Received 19 October 2016

Accepted 13 February 2017

Population genetic structure and comparative diversity of smallmouth bass *Micropterus dolomieu*: congruent patterns from two genomes

C. A. STEPIEN[1,2*], S. I. KARSIOTIS[1], T. J. SULLIVAN[1,3] and K E. KLYMUS[1,4]

[1]*Great Lakes Genetics/Genomics Laboratory, Department of Environmental Sciences, University of Toledo, Toledo, OH 43606, U.S.A. and* [2]*NOAA Pacific Marine Environmental Laboratory (PMEL), 7600 Sand Point Way NE, Seattle, WA 98115, U.S.A.*

*Author to whom correspondence should be addressed: Tel.: +1 206-526-6038; email: carol.stepien@noaa.gov; stepien.carol@gmail.com

[3]Present address: Department of Biology, University of Louisiana at Lafayette, Billeaud Hall, 410 East St Mary Blvd Lafayette LA 70504, U.S.A.

[4]Present address: U.S. Geological Survey, Columbia Environmental Research Center, Columbia, MO 65201, U.S.A.

24 Running title: *MICROPTERUS DOLOMIEU* POPULATION GENETICS

25

26

27

Genetic diversity and divergence patterns of smallmouth bass *Micropterus dolomieu* spawning groups are analysed across its northern native range with mtDNA cytochrome *b* gene sequences and eight unlinked nuclear DNA microsatellite loci. Results reveal high levels of genetic variability and significant differences in allelic representation among populations (mtDNA: mean ± S.E., $H_D = 0{\cdot}50 \pm 0{\cdot}06$, mean ± S.E., $\Phi_{ST} = 0{\cdot}41 \pm 0{\cdot}02$ and microsatellites: mean ± S.E. $H_O = 0{\cdot}46 \pm 0{\cdot}03$, mean ± S.E. $F_{ST} = 0{\cdot}25 \pm 0{\cdot}01$). The distributions of 28 variant mtDNA haplotypes, which differ by an average of 3·94 nucleotides (range = 1–8), denote divergent representation among geographic areas. Microsatellite data support nine primary population groups, whose high self-assignment probabilities likewise display marked divergence. Genetic patterns demonstrate: (1) high genetic diversity in both genomes, (2) significant divergence among populations, probably resulting from natal site homing and low lifetime migration, (3) support for three post-glacial refugia that variously contributed to the current northern populations, which remain evident today despite waterway connectivity and (4) a weak yet significant genetic isolation by geographic distance pattern, indicating that other processes affect the differences among populations, such as territoriality and site fidelity.

Key words: cytochrome *b*; Great Lakes; *Micropterus dolomieu*; microsatellites; phylogeography; smallmouth bass.

INTRODUCTION

Wide-ranging species often harbour a large number of genetically variable populations, whose broad-scale genetic patterns reflect their demographic and geographic histories and interactions (Slatkin, 1987; Avise, 2000; Manel *et al.,* 2003). Fine-scale population structure of such species is influenced by gene flow and differentiation patterns that are regulated by spatial patterns of habitat availability, connectivity among locations, dispersal ability and reproductive behaviour. Population divergence decreases with increased gene flow, facilitated by higher degrees of habitat connectivity and dispersal (Boizard *et al.,* 2009; Fitzpatrick *et al.,* 2014; Kutschera *et al*. 2016). In contrast, habitat isolation and behavioural site fidelity will reduce gene flow and increase population differences (Lindsay *et al.,* 2008; Nichols *et al.,* 2012; Baker *et al.,* 2013). Furthermore, compared with large populations, small or isolated ones often contain lower genetic diversity resulting from genetic drift, population bottlenecks and localized selection (Nei *et al.,* 1975; Hedrick & Kalinowski, 2000; Vandewoestjine *et al.,* 2008; Bouzat, 2010; Puckett *et al.,* 2014).

Broad-scale population genetic patterns generally are shaped by historical climatic and geographic influences on demography, including barriers to dispersal and connectivity. In temperate North America, the primary historic determinant was the Pleistocene glaciations from *c.* 2 580 000 to *c.* 12 000 years ago (ya), which markedly limited available habitat (Fig. 1; Hewitt, 2004; Gibbard *et al.,* 2010). Populations of freshwater fishes were concentrated in southerly glacial refugia, where they experienced isolation and bottlenecks, followed by subsequent founder effects with recolonization of northern habitats after glacial recession (Bernatchez, 1997; Hewitt, 2000; Gugerli *et al.,*

2001). Interestingly, recent studies have indicated a seat of high population genetic variability in several fish species in the Laurentian Great Lakes, which offered huge and abundant new habitat formed by the glaciers. This high diversity is exemplified in today's Great Lakes' populations of smallmouth bass *Micropterus dolomieu* Lacépède 1802 (Stepien *et al.,* 2007), yellow perch *Perca flavescens* (Mitchill 1814) (Sepulveda-Villet & Stepien, 2012; Sullivan & Stepien, 2014, 2015; Stepien *et al.* 2015*a*,*b*), walleye *Sander vitreus* (Mitchill 1818) (Stepien *et al*., 2009, 2015*b*; Haponski & Stepien, 2014), lake whitefish *Coregonus clupeaformis* (Mitchill 1818) (Bernatchez & Dodson, 1991) and brown bullhead *Ameiurus nebulosus* (LeSueur 1819) (Murdoch & Hebert, 1997). The pattern of substantial genetic diversity in recolonized areas is believed to reflect admixture of populations from several different glacial refugia (Bernatchez & Dodson, 1991; Stepien *et al.,* 2007, 2015*a*,*b*; Sepulveda-Villet & Stepien, 2012).

The genus *Micropterus* Lacépède 1802 (family Centrarchidae) contains 13 North American species of black basses (Baker *et al*., 2013), which are top carnivores and popular sport fishes (Bagley *et al.*, 2011). The wide-ranging smallmouth bass *M. dolomieu* possesses the most northerly range, which encompasses the entirety of the Great Lakes, along with adjacent rivers and streams (Bagley *et al.,* 2011). This broad range provides an interesting case study for investigating biogeographic patterns in formerly glaciated regions and contemporary influences on these patterns. The present study analyses the relative diversity and divergence from both the mitochondrial (mt) and nuclear genomes, with the aim of elucidating the interplay between historic and contemporary factors on population genetic structure. Since mtDNA is maternally inherited and largely free of recombination, its patterns of genetic variation are more conserved, prone to bottlenecks

93 and often retain the evidence of historic differentiation at a broader geographic scale. In
94 contrast, nuclear DNA microsatellites are bi-parentally inherited, have a higher mutation
95 rate and may better reflect contemporary population genetic relationships at finer temporal
96 and spatial scales.

97 An earlier study by Stepien *et al.* (2007) evaluated eight microsatellite loci of *M.*
98 *dolomieu* populations across North America, discerning pronounced divergences and
99 primary genetic barriers separating the Hudson River, Ohio River and upper Mississippi
100 River population groups. Stepien *et al.* (2007) noted that that its historical glacial refugia
101 signatures were similar to patterns delineated in other species, including *S. vitreus*
102 (Stepien & Faber 1998, Stepien *et al.* 2009), *A. nebulosus* (Murdoch & Hebert, 1997) and
103 *P. flavescens* (Sepulveda-Villet & Stepien 2012). Here, new analyses of those results are
104 compared and contrasted with a new mtDNA cytochrome *b* (*cytb*) sequence data set to
105 provide a more complete view of *M. dolomieu* population genetic patterns and
106 colonization history across the Great Lakes and adjacent river systems.

