multiple genomic signatures associated with historic and contemporary increases in *Ne* in Green Bay and Muskegon Lake, where (i) effective population sizes rapidly increased over recent generations, (ii) Tajima's D estimates were negative across the entire genome and (iii) a factor of 2.8–3.4 more polymorphic SNPs were found in Green Bay and Muskegon Lake than the main basin collection sites, where the majority of these additional SNPs occur at low frequency. The additional SNPs probably come from an excess of mutations that accumulated during population expansions that occurred over historic time scales (i.e. thousands of years). This last observation is consistent with a large population expansion over historic time scales (e.g. over thousands of years), as indicated by Stairway Plot 2 (electronic supplementary material, figure S9), where there has not been sufficient time for purifying selection to remove deleterious alleles [76]. Given the substantial water quality and habitat improvements in Green Bay and Muskegon Lake driven by environmental remediation initiatives implemented in the early 1970s (electronic supplementary material, figures S12 and S13, and see 'Supplementary Discussion'), we should not have been surprised to see an increase in *Ne*. Nevertheless, empirical documentation of contemporary increases in *Ne* remains rare, and this result was unexpected.

To what time period do our demographic inferences apply? Linkage disequilibrium-based methods of *Ne*, such as those depicted in figure 3, provide recent and contemporary estimates of effective population size [77]. We used the program GONE to estimate recent changes to *Ne* over time, and the results suggest that there have been contemporary increases in *Ne* in both Green Bay and Muskegon Lake. However, the exact date for when these increases in *Ne* began remains imprecise. A recent study with pink salmon (*Oncorhynchus gorbuscha*) and a known date of an extreme founder event illustrated that GONE may have missed the timing of the event by as much as six generations [45]. Nevertheless, the known abundance patterns in salmon and other species are well characterized with this approach [45,78], and our results clearly illustrate that effective population sizes have been increasing in both Muskegon Lake and Green Bay in the last 10–30 generations, with larger increases occurring in Muskegon Lake. Our inferences based on Tajima's D, the number of segregating loci and site frequency spectra suggest much larger increases in *Ne* that occurred over historic time scales (centuries, not decades; figure 2; electronic supplementary material, figure S9). This result contrasts with those from main basin sites, which all show increases in *Ne* starting more than 3000 years ago and no signs of a historic bottleneck (electronic supplementary material, figure S9). We speculate that both the Green Bay and the Muskegon Lake yellow perch populations were founded much more recently than the main basin populations or that both went through a large, unknown, bottleneck event approximately 2000 years ago.

We also found evidence of parallel evolution in response to reduced water clarity in both Green Bay and Muskegon Lake yellow perch. In comparison with the main basin of Lake Michigan, both Green Bay and Muskegon Lake exhibit high turbidity and low visual clarity, as indicated by shallower Secchi depths. Such differences in water clarity and ambient light conditions could have driven the significant allele frequency shifts in the coding sequences of photopigment encoding genes and triggered a set of biological processes contributing to light detection and absorption, visual perception and phototransduction, a result that implies parallel genetic adaptation in Green Bay and Muskegon Lake yellow perch in response to local environmental conditions. Differences in vision and eye size, while accounting for allometry, among those sample sites was not an *a priori* hypothesis, although it could have been had we thought about the differences in water clarity among the sample sites. Nevertheless, using whole-genome sequencing to guide phenotypic and evolutionary predictions does suggest a promising avenue for the future of genomic tools [79]. Future studies should investigate linkages between candidate genes and their patterns of expression in addition to the structural and functional analyses of the proteins they encode. Additional phenotyping efforts and comparisons among multiple drowned river mouth lakes would also be useful. The observation that both genetic and phenotypic changes to reduced water clarity occurred in genetically divergent populations suggests that a match between vision-related traits and environmental characteristics is an important component of fitness in yellow perch populations [80]; environmental changes that cause a mismatch may result in declines in abundance.