107 The project aim was to identify the factors that shaped the population genetic
108 patterns observed today for *M. dolomieu*, comparing and contrasting results from two
109 genomes. Similarities and differences of genetic and geographic factors from mtDNA *cytb*
110 sequences and eight nuclear microsatellites are investigated for 18 *M. dolomieu*
111 populations across once-glaciated areas (Fig. 1). The following questions are addressed:
112 are similar levels of genetic diversity found among *M. dolomieu* populations in both
113 genomes; are their patterns of genetic divergence congruent; is genetic variation
114 significantly partitioned among watersheds, lakes, rivers and basins; do patterns of

differentiation reflect isolation by geographic distance, measured by nearest waterway connectivity?

MATERIALS AND METHODS

SAMPLE COLLECTION, DNA EXTRACTION, AMPLIFICATION AND SEQUENCING

Spawning-condition adult *M. dolomieu* were assayed from 18 native population locations, avoiding areas that had been stocked and surveying a wide variety of habitats in the respective locales (Table I). Collections were made by federal, state and provincial agencies, colleagues and laboratory members (under their respective regulations and permits), with those by the Stepien Laboratory under permits issued from the Michigan Department of Natural Resources and Ohio Division of Wildlife (#140160), with the University of Toledo approved Institutional Animal Care and Use Committee (IACUC) protocol (#105400). Fin-clip tissues were preserved in 95% ethanol in the field and stored at room temperature in the Stepien Laboratory. Genomic DNA was extracted and purified using DNeasy Qiaquick kits (Qiagen; www.qiagen.com). Aliquots were labelled, frozen at –80° C and archived.

The entire mtDNA cytochrome oxidase subunit *b* (*cytb*) gene (1140 bp) was sequenced and analysed from 225 individuals, which were a representative subset of 357 individuals whose nuclear microsatellite variation was the focus of Stepien *et al*. (2007). The *cytb* gene was amplified using the polymerase chain reaction (PCR) and sequenced in both directions using the primers *Song-F* and *Song-R* (Song *et al.,* 1998), *Basscytbf-1*

(Near *et al.,* 2003) and *Smb-2r* (5′-CCCTTAGTAACTCCGCCACAC–3′, developed by the Stepien Laboratory). Sequencing reactions were performed at the Cornell University Life Sciences Core Laboratories Center (http://www.biotech.cornell.edu/biotechnology-resource-center-brc), which used automated Applied Biosystems 3730 (ABI; www.appliedbiosystems.com) with big dye terminator chemistry sequencing. The sequences were then aligned with BIOEDIT 7.25 (http://www.mbio.ncsu.edu/bioedit/bioedit.html; Hall, 1999;). Unique haplotypes were identified and compared with Stepien Laboratory databases and other *M. dolomieu* sequences (Near *et al.,* 2003; Roe *et al.,* 2002) from GenBank (http://www.ncbi.nlm.nih.gov/Genbank). Accession numbers for haplotype sequences in the present study are: 1 = KU171303, 2 = KU171304, 3 = KU171305, 4 = KU171306, 5 = KU171307, 6 = KU171308, 7 = KU171309, 8 = KU171310, 9 = KU171311, 10 = KU171312, 11 = KU171313, 12 = KU171314, 13 = KU171315, 14 = KU171316, 15 = KU171317, 16 = KU171318, 17 = KU171319, 18 = KU171320, 19 = KU171321, 20 = KU171322, 21 = KU171323, 22 = KU171324, 23 = KU171325, 24 = KU171326, 25 = KU171327, 26 = KU171328, 27 = KU171329 and 28 = KU171330.

Eight nuclear microsatellite DNA loci: Mdo-2, Mdo-8, Mdo-9 (Malloy *et al*., 2000) Mdo-3, Mdo-5, Mdo-11 (Malloy *et al.,* 2000; Coughlin *et al.,* 2003), RB-7 (DeWoody *et al.,* 1998) and MS-19 (DeWoody *et al.,* 2000) were reanalysed from populations reported in Stepien *et al.* (2007).

MITOCHONDRIAL DNA CYTOCHROME *B* DATA ANALYSIS

Phylogenetic relationships among mtDNA *cytb* haplotypes were analysed with MRBAYES 3.2.2 using a metropolis-coupled Markov chain Monte Carlo (MCMC) approach (Huelsenbeck & Ronquist, 2001; Ronquist *et al.,* 2011; http://mrbayes.sourceforge.net). The corrected Akaike information criterion (AICc) from jMODELTEST 2 (Guindon & Gascuel, 2003; Darriba *et al.,* 2012; https://github.com/ddarriba/jmodeltest2) was employed to assess the most appropriate model of substitution, which selected the Tamura & Nei (1993) model, with unequal frequencies and invariable sites ($I = 0.893$). Analyses were run for 10 million generations, with sampling every 100 iterations. Metropolis-coupled MCMC burn-in used log-likelihood values per generation to identify when stationarity was reached; 25% of the generations were discarded, along with trees and values prior to burn-in. A 50% majority-rule consensus tree was based on the remaining generations and branch support was obtained *via* posterior probability distribution (Holder & Lewis, 2003). The tree was rooted to the putative sister species spotted bass *Micropterus punctulatus* (Rafinesque 1819) (Johnson *et al.,* 2001; Near *et al*. 2004, 2005; Bagley *et al.*, 2011), using GenBank #HM070928.1 from Bagley *et al.* (2011). Relationships among the haplotypes were further depicted and compared using a statistical parsimony haplotype network in PopART (http://popart.otago.ac.nz/index.shtml; Leigh & Bryant, 2015).

MICROSATELLITE DNA DATA ANALYSIS

Genotype frequencies for previously collected microsatellite data by Stepien *et al.* (2007) were re-evaluated for conformance to Hardy–Weinberg equilibrium (HWE) expectations at each locus, with significance estimated with GENEPOP using MCMC and 1000 randomizations (Guo & Thompson, 1992). Deviations from HWE were tested for heterozygote deficiency and null (non-amplified) alleles in MICRO-CHECKER 2.2.3 (http://www.norwichresearchpark.com/research1/researchgroups/elsa/software/microchecker.aspx; van Oosterhout *et al.,* 2004). Genotype frequencies also were analysed for evidence of linkage disequilibrium (LD). Significance levels for HWE and LD were adjusted using sequential Bonferroni correction (Rice, 1989). LOSITAN (http://popgen.net/soft/lositan/; Beaumont & Nichols, 1996; Antao *et al.,* 2008) was used to evaluate whether loci were under possible selection.

Numbers of alleles (N_A) and allelic richness (A_R; number of alleles per locus, adjusted for sample size using rarefaction following Mousadik & Petit (1996)) were calculated with FSTAT. Numbers of full-siblings were estimated with COLONY 2.0.5.0 (http://www.zsl.org/science/software/colony; Jones & Wang, 2009). Levels of relatedness further were evaluated by computing population F_{IS} values in GENEPOP and by computing internal relatedness (IR; Amos *et al.* 2001).

Demographic partitioning of genetic variation was examined using STRUCTURE 2.3.4 (http://pritchardlab.stanford.edu/structure.html; Pritchard *et al.,* 2000; Pritchard & Wen, 2004), which assessed membership of individuals to $K = 1$ to $K = 18$ possible population groups (up to the total N of populations), independent of sampling location identity. Results were analysed using 10 independent runs for each, with burn-ins of 100 000 and 500 000 MCMC replicates, an admixture model, initial inferred ± = 1.0,

assumption of correlated alleles, prior F_{ST} mean = 0.01 and a S.D. prior = 0.05. Optimal K was determined *via* the ´-*K* likelihood method (Evanno *et al.,* 2005). GENECLASS2 (http://www1.montpellier.inra.fr/URLB/geneclass/geneclass.html; Piry *et al.,* 2004) assessed self-assignment of individuals to population samples, using a simulated population size of 10 000 individuals per location and a 0.01 rejection level (Cornuet *et al.,* 1999).

ANALYSES FOR BOTH GENOMES

Numbers of private (unique) haplotypes (N_{PH}) and private alleles (N_{PA}), *i.e.* those occurring only in a single spawning group, were determined with CONVERT 1.31 (http://www.agriculture.purdue.edu/fnr/html/faculty/rhodes/students and staff/glaubitz/software.htm; Glaubitz, 2004). *F*-statistics and their associated levels of significance were compared with FSTAT 2.9.3.2 (http://www2.unil.ch/popgen/softwares/fstat.htm; Goudet, 2002) to discern whether population samples differed genetically using pairwise ¸ $_{ST}$ (Weir & Cockerham, 1984) and exact *G* nonparametric comparisons; both were executed in GENEPOP 4.2 (http://kimura.univ-montp2.fr/~rousset/Genepop.htm; Raymond & Rousset, 1995; Rousset, 2008). Exact *G*-test comparisons were performed since they are non-parametric, have fewer assumptions and are less sensitive to sample-size effects. The *F*-based tests aided comparisons with levels of variation from other population genetic studies. Significance-associated probabilities of both types of tests were adjusted *via* sequential Bonferroni corrections to minimize type I errors (Rice, 1989).

To further delineate patterns among populations, the relationships between genetic distances [$_{ST}(1-_{ST})^{-1}$] and geographic distances were evaluated separately for both genomic marker sets using the ln of geographic distances, measured as the shortest waterway distances (km). Regression significance was interpreted from 1000 GENEPOP permutations. Relationships between adjacent sampling sites additionally were assessed with Barrier 2.2 (http://ecoanthropologie.mnhn.fr/software/barrier.html; Manni *et al.,* 2004) to identify genetically discontinuous assemblages, independent from *a priori* knowledge of relationships. The barriers were ranked by bootstrap support from 2000 iterations in GENELAND 4.0 (http://www2.imm.dtu.dk/~gigu/Geneland/; Guillot *et al.,* 2005*a*,*b*, 2008) using R 2.13.1 (www.r-project.org). Barriers with 50% bootstrap support and a majority of loci supporting were reported.

Hierarchical partitioning of genetic variation among possible grouping scenarios was tested with analysis of molecular variance (AMOVA; Excoffier *et al.,* 1992) in ARLEQUIN 3.5.1.3 (Excoffier & Lischer, 2010). Alternatives that were evaluated included distribution of variation among the population groups located in either lakes or rivers and between sites within *v.* outside of the Great Lakes drainage.

RESULTS

GENETIC DIVERSITY (QUESTION 1)

All microsatellite loci were found to be unlinked and genotype frequencies conformed to HWE expectations after Bonferroni correction. There were 28 *cytb* haplotypes (mean ± S.E. H_D = 0·50 ± 0·06) and 478 microsatellite alleles (mean ± S.E. H_O =

247 0·46 ± 0·03) among *M. dolomieu* from the 18 population sampling sites (Table I). Mean
248 values per sampling location were 3·4 haplotypes (range = 1–7) and 26·6 microsatellite
249 alleles (range = 14–38; Table I).

250 Haplotype 1 was the most common in *M. dolomieu* (28% overall), occurring at
251 sampling sites G–R (ranging from Georgian Bay in Lake Huron (site G) through to Paint
252 Creek in the Ohio River (site R)), followed by haplotype 2 (14% overall) that was
253 observed at sites A, G–I and L–O (Fig. 2 and Supporting Information Table SI. All
254 haplotypes were closely related, diverging from one another by a mean of 3·94 nucleotides
255 (range = 1–8). No haplotypes were separated by more than four nucleotide positions from
256 the most frequent haplotype 1. Just four haplotypes varied by four substitutions from
257 haplotype 1; two of these were located in the upper Mississippi drainage and were closely
258 related to one another (haplotypes 7 and 20), whereas the other two coincided with
259 haplotype 1 in Cattaraugus Creek from eastern Lake Erie (site L, haplotype 19) and in
260 Sandusky Bay from central Lake Erie (site I, haplotype 25). Ten of the 18 population
261 samples contained at least one private haplotype, with the highest proportion (0·60)
262 located in the Cannon River (site A) of the upper Mississippi River system. The greatest
263 number of haplotypes and haplotypic diversity comprised seven haplotypes and H_D = 0·85
264 for the Grand River collection from eastern Lake Erie (site K). Other spawning
265 populations with high haplotypic variability were Georgian Bay in Lake Huron (site G; H_D
266 = 0·83 and six haplotypes) and Saginaw Bay in Lake Huron (site F; H_D = 0·71, four
267 haplotypes).

268 Similarly, mean ± S.E. nuclear microsatellite heterozygosity (H_O) for all population
269 samples was 0·46 ± 0·03, ranging from 0·15 in Little Moose Lake (site P) to 0·60 for St

Louis Bay (site D). The highest proportions of private alleles (0·08) characterised the Chagrin River (site J) and St Croix River (site C). Other locations with private alleles included Canon River (site A), Apple River (site B), Anchor Bay (site H), Hudson River (site Q) and Paint Creek (site R). Proportions of full siblings inferred per site varied from none (sites A–D, F, I, J, L, N and R) to 0·67 (sites E and P), with mean ± S.E. = 0·15 ± 0·05 over the entire range (Table I). For the microsatellite data, high levels of relatedness (*i.e.* inbreeding) were inferred based on positive values from internal relatedness estimations for all population samples surveyed (Table I). F_{IS} estimates showed a less uniform pattern, with some (but not all) population samples having positive values (again indicating possible inbreeding).

GENETIC DIVERGENCE

Genetic divergences among *M. dolomieu* populations across major regions were supported by all comparisons in both data sets (Table II). Fine-scale analyses found that most population samples were genetically distinct, with 67·3% of the *cytb* and 90·2% of the microsatellite comparisons being significantly different based on *F*-statistics (Supporting Information Table SII). According to exact *G*-test comparisons, 75·0% of the *cytb* and 92·0% of the microsatellite comparisons significantly differed (Supporting Information Table SIII). Of these comparisons, results from the *cytb* and microsatellite data sets were mostly congruent, with the *F*-statistics data set finding the same 99 comparisons significant (64·7%) after sequential Bonferroni correction and the *G*-test comparisons revealing the same 120 comparisons significant with correction (78·4%).

292 Borderline and non-significant values may have been influenced by smaller sample sizes.
293 The microsatellite data delineated 39 (25·5%) instances of significance where mtDNA did
294 not, whereas mtDNA found four instances of significant differences (2·6%) where
295 microsatellites did not.

296 PARTITIONING OF VARIANCE AMONG CATCHMENTS, LAKES, BASINS and
297 RIVERS

298 Occurrences and frequencies of mtDNA *cytb* haplotypes [Figs. 2 and 3(a) and
299 Supporting Information Table SI often were distinctive among populations and regions.
300 Haplotype 1 was widespread (64/225, *i.e.* 28% of individuals), being prevalent throughout
301 most of the Great Lakes drainage, as well as in the Hudson River (site Q) and Paint Creek
302 of the Ohio River (site R). Haplotype 2 (32/225, *i.e.* 14% of individuals) was distributed
303 sporadically across a wide geographic range throughout most of the Great Lakes drainage
304 and as far west as the Cannon River (site A). Haplotype 4 was most common in the upper
305 Mississippi River drainage and also occurred in Lake Superior. Haplotypes unique to
306 specific regions included: 7, 20 and 23 in the Mississippi River Drainage, 17 in Lake
307 Superior, 11,16, 21 and 27 in Lake Huron, 22 in Lake St Clair, 9, 10, 14, 15, 19, 24, 25, 26
308 and 28 in Lake Erie, 13 in the St Lawrence River and 12 in Paint Creek.