Several caveats are worth discussing. First, linkage disequilibrium-based estimates of effective population sizes may not accurately estimate effective population sizes prior to bottleneck events [45,81]. As such, the effective population sizes of yellow perch in Green Bay and Muskegon Lake prior to around 30 generations ago were likely much larger than present day estimates but cannot be estimated using those approaches. Also, contemporary effective population size may be underestimated by linkage disequilibrium-based methods due to overlapping generations of yellow perch and the mixed-age adult samples under study [82]. Demographic estimates by GONE may also be unreliable for the most recent several generations. However, despite the wide range of estimates for recent generations, the general trend still indicates an increase in effective population size that appears to have persisted in Green Bay and Muskegon Lake for the past 10–30 generations. Perhaps most importantly, some of our demographic results could also be explained by gene flow from divergent populations [83,84]; however, we do not think this to be the case for three reasons. First, PCA, phylogenetic and admixture analyses support a clear separation of Green Bay, Muskegon Lake and Lake Michigan main basin populations and suggest the absence of admixed individuals in both Green Bay and Muskegon Lake. Even if all the main basin loci were introduced into Muskegon Lake samples via gene flow (which is unlikely), this mechanism could only account for 24% of the 2 275 935 SNPs found in Muskegon Lake (i.e. 535 860 SNPs shared by Muskegon Lake and Grand Haven divided by 2 275 935 Muskegon Lake SNPs; figure 2a). Second, over shorter time scales, estimates of contemporary admixture within Green Bay and Muskegon Lake samples suggest a single, randomly mating population (as do the small number of loci out of Hardy–Weinberg equilibrium; electronic supplementary material, figure S4 and table S4). Third, over longer time scales, the high rates of introgression required to explain these results would also be likely to swamp any signal of parallel genetic adaptation (figure 4) and would also leave more time for purifying selection to remove deleterious loci. Note that other factors that can increase *Ne* through time (e.g. decreased variance in reproductive success) are unlikely to be accompanied by the other genomic patterns we report (i.e. negative Tajima's D and larger numbers of SNPs). Thus, we think the most parsimonious explanation for our results may be large, simultaneous increases in population sizes in Green Bay and Muskegon Lake yellow perch. Interestingly, fisheries independent data suggest that yellow perch have been declining recently in the main basin [85] and the positive, albeit small, genome-wide estimates of Tajima's D suggest that these populations should be closely monitored in the coming decades. While we are aware that correlation does not equal causation, we speculate that the recent increases in *Ne* in Green Bay and Muskegon Lake could be a result of substantial water quality and habitat improvements driven by a set of environmental remediation initiatives implemented in the early 1970s, as evidenced by the strong relationship between decreases in nutrient loads and concomitant increases in *Ne* [86,87] (electronic supplementary material, figures S12 and S13, and 'Supplementary Discussion').

In conclusion, we found that two populations of an economically important fish species have had large, simultaneous increases in effective population sizes both historically and over contemporary time scales. The two populations are separated by over 350 km and are genetically divergent from one another, such that they represent two distinct evolutionary lineages (figure 1b). We also found evidence of parallel adaptation to reduced visual clarity in each of the habitats, which may also have been further magnified by centuries-long anthropogenic disturbances and pollution, though we suspect that this genetic adaptation occurred over a longer time scale. The contemporary estimates of *Ne* are orders of magnitude lower than historic

estimates, suggesting that recent anthropogenic disturbances may have reduced genetic diversity in Green Bay and Muskegon Lake. While the recent increases in *Ne* suggest that the largest threats to these populations may have passed, the overall point estimates remain low. We suggest that conservation and management efforts targeting yellow perch in these regions should strive to maintain genetic diversity and obtain accurate estimates of population abundances through time. Taken together, these results suggest that the synchrony of key ecological and evolutionary processes can drive parallel demographic and evolutionary trajectories across independent populations. How frequently such synchrony occurs remains an open question in the fields of ecology and evolution.

Ethics. Field collections of yellow perch were authorized by issuance of state-specific collector's permits and approved by Purdue University's Animal Care and Use Committee protocol 1112000400.

Data accessibility. All raw data generated for this project are stored in the NCBI Short Read Archive (SRA) PRJNA1037191 Great Lakes yellow perch whole genome sequencing. Additionally, unfiltered and filtered VCF files (as described in electronic supplementary material, Table S2) are also available [88]. The scripts used in this project are hosted in a public repository [89].