309 The parsimony network (Fig. 2) indicated the prevalence of haplotype 1, which
310 was distributed from Georgian Bay in Lake Huron (site G) to Paint Creek in the Ohio
311 River (site R). Nine haplotypes (32% of haplotypes) varied by just single nucleotides from
312 haplotype 1, including haplotypes 3, 9, 10, 14, 22 and 28, which also were the most
313 abundant. The most genetically distant haplotypes (from haplotype 1) were 7, 19, 20 and
314 25, diverging by four nucleotide substitutions. The Mississippi River grouping (sites A–C,

315 haplotypes 4, 7, 20, 23) was the most defined cluster from haplotype 1, differing by three
316 to four steps.

317 The Bayesian phylogenetic tree of the 28 haplotypes (Fig. 4) illustrated relatively
318 close relationships among most haplotypes. There were three well-supported clades; one
319 linked three haplotypes (5, 18 and 19) uniquely found in Paint Creek (site R), another
320 contained six haplotypes (6, 7, 23, 4, 20 and 27) that mostly occurred in the upper
321 Mississippi River and Lake Superior (sites A–F, J) and a third clade possessed six
322 haplotypes (2, 13, 15, 16, 15 and 26) mostly from the lower Great Lakes (sites H–O).
323 Haplotypes 7 and 23 from the Cannon River (site A) also grouped together as a well-
324 supported clade.

325 Bayesian analyses of microsatellite data using STRUCTURE [Fig. 3(b)] identified
326 $K = 5$, 9 and 14 population clusters (Supporting Information Figs. S1 and S2), with $K = 5$
327 being the optimal number determined by the delta K method (Evanno *et al.,* 2005). Neither
328 microsatellite nor *cytb* sequences differentiated among adjacent sites in Lake Erie. Results
329 from the GENECLASS analysis indicated high self-assignment at most spawning
330 locations, with the exceptions of Georgian Bay (site G), Sandusky Bay (site I),
331 Cattaraugus Creek (site L) and Cranberry Lake (site O), in which higher proportions of
332 individuals assigned to other spawning groups (Table SIV). Overall, however, these
333 results were consistent with those found by STRUCTURE (FIG. S2).

334 AMOVA partitioning of genetic variation supported significant primar divergence
335 between the two population groupings from within *v*. outside the Great Lakes' drainage,
336 along with distinctive differentiation among their component populations (Table III).

There was no significant differentiation between population samples from lakes *v.* rivers (Table III).

Populations from outside the Great Lakes catchment were separated by the strongest barriers to gene flow inferred from Barrier analysis (Fig. 1). *Micropterus dolomieu* from the Hudson River (site Q) appeared very distinctive from other populations, having the most pronounced genetic barrier based on microsatellite evidence (I, 6 loci) and the second-greatest barrier based on *cytb* evidence (*cytb* 2). The upper Mississippi River drainage (sites A–C) was denoted by the strongest *cytb* inferred barrier (*cytb* 1) from the other populations and the third-greatest microsatellite inferred barrier (III, 7 loci). The second-largest microsatellite inferred barrier (II, 6 loci) divided the Cranberry Lake population (site O) from the others, a difference not shown by the *cytb* sequences. Within the Great Lakes, the strongest division was between populations from the upper (Superior, Michigan, Huron, St Clair; sites D–H) *v.* the lower Great Lakes (Erie, Ontario; sites I–N), which was recovered with both genomes (IV, 6 loci and *cytb* 3). The upper Mississippi River populations were subdivided further, with Cannon River (site A) distinguished by the fifth strongest *cytb* inferred barrier (*cytb* 5) and Cannon River and Apple Creek (Sites A and B) characterized by the fifth largest microsatellite inferred barrier (V, 5 loci). The Lake St Clair population was divided from the other upper Great Lakes (VI, 7 loci and *cytb* 4). The remaining barriers separated out the St Louis Bay (site D; VII, 6 loci and *cytb* 6), Ohio River (site R; *cytb* 7), St Lawrence River (sites O and P; VIII, 5 loci and *cytb* 9), Whitefish Bay (site E; X, 6 loci, *cytb* 8), Lake Ontario (site N; IX, 5 loci) and Sandusky Bay populations (site I; *cytb* 10).

ISOLATION BY GEOGRAPHIC DISTANCE

Relationships among *M. dolomieu* populations followed a pattern of genetic isolation by geographic distance, with both the mtDNA and the microsatellite data sets (Fig. 5). Although the largest genetic differences between populations matched those with large geographic demarcations, there were also a number of sampling sites separated by large geographic distances that exhibited little to no genetic difference. Many higher-than-average pairwise comparisons for the *cytb* data were associated with sites B (Apple Creek) and C (St Croix River) for 8/10 comparisons and with site P (Little Moose Lake) for 4/10 comparisons [Fig. 5(a)]. All 11 higher-than-average pairwise comparisons for microsatellites were associated with site P [Fig. 5(b)].

DISCUSSION

GENETIC DIFFERENTIATION AND RELATION TO QUESTIONS

This investigation discerns strong patterns of differentiation among spawning groups of *M. dolomieu* from 18 population sites across the North American Great Lakes and adjacent areas, based on both mitochondrial and nuclear DNA data. These results help to answer the original questions.

Do similar levels of genetic diversity characterize M. dolomieu populations with both genomes?

Results demonstrate that overall levels of genetic diversity are very similar between the two genomes. High nuclear microsatellite diversity, however, did not always coincide

with high mtDNA variability and *visa versa*. Notably, the population from Georgian Bay in Lake Huron (site G) possessed high mtDNA diversity (0·83) yet had modest microsatellite variability (0·38), whereas population samples from Apple Creek (B) and the St Croix River (C) in the upper Mississippi River system were monomorphic in mtDNA (0·00), yet contained average levels of nuclear microsatellite variability (0·49 and 0·55, respectively). It is possible that this discord reflects sample-size effects or the respective evolutionary mutational scales of mitochondrial *v*. nuclear DNA genomes, which would provide an intriguing avenue for further investigation.

Are patterns of genetic divergence similar between the two M. dolomieu genomes?

Both genomes do indicate that populations from all major geographic regions diverge significantly from each other (Tables II and III) and almost all local population samples also differ significantly (Tables SI-IV). More fine-scale differences were resolved with the microsatellite results, probably due to their having multiple loci and more alleles and consequently greater discrimination power. Results from the mtDNA analysis probably reflect population origins that trace to historical expansion from glacial refugia rather than recent genetic differentiation. Nevertheless, both data sets reveal congruent patterns.

Is genetic variation significantly partitioned among catchments, rivers, lakes and basins?

Results do indeed indicate that the strongest barriers to gene flow separate populations outside of the Great Lakes from those within. Significant divergences among many population sites distinguish *M dolomieu* from river systems, lakes and basins. Differential occurrence and distribution of recovered mitochondrial haplotypes denote indigenous population groups from lakes, rivers and geographic regions that may reflect the influences of their historical glacial refugia origins and colonization pathways.

Are patterns of differentiation related to genetic isolation by geographic distance using both genomes?

Both data sets support the hypothesis of differentiation with geographic distance. This indicates that geographic demarcation among population sites significantly influenced the genetic divergence among spawning groups. Since this correlation was relatively weak, however, colonization from glacial refugia also probably affected the patterns of relationships among populations.

COMPARISONS WITH DIVERSITY OF OTHER GREAT LAKES FISHES

Micropterus dolomieu populations in the Great Lakes (mean ± S.E. *H*o = 0·46 ± 0·03; range = 0·15–0·60) display lower observed microsatellite heterozygosity than did some other popular fishery species with overlapping geographic ranges, such as *P. flavescens* (mean ± S.E. *H*o = 0·53 ± 0·01, range = 0·39–0·63; Sepulveda-Villet & Stepien, 2012; Stepien *et al.* 2015*a*,*b*) and *S. vitreus* (mean ± S.E. *H*o = 0·68 ± 0·03, range = 0.52–0.77, Haponski & Stepien, 2014; Stepien *et al.* 2015*b*). This may be related to smaller contemporary spawning population sizes and less genetic migration of *M. dolomieu* individuals, leading to more apparent inbreeding and reduced genetic variability *via* drift. mtDNA haplotypic diversity of *M. dolomieu* (mean ± S.E. H_D = 0·50 ± 0·06, range = 0–0·85) is almost as high as *S. vitreus* (mean ± S.E. H_D = 0·53 ± 0·02, range = 0·15–0·88), with both being considerably greater than for *P. flavescens* (mean ± S.E. H_D = 0·31 ± 0·05, range = 0–0·82). Each of these three species possessed a single dominant haplotype in the Great Lakes. This pattern probably resulted from their common phylogeographic and

428 glacial colonization histories and it is interesting that all three species appear to have
429 evolved similarly in response to these historical events.

430 Proportions of full siblings in *M. dolomieu* samples (mean ± S.E. p_S = 0·15 ± 0·05,
431 range = 0–0·67) were similar to those for *P. flavescens* (mean ± S.E. p_S = 0·16 ± 0·02,
432 range = 0–0·38; Sullivan & Stepien, 2015) and about twice those characterizing *S. vitreus*
433 populations (mean ± S.E. p_S = 0·08 ± 0·02, range = 0–0·76; Haponski & Stepien, 2014).
434 This appears to indicate a tendency for proximate fish to be members of family groups and
435 may reflect how limited overall lifetime migration patterns affect gene flow among
436 locales. This is supported by the consistent positive internal relatedness (IR) values
437 estimated for each population sample. The moderate to high relatedness in these sampled
438 populations might indicate inbreeding, although significant loss of diversity does not seem
439 to have occurred. *Micropterus dolomieu* appears to disperse the least of these three
440 species, having a maximum lifetime dispersal reported as 10 km (Lyons & Kanehl, 2002).
441 In comparison, *P. flavescens* may migrate up to 48 km (with the occasional individual
442 travelling as far as 200 km; Rawson, 1980) and *S. vitreus* up to 300 km (Colby *et al.,*
443 1979). A smaller lifetime dispersal distance probably lessens gene flow among regions
444 and may increase inbreeding among localized populations. For example, Ridgway *et al.*
445 (1991) discovered that adult male *M. dolomieu* displayed high breeding site fidelity, with
446 81% of the males building nests within 200 m of their previous year's nest site. Since nest
447 locations were stable, lack of migration by males may limit gene flow and preserve private
448 alleles within a given population. Similarly, Gross & Kapuscinski (1997) discerned that
449 over 50% of spawned individuals were produced by < 10% of available males in Lake
450 Opeongo, Ontario, Canada. It is possible the pronounced success of few individuals could

contribute to the high degree of relatedness detected in results presented here. As discussed by Bouzat (2010), increased inbreeding often is associated with lower genetic diversity; but this may not always be the case. *Micropterus dolomieu* may use cues (*e.g.* olfactory and visual) to distinguish close relatives and avoid inbreeding, thereby reducing the possible negative effects of inbreeding depression. For example, such sensory clues appear to allow the perch *Perca fluviatilis* L. 1758 to identify kin from non-relatives and circumvent inbreeding (Behrmann-Goedel *et al*., 2006; Behrmann-Godel & Gerlach, 2008). Whether such factors also may occur in *M. dolomieu* may be important for the long-term sustainability of population groups and should be investigated further.

COMPARISONS WITH THE DIVERGENCE PATTERNS OF OTHER GREAT LAKES FISHES

Populations of *M. dolomieu* display evidence for origins from the three glacial refugia (Missourian, Mississippian and Atlantic) in this study, which also characterized *S. vitreus* (Stepien *et al.,* 2009, 2015*a*, *b*; Haponski & Stepien, 2014), *P. flavescens* (Sepulveda-Villet & Stepien, 2012, Sullivan & Stepien, 2014, 2015) and several other Great Lakes fishes (April *et al.,* 2013). Those refugia provided retreat from climatic, habitat and geophysical changes that occurred during the Pleistocene glaciations (Hewitt, 2004; Provan & Bennet, 2008). The upper Mississippi River populations of *M. dolomieu* traced to the Missourian refugium, as likewise was detected for *A. nebulosus* of western Lake Superior (Murdoch & Hebert, 1997). The majority of all Great Lakes *M. dolomieu* populations appear to have originated from the Mississippian refugium, as was the case for most Great Lakes' species (Mandrak & Crossman, 1992; Todd & Hatcher, 1993). Eastern populations of *M. dolomieu* from the St Lawrence and Hudson Rivers traced their

474 colonization to the Atlantic refugium, which provided protection in diverse habitats east of
475 the Appalachian Mountains (Bernatchez, 1997). Similar to findings for *P. flavescens* by
476 Sepulveda-Villet & Stepien (2012), haplotype 1 was probably widespread prior to the
477 glaciations, leading to its currently ubiquitous distribution in sites that were colonized
478 from both the Mississippian and Atlantic refugia. These parallel broad-scale patterns for
479 *M. dolomieu*, *S. vitreus* and *P. flavescens* indicate their shared phylogeographic history
480 (*i.e.* congruent glacial refugia and colonization pathways) and point to the importance of
481 maintaining these genetic demarcations among key population groups for conservation
482 planning to protect the natural population structure of these exploited fishery species.

483 The present study discerns that large genetic breaks distinguish populations outside
484 the Great Lakes from those within and delineate those in the upper from the lower Great
485 Lakes. This pattern is similar to that recovered for *P. flavescens* (Sepulveda-Villet &
486 Stepien, 2012), but varies from *S. vitreus* (Haponski & Stepien, 2014). The latter
487 possessed many barriers separating populations outside of the Great Lakes, as well as
488 several within them, but lacked the demarcation between the upper and the lower Great
489 Lakes. This is probably due to differences in migration patterns and spawning-site fidelity
490 among the three species.

491

492 RELEVANCE TO OTHER *M. DOLOMIEU* STUDIES

493 Although *M. dolomieu* is an economically and ecologically important species in
494 the Great Lakes and elsewhere, its genetic diversity and broad and fine-scale population
495 relationships are relatively understudied. Borden & Stepien (2006) described moderate
496 variability and some genetic divergence among groups using partial mtDNA sequences

497 and microsatellite loci, but was limited by low-resolution power and small sample sizes.
498 Stepien *et al.* (2007) examined broad-scale patterns using eight microsatellite loci and
499 larger samples, discerning differences among the Great Lakes populations, as well as for
500 river *v.* lake sites. Bagley *et al.* (2011) examined mtDNA *cytb* haplotypes of populations
501 in Arkansas and southern Missouri, U.S.A. That study recovered 17 distinct haplotypes
502 (three of which were shared with the present analysis: haplotypes 3–5). Hallerman *et al.*
503 (2015) discovered that *M. dolomieu* from the Tennessee River drainage were markedly
504 differentiated from populations in the New and James River drainages, which may be
505 related to different glacial refugia origins (Mississippian *v*. Highlands) as well as artefacts
506 of past stocking supplementation.