Supplementary material is available online [90].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. X.Y.: conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing—original draft, writing—review and editing; C.E.S.: data curation; M.M.S.: formal analysis; P.T.E.: formal analysis; T.J.H.: formal analysis; C.R.R.: data curation, formal analysis, funding acquisition; T.O.H.: conceptualization, funding acquisition; M.R.C.: conceptualization, formal analysis, funding acquisition, investigation, methodology, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was funded by the Great Lakes Fishery Commission (Project IDs: 2018_CHR_44072, 2021_CHR_440970) with additional support from the Great Lakes Fishery Trust (#1456 and #1952) to sample Muskegon Lake.

Acknowledgements. We thank the Michigan Department of Natural Resources, the Indiana Department of Natural Resources, the Wisconsin Department of Natural Resources, the Illinois Natural History Survey and the Grand Traverse Band of Ottawa and Chippewa Indians for assistance with sample collection. In particular, we thank Charles Roswell, Troy Zorn, Tammie Paoli, Dave Clapp, Pat O'Neil, Dave Fielder, David Schindelholz, Erik Olsen, Jayson Beugly, Josey Cline, Taylor Senegal, Greg Chorak and Scott Koenigbauer. We thank Kurt Thompson for assistance with the map of Muskegon Lake. Secchi depth data were collected and provided by Sarah Bartlett and the NEW Water monitoring programme in Green Bay and Sarah Stein and Les Warren of Purdue University.

References

1. Wright S. 1931 Evolution in mendelian populations. *Genetics* **16**, 97–159. (doi:10.1093/genetics/16.2.97)
2. Hedgecock D, Pudovkin AI. 2011 Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bull. Mar. Sci.* **87**, 971–1002. (doi:10.5343/bms.2010.1051)
3. Hedgecock D. 1994 Does variance in reproductive success limit effective population sizes of marine organisms? In *Genetics and evolution of aquatic organisms* (ed. A Beaumont), pp. 122–134. New York, NY: Springer.
4. Hauser L, Adcock GJ, Smith PJ, Ramiréz JHB, Carvalho GR. 2002 Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proc. Natl Acad. Sci. USA* **99**, 11742–11747. (doi:10.1073/pnas.172242899)
5. Hoarau G, Boon E, Jongma DN, Ferber S, Palsson J, Van der Veer HW, Rijnsdorp AD, Stam WT, Olsen JL. 2005 Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa L.*). *Proc. R. Soc. B* **272**, 497–503. (doi:10.1098/rspb.2004.2963)
6. Miller LM, Kapuscinski AR. 1997 Historical analysis of genetic variation reveals low effective population size in a northern pike (*Esox lucius*) population. *Genetics* **147**, 1249–1258. (doi:10.1093/genetics/147.3.1249)
7. Turner TF, Wares JP, Gold JR. 2002 Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* **162**, 1329–1339. (doi:10.1093/genetics/162.3.1329)
8. Hutchinson WF, van Oosterhout C, Rogers SI, Carvalho GR. 2003 Temporal analysis of archived samples indicates marked genetic changes in declining north sea cod (*Gadus morhua*). *Proc. R. Soc. B* **270**, 2125–2132. (doi:10.1098/rspb.2003.2493)
9. Waples RS. 2016 Tiny estimates of the Ne/N ratio in marine fishes: are they real? *J. Fish Biol.* **89**, 2479–2504. (doi:10.1111/jfb.13143)
10. Waples RS, Grewe PM, Bravington MW, Hillary R, Feutry P. 2018 Robust estimates of a high Ne/N ratio in a top marine predator, southern bluefin tuna. *Sci. Adv.* **4**, eaar7759. (doi:10.1126/sciadv.aar7759)
11. Palstra FP, Fraser DJ. 2012 Effective/census population size ratio estimation: a compendium and appraisal. *Ecol. Evol.* **2**, 2357–2365. (doi:10.1002/ece3.329)
12. Ruzzante DE, McCracken GR, Parmelee S, Hill K, Corrigan A, MacMillan J, Walde SJ. 2016 Effective number of breeders, effective population size and their relationship with census size in an iteroparous species, *Salvelinus fontinalis*. *Proc. R. Soc. B* **283**, 20152601. (doi:10.1098/rspb.2015.2601)
13. Clarke SH *et al*. 2024 Global assessment of effective population sizes: consistent taxonomic differences in meeting the 50/500 rule. *Mol. Ecol.* **33**, e17353. (doi:10.1111/mec.17353)
14. Hahn M. 2018 *Molecular population genetics*. Oxford, UK: Oxford University Press.
15. Waples RS. 2022 What is *Ne*, anyway? *J. Hered.* (ed. W Murphy), **113**, 371–379. (doi:10.1093/jhered/esac023)
16. Capblancq T, Butnor JR, Deyoung S, Thibault E, Munson H, Nelson DM, Fitzpatrick MC, Keller SR. 2020 whole-exome sequencing reveals a long-term decline in effective population size of red spruce (*Picea rubens*) . *Evol. Appl.* **13**, 2190–2205. (doi:10.1111/eva.12985)
17. Sharma R *et al*. 2012 Effective population size dynamics and the demographic collapse of Bornean orang-utans. *PLoS One* **7**, e49429. (doi:10.1371/journal.pone.0049429)
18. Okello JBA, Wittemyer G, Rasmussen HB, Arctander P, Nyakaana S, Douglas-Hamilton I, Siegismund HR. 2008 Effective population size dynamics reveal impacts of historic climatic events and recent anthropogenic pressure in African elephants. *Mol. Ecol.* **17**, 3788–3799. (doi:10.1111/j.1365-294X.2008.03871.x)
19. Ólafsdóttir GÁ, Westfall KM, Edvardsson R, Pálsson S. 2014 Historical DNA reveals the demographic history of Atlantic cod (*Gadus morhua*) in medieval and early modern Iceland. *Proc. Biol. Sci.* **281**, 20132976. (doi:10.1098/rspb.2013.2976)