507 The present investigation indicates that *M. dolomieu* possesses moderately high
508 genetic diversity, strong broad-scale patterns related to post-glacial colonization and
509 pronounced fine-scale population divergence, which are probably due to a combination of
510 genetic isolation by geographic distance, limited migration and spawning-site fidelity.
511 These results, coupled with those of earlier studies, indicate that *M. dolomieu* population
512 structure and genetic diversity stem from interplay between phylogeographic history and
513 contemporary processes that are in many ways similar to other North American fish
514 species (Stepien *et al.,* 2009, 2015*a*, *b*; Sepulveda-Villet & Stepien, 2012; Haponski &
515 Stepien, 2014; Sullivan & Stepien 2014, 2015).

516 In conclusion, this investigation identifies a number of important aspects
517 concerning the broad and fine-scale population genetic patterns of *M. dolomieu* spawning
518 groups. First, moderate genetic diversity and strong population genetic divergence
519 characterize both genomes. Broad-scale population genetic patterns of *M. dolomieu*

indicate the probable influence of post-glacial recolonization patterns tracing to Missourian, Mississippian and Atlantic refugia. Robust fine-scale genetic patterns are discerned among population groups, which are related to a combination of factors, including geographic demarcation, limited dispersal and spawning-site fidelity. These results denote the importance of both historical and contemporary influences on population structure in an exploited fishery species. Lastly, this enhanced understanding of the contemporary influences on population structure, specifically the interplay among behavioural philopatry, inbreeding and genetic diversity may provide useful information to help manage the demographic variability and viability of *M. dolomieu* effectively.

Research was funded by grants to C.A.S. from NOAA Ohio Sea Grant R/LR-13, USEPA GL-00E01149-0 (which partially supported K.E.K.) and GL-00E01898, NSF URM #0829252 (Undergraduate Research & Mentoring in Environmental Biology at the Land-Lake Ecosystem Interface, which supported S.I.K.) and NSF GK-12 DGE#0742395 (which supported T.J.S.). Samples either were collected by Stepien's Great Lakes Genetics–Genomics Laboratory members or were provided by various colleagues and agencies, including: D. Bright, R. Butryn, A. Ford, D. Gonder, M. Greenlee, M. Jedlicka, O. Johnson, J. Junosko, D. Kiss, B. Kulik, J. Lindgren, T. Parrett, M. Sanderson, K.

Schmidt, G. Steinhart, D. Swartel, M. Thomas and W. Zawiski, as well as Ontario Ministry of Natural Resources, New York Department of Environmental Conservation and Michigan Department of Natural Resources. Stepien's Great Lakes Genetics–Genomics Laboratory members J. Brown and D. Murphy aided primer design, with A. Haponski, V. Palsule, L. Pierce, C. Prichard and S. Woolwine providing valuable comments on manuscript drafts. M. Gray, R. Lohner, S. McBride, P. Uzmann and V. Sigler gave logistic assistance. A. Elz in CAS' new NOAA PMEL laboratory provided final proof-reading corrections. This is contribution 4609 from NOAA Pacific Marine Environmental Laboratory (PMEL).

Supporting Information

Supporting information can be found in the online version of this paper.

Table SI. Frequencies of *Micropterus dolomieu* population samples' mtDNA *cytb* sequence haplotypes.

Table SII. Genetic divergence patterns among smallmouth bass *Micropterus dolomieu* population samples (see Fig. 1 for locations) measured by pairwise comparisons (, $_{ST}$) based on mtDNA *cytb* sequence variation (below diagonal) and on eight nuclear DNA microsatellite loci (above diagonal). Values with black backgrounds remained significant upon sequential Bonferroni correction, and values with grey background were significant prior to sequential Bonferroni correction, but not after. Values shown on white were not significant.

564 **Table SIII.** Exact *G*-test values among all locations (A–R, see Fig. 1) based on eight
565 nuclear microsatellite loci (above diagonal) and mitochondrial DNA *cytb* sequences
566 (below diagonal). Black background indicates significant values ($P < 0.05$).
567
568 **Table SIV.** Results of Bayesian assignments to *Micropterus dolomieu* population samples
569 A–R (see Fig. 1) based on eight microsatellite loci using GENECLASS2. Underlined
570 values denote the percentage of individuals that self-assigned to source collection.
571
572 **Fig. S1.** Results for each possible number of *K* clusters based on STRUCTURE
573 HARVESTER (Earl & von Holdt, 2012) are shown for ” *K*, indicating greatest support for
574 *K*=5.
575
576 **Fig. S2**. Smallmouth bass *Micropterus dolomieu* population samples A–R (see Fig. 1)
577 estimates from Bayesian assignments using STRUCTURE 2.3.4, (a) $K = 9$ and (b) $K = 14$
578 for eight microsatellite loci. Individuals are represented by thin vertical bars coloured
579 according to inferred group membership. Vertical black lines separate different population
580 samples (spawning groups).
581
582
583
584
585
586
587
588

589 **References**

590

591 Amos, W., Worthington W. J., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D., &
592 Coulson, T. (2001). The influence of parental relatedness on reproductive success.
593 *Proceedings of the Royal Biological Society B*. **268**, 2021–2027.
594 Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A. & Luikart, G. (2008).
595 LOSITAN: A workbench to detect molecular adaptation based on
596 a F_{ST} outlier method *BMC Bioinformatics* **2008**, 323. doi:10.1186/1471-2105- 9-
597 323.
598 April, J., Hanner, R. H., Dion-Côté A. M., & Bernatchez, L. (2013). Glacial cycles as an
599 allopatric speciation pump in north-eastern American freshwater fishes. *Molecular*
600 *Ecology* **22**, 409–422.
601 Avise, J., C. (2000). *Phylogeography: the History and Formation of Species.* Cambridge,
602 MA: Harvard University Press.
603 Bagley, J. C., Mayden, R. L., Roe, K. J., Holznagel, W. & Harris, P.M. (2011).
604 Congeneric phylogeographical sampling reveals polyphyly and novel biodiversity
605 within black basses (Centrarchidae: *Micropterus*). *Biological Journal of the*
606 *Linnaean Society* **104**, 346–363.
607 Baker, C. S., Steel, D., Calambokidis, J., Falcone, E., González-Peral, U., Barlow, J.,
608 Burdin, A. M., Clapham, P. J., Ford, K. K. B., Gabriele, C. M., Mattila, D., Rojas-
609 Bracho, L., Straley, J. M., Taylor, B. L., Urbán, J., Wade, P. R., Weller, D.,
610 Witteveen, B. H., & Yamaguchi, M. (2013). Strong maternal fidelity and natal

philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology Progress Series* **494**, 291–306.

Baker, W.H., Blanton, R.E., & Johnson, C. E. (2013). Diversity within the redeye bass, *Micropterus coosae* (Perciformes: Centrarchidae) species group, with descriptions of four new species. *Zootaxa* **3635**, 379–401.

Beaumont, M. A. & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B* **263**, 1619–1626.

Behrmann-Godel, J., Gerlach, G., & Eckmann, R. (2006). Kin and population recognition in sympatric Lake Constance perch (*Perca fluviatilis* L.): can assortative shoaling drive population divergence*? Behavioral Ecology and Sociobiology* **59**, 461–468.

Behrmann-Godel, J. & Gerlach, G. (2008). First evidence for postzygotic reproductive isolation between two populations of European perch (*Perca fluviatilis* L.) within Lake Constance. *Frontiers in Zoology* **5**, 1–8.

Bernatchez, L. & Dodson, J. J. (1991). Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. *Evolution* **45**, 1016–1035.

Bernatchez L. (1997). Mitochondrial DNA analysis confirms the existence of two glacial races of rainbow smelt *Osmerus mordax* and their reproductive isolation in the St. Lawrence River estuary (Quebec, Canada). *Molecular Ecology* **6**, 73–83.