20. Lonsinger RC, Adams JR, Waits LP. 2018 Evaluating effective population size and genetic diversity of a declining kit fox population using contemporary and historical specimens. *Ecol. Evol.* **8**, 12011–12021. (doi:10.1002/ece3.4660)
21. Abadía-Cardoso A, Freimer NB, Deiner K, Garza JC. 2017 Molecular population genetics of the northern elephant seal *Mirounga angustirostris*. *J. Hered.* **108**, 618–627. (doi:10.1093/jhered/esx053)
22. Pita A, Pérez M, Velasco F, Presa P. 2017 Trends of the genetic effective population size in the southern stock of the European hake. *Fish. Res.* **191**, 108–119. (doi:10.1016/j.fishres.2017.02.022)
23. Ovenden JR, Leigh GM, Blower DC, Jones AT, Moore A, Bustamante C, Buckworth RC, Bennett MB, Dudgeon CL. 2016 Can estimates of genetic effective population size contribute to fisheries stock assessments? *J. Fish Biol.* **89**, 2505–2518. (doi:10.1111/jfb.13129)
24. Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, Ruegg K, Palstra F. 2011 Understanding and estimating effective population size for practical application in marine species management. *Conserv. Biol.* **25**, 438–449. (doi:10.1111/j.1523-1739.2010.01637.x)
25. Shrimpton JM, Heath DD. 2003 Census vs. effective population size in chinook salmon: large- and small-scale environmental perturbation effects. *Mol. Ecol.* **12**, 2571–2583. (doi:10.1046/j.1365-294x.2003.01932.x)
26. Horreo JL, Machado-Schiaffino G, Griffiths AM, Bright D, Stevens JR, Garcia-Vazquez E. 2011 Atlantic salmon at risk: apparent rapid declines in effective population size in southern European populations. *Trans. Am. Fish. Soc.* **140**, 605–610. (doi:10.1080/00028487.2011.585574)
27. Kuparinen A, Hutchings JA, Waples RS. 2016 Harvest-induced evolution and effective population size. *Evol. Appl.* **9**, 658–672. (doi:10.1111/eva.12373)
28. Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR. 2014 NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (*Ne*) from genetic data. *Mol. Ecol. Resour.* **14**, 209–214. (doi:10.1111/1755-0998.12157)
29. Nomura T. 2008 Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evol. Appl.* **1**, 462–474. (doi:10.1111/j.1752-4571.2008.00015.x)
30. Cowen R, Gawarkiewicz G, Pineda J, Thorrold S, Werner F. 2007 Population connectivity in marine systems: an overview. *Oceanography* **20**, 14–21. (doi:10.5670/oceanog.2007.26)
31. Cowen RK, Sponaugle S. 2009 Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.* **1**, 443–466. (doi:10.1146/annurev.marine.010908.163757)
32. Waples RS. 1998 Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* **89**, 438–450. (doi:10.1093/jhered/89.5.438)
33. Christie MR, Tissot BN, Albins MA, Beets JP, Jia Y, Ortiz DM, Thompson SE, Hixon MA. 2010 Larval connectivity in an effective network of marine protected areas. *PLoS One* **5**, e15715. (doi:10.1371/journal.pone.0015715)
34. Waples RS. 2024 Practical application of the linkage disequilibrium method for estimating contemporary effective population size: a review. *Mol. Ecol. Resour.* **24**, e13879. (doi:10.1111/1755-0998.13879)
35. Palstra FP, Ruzzante DE. 2008 Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol. Ecol.* **17**, 3428–3447. (doi:10.1111/j.1365-294x.2008.03842.x)
36. Jorde PE, Ryman N. 1995 Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics* **139**, 1077–1090. (doi:10.1093/genetics/139.2.1077)
37. Liu X, Fu YX. 2015 Exploring population size changes using SNP frequency spectra. *Nat. Genet.* **47**, 555–559. (doi:10.1038/ng.3254)
38. Liu X, Fu YX. 2020 Stairway plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biol.* **21**, 280. (doi:10.1186/s13059-020-02196-9)
39. Santiago E, Novo I, Pardiñas AF, Saura M, Wang J, Caballero A. 2020 Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Mol. Biol. Evol.* **37**, 3642–3653. (doi:10.1093/molbev/msaa169)
40. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009 Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695. (doi:10.1371/journal.pgen.1000695)
41. Li H, Durbin R. 2011 Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496. (doi:10.1038/nature10231)
42. Laurent V, Planes S. 2007 Effective population size estimation on *Sardina pilchardus* in the bay of biscay using a temporal genetic approach. *Biol. J. Linnean Soc.* **90**, 591–602. (doi:10.1111/j.1095-8312.2007.00747.x)
43. Westbury MV *et al*. 2018 Extended and continuous decline in effective population size results in low genomic diversity in the world's rarest hyena species, the brown hyena. *Mol. Biol. Evol.* (ed. B Shapiro), **35**, 1225–1237. (doi:10.1093/molbev/msy037)
44. Diaz M, Wethey D, Bulak J, Ely B. 2000 Effect of harvest and effective population size on genetic diversity in a striped bass population. *Trans. Am. Fish. Soc.* **129**, 1367–1372. (doi:10.1577/1548-8659(2000)129<1367:EOHAEP>2.0.CO;2)
45. Sparks MM, Schraidt CE, Yin X, Seeb LW, Christie MR. 2024 Rapid genetic adaptation to a novel ecosystem despite a large founder event. *Mol. Ecol.* **33**, e17121. (doi:10.1111/mec.17121)
46. Hoey JA, Able KW, Pinsky ML. 2022 Genetic decline and recovery of a demographically rebuilt fishery species. *Mol. Ecol.* **31**, 5684–5698. (doi:10.1111/mec.16697)
47. Wang J, Santiago E, Caballero A. 2016 Prediction and estimation of effective population size. *Heredity* **117**, 193–206. (doi:10.1038/hdy.2016.43)
48. Hedrick P. 2005 Large variance in reproductive success and the Ne/N ratio. *Evolution* **59**, 1596–1599. (doi:10.1554/05-009)
49. Allendorf FW. 1993 Delay of adaptation to captive breeding by equalizing family size. *Conserv. Biol.* **7**, 416–419. (doi:10.1046/j.1523-1739.1993.07020416.x)
50. Borlase SC, Loebel DA, Frankham R, Nurthen RK, Briscoe DA, Daggard GE. 1993 modeling problems in conservation genetics using captive *Drosophila* populations: consequences of equalization of family sizes . *Conserv. Biol.* **7**, 122–131. (doi:10.1046/j.1523-1739.1993.07010122.x)
51. Waples RS, Hindar K, Karlsson S, Hard JJ. 2016 Evaluating the ryman-laikre effect for marine stock enhancement and aquaculture. *Curr. Zool.* **62**, 617–627. (doi:10.1093/cz/zow060)
52. Ryman N, Laikre L. 1991 Effects of supportive breeding on the genetically effective population size. *Conserv. Biol.* **5**, 325–329. (doi:10.1111/j.1523-1739.1991.tb00144.x)
53. Great Lakes Commission. Lake Michigan. See https://www.glc.org/lakes/lake-michigan/ (accessed 22 September 2024).
54. NOAA. National ecosystem status. See https://ecowatch.noaa.gov/index.php/regions/great-lakes/lake-michigan (accessed 22 September 2022).
55. Herdendorf CE. 1990 Great Lakes estuaries. *Estuaries* **13**, 493. (doi:10.2307/1351795)
56. Mader MM, Ruetz CR, Woznicki SA, Steinman AD. 2023 Land cover and water quality of drowned river mouths: evidence of an environmental gradient along the eastern lake Michigan shoreline. *J. Great Lakes Res.* **49**, 102237. (doi:10.1016/j.jglr.2023.09.008)
57. Senegal TJ, Ruetz CR III, Chorak GM, Janetski DJ, Clapp DF, Bowen GJ, Höök TO. 2020 Differential habitat use patterns of yellow perch *Perca flavescens* in eastern lake Michigan and connected drowned river mouth lakes. *J. Great Lakes Res.* **46**, 1412–1422. (doi:10.1016/j.jglr.2020.06.021)
58. Chorak GM, Ruetz CR III, Thum RA, Partridge CG, Janetski DJ, Höök TO, Clapp DF. 2019 Yellow perch genetic structure and habitat use among connected habitats in eastern Lake Michigan. *Ecol. Evol.* **9**, 8922–8932. (doi:10.1002/ece3.5219)
59. Schraidt CE, Ackiss AS, Larson WA, Rowe MD, Höök TO, Christie MR. 2023 Dispersive currents explain patterns of population connectivity in an ecologically and economically important fish. *Evol. Appl.* **16**, 1284–1301. (doi:10.1111/eva.13567)