Boizard, J., Magnan, P. & Angers, B. (2009). Effects of dynamic landscape elements on fish dispersal: the example of the creek chub (*Semotilus atromaculatus*). *Molecular Ecology* **18**, 430–441.

633 Borden, W. C. & Stepien, C. A. (2006). Discordant population genetic structure of
634 smallmouth bass, *Micropterus dolomieu* Lacépède in Lake Erie based on
635 mitochondrial DNA sequences and nuclear DNA microsatellites. *Journal of Great*
636 *Lakes Research* **32**, 242–257.
637 Borden, W. C. & Krebs, R. A. (2009). Phylogeography and postglacial dispersal of
638 smallmouth bass (*Micropterus dolomieu*) into the Great Lakes. *Canadian Journal*
639 *of Fish and Aquatic Sciences* **66**, 2142–2156.
640 Bouzat, J. L. (2010). Conservation genetics of population bottlenecks: the role of chance,
641 selection and history. *Conservation Genetics* **11**, 463–478.
642 Colby, P. J., McNicol, R. E., & Ryder, R. A. (1979). Synopsis of biological data on the
643 walleye *Stizostedion v. vitreum* (Mitchell 1818*). FAO Fisheries Synopsis* FIR
644 **119**/S119 Rome.
645 Cornuet J. M., Piry S., Luikart G., Estoup A. & Solignac M. (1999). New methods
646 employing multilocus genotypes to select or exclude populations as origins of
647 individuals. *Genetics* **153**, 1989–2000.
648 Coughlin, W. D., Echelle, A. A., Van Den Bussche, R. A.. Cofer, L. M. Fisher, W. L. &
649 Propst, D. L. (2003). Genetic structure of spotted bass (*Micropterus punctualus*). in
650 the Red and Arkansas river basins: microsatellite and mitochondrial DNA
651 variation. *The Southwest Naturalist* **48**, 526–533.
652 Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jMODELTEST 2: more
653 models new heuristics and parallel computing. *Nature Methods* **9**, 772.
654 DeWoody, J. A., Fletcher, D. E., Wilkins, S. D., Nelson, W. S. & Avise J. C. (1998).
655 Molecular genetic dissection of spawning, parentage and reproductive tactics in a
656 population of redbreast sunfish, *Lepomis auritus*. *Evolution* **52**, 1802–1810.

DeWoody, J. A., Fletcher, D. E., Wilkins, S. D., Nelson, W. S. & Avise J. C. (2000). Genetic monogamy and biparental care in an externally fertilizing fish, the largemouth bass (*Micropterus salmoides*). *Proceedings of the Royal Society B.* **267**, 2431–2437.

Earl, D. A. & B. M. von Holdt. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**, 359–361.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611–2620.

Excoffier, L. & Lischer, H. E. L. (2010) Arlequin suite v3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–567.

Excoffier, L., Smouse, P.E. & Quattro, J. M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA data. *Genetics* **131**,479–491.

Fitzpatrick, S. W., Crockett, H. & Funk, W. C. (2014). Water availability strongly impacts population genetic patterns of an imperiled Great Plains endemic fish. *Conservation Genetics* **15**, 771–788.

Gibbard, P. L., Head, M. J. & Walker, J. C. (2010). Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science* **25**, 96–102.

679 Glaubitz J. C. (2004). CONVERT: a user-friendly program to reformat diploid genotypic
680 data for commonly used population genetic software packages. *Molecular Ecology*
681 *Notes* **4**, 309–310.
682 Gross, M. & Kapuscinski, A. (1997). Reproductive success of smallmouth bass estimated
683 and evaluated from family-specific DNA fingerprints. *Ecology,* **78**, 1424-1430.
684 Guillot, G., Estoup A., Mortier F. & Cosson J. F. (2005a). A spatial statistical model for
685 landscape genetics. *Genetics* **170**, 1261–1280.
686 Guillot, G., Mortier F. & Estoup A. (2005b). GENELAND: a computer package for
687 landscape genetics. *Molecular Ecology Notes* **5**, 712–715.
688 Guillot, G. (2008). Inference of structure in subdivided populations at low levels of
689 genetic differentiation – the correlated allele frequencies model revisited.
690 *Bioinformatics* **24**, 2222–2228.
691 Guindon, S. & Gascuel, O. (2003). A simple, fast, accurate algorithm to estimate large
692 phylogenies by maximum likelihood. *Systematic Biology* **52**, 696–704.
693 Guo, S. W. & Thompson E. A. (1992). Performing the exact test of Hardy-Weinberg
694 proportion for multiple alleles. *Biometrics* **48**, 361–372.
695 Gugerli, F., Sperisen, C., Büchler, U., Magni, F., Geburek, T., Jeandroz, S. & Senn, J.
696 (2001). Haplotype variation in a mitochondrial tandem repeat of Norway spruce
697 (*Picea abies*) populations suggests a serious founder effect during postglacial re-
698 colonization of the western Alps. *Molecular Ecology* **10**, 1255–1263.
699 Hallerman, E., Johnson, N. & Dutton, D. (2015). Implications of microsatellite DNA
700 variation upon management of Virginia smallmouth bass populations. In *Black*
701 *Bass Diversity: Multidisciplinary Science for Conservation* (Tringali, M., Long, J.

M., Birdsong, T. W. & Allen, M.S., eds), pp. 615-633. *American Fisheries Society, Symposium* **82**. Bethesda, MD: AFS.

Haponski, A. E. & Stepien, C. A. (2014). A population genetic window onto the past and future of the walleye *Sander vitreus*: relation to historic walleye and the extinct "blue pike" *S. v. "glaucus". BMC Evolutionary Biology* **2014**, 14:133.

Hedrick, P. W. & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecological Systems* **31**, 139–62.

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907–913.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B* **359**, 183–195.

Hall, T. A. (1999). BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95.

Holder, M. & Lewis P. O. (2003). Phylogeny estimation: traditional and Bayesian approaches. *Natural Reviews in Genetics* **4**, 275– 284.

Huelsenbeck, J. P. & Ronquist F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.

Johnson, R. L., Magee, J. B., & Hodge, T. A. (2001). Phylogenetics of freshwater black basses (Centrarchidae: *Micropterus*) inferred from restriction endonuclease analysis of mitochondrial DNA. *Biochemical Genetics* **39**, 395–406.

Jones, O. & Wang, J. (2009). COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**, 551–555.

724 Kutschera, V., Frosch C., Janke A., Skirnission K., Bidon T., Lecomte N., Fain S., Eiken
725 H., Arnason U. & Laidre K. (2016). High genetic variability of vagrant polar
726 bears illustrates importance of population connectivity in fragmented sea ice
727 habitats. *Animal Conservation* **19**, 337-349. doi:10.1111/acv.12250
728 Leigh, J. W. & Bryant, D. (2015). POPART: Full-feature software for haplotype network
729 construction. *Methods in Ecology and Evolution* **6**, 1110–1116.
730 Lindsay, D. L., Barr, K. R., Lance, R. F., Tweddale, S. A., Hayden, T. J. & Leberg, P. L.
731 (2008). Habitat fragmentation and genetic diversity of an endangered, migratory
732 songbird, the golden-cheeked warbler (*Dendroica chrysoparia). Molecular*
733 *Ecology* **17**, 2122–2133.
734 Lyons, J. & Kanehl, P. (2002). Seasonal movements of smallmouth bass in streams. In:
735 *Black Bass: Ecology, Conservation and Management* (Philipp, D.P. & Ridgway,
736 M.S., eds), pp. 149–160. Bethesda, MD: American Fisheries Society
737 Malloy, T. P., Van Den Bussche, R. A., Coughlin, W. D. & Echelle, A. A. (2000).
738 Isolation and characterization of microsatellite loci in smallmouth bass,
739 *Micropterus dolomieu* (Teleostei: Centrarchidae) and cross-species amplification
740 in spotted bass, *M. punctulatus*. *Molecular Ecology* **9**, 1946–1948.
741 Mandrak, N. E. & Crossman, E. J. (1992). Postglacial dispersal of freshwater fishes into
742 Ontario. *Canadian Journal of Zoology* **70**, 2247–2259.
743 Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics:
744 combining landscape ecology and population genetics. *Trends in Ecology and*
745 *Evolution* **18**, 189–197.
746 Manni, F. E., Guerard, E. & Heyer, E. (2004). Geographical patterns of (genetic,

morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology* **76**, 173-190. doi: 10.1353/hub.2004.0034

El Mousadik, A. & Petit, R. J. (1996). Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology* **5**, 547–555.