60. Miller LM. 2003 Microsatellite DNA loci reveal genetic structure of yellow perch in Lake Michigan. *Trans. Am. Fish. Soc.* **132**, 503–513. (doi:10.1577/1548-8659(2003)132<0503:MDLRGS>2.0.CO;2)
61. Evans MS. 1986 Recent major declines in zooplankton populations in the inshore region of lake Michigan: probable causes and implications. *Can. J. Fish. Aquat. Sci.* **43**, 154–159. (doi:10.1139/f86-017)
62. Feron R *et al*. 2020 Characterization of a Y-specific duplication/insertion of the anti-mullerian hormone type II receptor gene based on a chromosome-scale genome assembly of yellow perch, *Perca flavescens*. *Mol. Ecol. Resour.* **20**, 531–543. (doi:10.1111/1755-0998.13133)
63. McKenna A *et al*. 2010 The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. *Genome Res.* **20**, 1297–1303. (doi:10.1101/gr.107524.110)
64. Hemstrom W, Grummer JA, Luikart G, Christie MR. 2024 Next-generation data filtering in the genomics era. *Nat. Rev. Genet.* **25**, 750–767. (doi:10.1038/s41576-024-00738-6)
65. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
66. Tamura K, Stecher G, Kumar S. 2021 MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027. (doi:10.1093/molbev/msab120)
67. R Core Team. 2020 R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. See https://www.R-project.org.
68. Peakall R, Smouse PE. 2012 GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537–2539. (doi:10.1093/bioinformatics/bts460)
69. Weir BS, Cockerham CC. 1984 Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370. (doi:10.1111/j.1558-5646.1984.tb05657.x)
70. Alexander DH, Novembre J, Lange K. 2009 Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664. (doi:10.1101/gr.094052.109)
71. Wickham H. 2009 *Ggplot2: elegant graphics for data analysis*. New York, NY: Springer.
72. Korunes KL, Samuk K. 2021 Pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol. Ecol. Resour.* **21**, 1359–1368. (doi:10.1111/1755-0998.13326)
73. Foster ZSL, Sharpton TJ, Grünwald NJ. 2017 Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput. Biol.* **13**, e1005404. (doi:10.1371/journal.pcbi.1005404)
74. Willoughby JR, Harder AM, Tennessen JA, Scribner KT, Christie MR. 2018 Rapid genetic adaptation to a novel environment despite a genome-wide reduction in genetic diversity. *Mol. Ecol.* **27**, 4041–4051. (doi:10.1111/mec.14726)
75. Bernatchez L, Ferchaud AL, Berger CS, Venney CJ, Xuereb A. 2024 Genomics for monitoring and understanding species responses to global climate change. *Nat. Rev. Genet.* **25**, 165–183. (doi:10.1038/s41576-023-00657-y)
76. Marth GT, Czabarka E, Murvai J, Sherry ST. 2004 The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. *Genetics* **166**, 351–372. (doi:10.1534/genetics.166.1.351)
77. Waples RS, Do C. 2010 Linkage disequilibrium estimates of contemporary n e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol. Appl.* **3**, 244–262. (doi:10.1111/j.1752-4571.2009.00104.x)
78. Kardos M *et al*. 2023 Inbreeding depression explains killer whale population dynamics. *Nat. Ecol. Evol.* **7**, 675–686. (doi:10.1038/s41559-023-01995-0)
79. Garner BA *et al*. 2016 Genomics in conservation: case studies and bridging the gap between data and application. *Trends Ecol. Evol.* **31**, 81–83. (doi:10.1016/j.tree.2015.10.009)
80. Andersson ML, Scharnweber K, Eklöv P. 2024 Environmental and ecological drivers of eye size variation in a freshwater predator: a trade-off between foraging and predation risk. *Funct. Ecol.* **38**, 2470–2477. (doi:10.1111/1365-2435.14655)
81. Reid BN, Pinsky ML. 2022 Simulation-based evaluation of methods, data types, and temporal sampling schemes for detecting recent population declines. *Integr. Comp. Biol.* **62**, 1849–1863. (doi:10.1093/icb/icac144)
82. Waples RS, Antao T, Luikart G. 2014 Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics* **197**, 769–780. (doi:10.1534/genetics.114.164822)
83. Novo I, Ordás P, Moraga N, Santiago E, Quesada H, Caballero A. 2023 Impact of population structure in the estimation of recent historical effective population size by the software gone. *Genet. Sel. Evol.* **55**, 86. (doi:10.1186/s12711-023-00859-2)
84. Paz-Vinas I, Quéméré E, Chikhi L, Loot G, Blanchet S. 2013 The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Mol. Ecol.* **22**, 3279–3291. (doi:10.1111/mec.12321)
85. Clapp D, Makauskas D. 2021 Status of yellow perch in Lake Michigan—2019–2020. Report to the Lake Michigan Committee. See https://www.glfc.org/pubs/lake_committees/michigan/yellow_perch/Status%20of%20YEP%20in%20LM%202019-2020%20(2021).pdf.
86. Harris HJ, Wenger RB, Sager PE, Val Klump J. 2018 The green bay saga: environmental change, scientific investigation, and watershed management. *J. Great Lakes Res.* **44**, 829–836. (doi:10.1016/j.jglr.2018.08.001)
87. Liu B, McLean CE, Long DT, Steinman AD, Stevenson RJ. 2018 Eutrophication and recovery of a lake inferred from sedimentary diatoms originating from different habitats. *Sci. Total Environ.* **628–629**, 1352–1361. (doi:10.1016/j.scitotenv.2018.02.174)
88. Yin X. 2023 Data for: Great Lakes yellow perch vcf files. Figshare. (doi:10.6084/m9.figshare.24517471)
89. Yin X. 2024 Online resources for: Great Lakes yellow perch scripts. Figshare. (doi:10.6084/m9.figshare.27830115)
90. Yin X, Schraidt CE, Sparks MM, Euclide PT, Hoyt TJ, Ruetz CR, Höök TO, Christie MR. 2024 Supplementary material from: Parallel genetic adaptation amidst a background of changing effective population sizes in divergent yellow perch (*Perca flavescens*) populations. Figshare. (doi:10.6084/m9.figshare.c.7597090)