Murdoch, M. H. & Hebert, P. D. N. (1997). Mitochondrial DNA evidence of distinct glacial refugia for brown bullhead (*Ameiurus nebulosus)* in the Great Lakes. *Canadian Journal of Fish and Aquatic Sciences* **54**, 1450–1460.

Near, T. J., Kassler, T. W., Koppelmann, J. B., Dillman, C. B. & Philipp, D. P. (2003). Speciation in North American black basses, *Micropterus* (Actinopterygii: Centrarchidae). *Evolution* **57**, 1610–1621.

Near, T. J., Bolnick, D. I., & Wainwright, P.C. (2004). Investigating phylogenetic relationships of sunfishes and black basses (Actinopterygii: Centrarchidae) using DNA sequences from mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution* **32**, 344–357.

Near, T. J., Bolnick, D. I. & Wainwright, P. C. (2005). Fossil calibrations and molecular divergence time estimates in centrarchid fishes (Teleostei: Centrarchidae). *Evolution* **59**, 1768–1782.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10.

Nichols, H. J., Jordan, N. R., Jamie, G. A., Cant, M. A. & Hoffman, J. I. (2012). Fine-scale spatiotemporal patterns of genetic variation reflect budding dispersal coupled with strong natal philopatry in a cooperatively breeding mammal. *Molecular Ecology* **21**, 5348–5362.

van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535–538.

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L. & Estoup, A. (2004). GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536–539.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Provan, J., & Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution* **23**, 564–571.

Puckett, E. E., Kristensen, T. V., Wilton, C. M., Lyda, S. B., Noyce, K. V., Holahan, P. M., Leslie, D. M., Jr., Beringer J., Belant, J. L., White, D., Jr. & Eggert, L. S. (2014). Influence of drift and admixture on population structure of American black bears (*Ursus americanus*) in the central interior highlands, USA, 50 years after translocation, *Molecular Ecology* **23**, 2414–2427.

Rawson, M. R. (1980). Lake Erie Fisheries Investigations: yellow perch movements. Federal Aid in Fish Restoration Project Number F–35–R–18, Columbus OH: Ohio Department of Natural Resources. Division of Wildlife. 25 pp.

Raymond, M. & Rousset, F. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248–249.

Rice, W. R. (1989). Analysing tables of statistical tests. *Evolution* **43**, 223–225.

791 Ridgway, M. S., MacLean, J. A. & MacLeod, J. C. (1991). Nest-site fidelity in a
792 centrarchid fish, the smallmouth bass (*Micropterus dolomieu*). *Canadian Journal*
793 *of Zoology* **69**, 3103–3105
794 Roe, K. J., Harris, P. M. & Mayden, R. L. (2002). Phylogenetic relationships of the genera
795 of North American sunfishes and basses (Percoidei: Centrarchidae) as evidenced
796 by the mitochondrial cytochrome *b* gene. *Copeia* **2002**, 897–905.
797 Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget,
798 B., Liu, L., Suchard, M. A. & Huelsenbeck, J. P. (2011). MRBAYES 3.2: Efficient
799 Bayesian phylogenetic inference and model choice across a large model space.
800 *Systematic Biology*. **61**, 539–542.
801 Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP
802 software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–106.
803 Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*
804 **236**, 787–792
805 Sepulveda-Villet, O. J. & Stepien, C. A. (2012). Waterscape genetics of the yellow
806 perch (*Perca flavescens*): patterns across large connected ecosystems and isolated
807 relict populations. *Molecular Ecology* **21**, 5795–5826.
808 Stepien, C. A. & Faber, J. E. (1998). Population genetic structure, phylogeography and
809 spawning philopatry in walleye (*Stizostedion vitreum*) from mitochondrial DNA
810 control region sequences. *Molecular Ecology* **7**, 1757–1769.
811 Stepien, C. A., Murphy, D. J. & Strange, R. M. (2007). Broad- to fine-scale population
812 genetic patterning in the smallmouth bass *Micropterus dolomieu* across the

Laurentian Great Lakes and beyond: an interplay of behaviour and geography. *Molecular Ecology* **16**, 1605–1624.

Stepien, C. A., Murphy, D. J., Lohner, R. N., Sepulveda–Villet, O. J. & Haponski, A. E. (2009). Signatures of vicariance, postglacial dispersal and spawning philopatry: Population genetics and biogeography of the walleye *Sander vitreus*. *Molecular Ecology* **18**, 3411–3428.

Stepien, C.A., Behrmann-Godel, J. & Bernatchez, L. (2015*a*). Evolutionary relationships, population genetics and ecological and genomic adaptations of perch (*Perca*). In *Biology of Perch* (Couture, P. & G. Pyle, eds), pp. 7–46. Boca Raton, FL: CRC Press.

Stepien, C. A., O. J. Sepulveda-Villet, O. J & Haponski, A. E. (2015*b*). Comparative genetic diversity, population structure and adaptations of walleye and yellow perch across North America. In: *Biology and Culture of Percid Fishes – Principles and Practices* (Kestlemont, P. K. Dabrowski, K. & Summerfelt, R. C., eds), pp. 643–690. Berlin: Springer. doi 10.1007978-94-017-7227-3.

Song, C. B., Near, T. J. & Page, L. M. (1998). Phylogenetic relations among percid fishes as inferred from mitochondrial cytochrome *b* DNA sequence data. *Molecular Phylogenetics and Evolution* **10**, 343–353.

Sullivan T. J. & Stepien, C. A. (2014). Genetic diversity and divergence of yellow perch (*Perca flavescens*) spawning populations across the Huron–Erie Corridor, from Lake Huron through western Lake Erie. *Journal of Great Lakes Research* **40**, 101–109.

Sullivan, T. J. & Stepien, C. A. (2015). Temporal population genetic structure of yellow perch *Perca flavescens* spawning groups in the lower Great Lakes. *Transactions of the American Fishery Society*. **144**, 211–226.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**, 512–526.

Todd, T. N. & Hatcher, C. O. (1993). Genetic variability and glacial origins of yellow perch (*Perca flavescens*) in North America. *Canadian Journal of Fisheries and Aquatic Sciences* **50**, 1828–1834.

Vandewoestijne, S., Schtickzelle1, N. & Baguette, M. (2008). Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biology* **6**, 46. doi:10.1186/1741-7007-6-46.

Weir, B. S. & Cockerham, C. C. (1984). Estimating *F*–statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Electronic Reference

Goudet, J. (2002). *F*stat version 2.9.3.2. Available at: http://www2.unil.ch/popgen/softwares/fstat.htm. (accessed 10 February 2017).

Pritchard, J. K. & Wen, W. (2004). Documentation for STRUCTURE Software: Ver 2.3.3. Department of Human Genetics, University of Chicago, Chicago, Illinois. Available at: http://pritchardlab.stanford.edu/structure.html (accessed 10 February 2017).

858