CONTRASTING THE ROLE OF HISTORIC FACTORS IN PHYLOGEOGRAPIC PATTERNS IN THE NATIVE JOHNNY DARTER (*Etheostoma nigrum*) AND INVASIVE ROUND GOBY (*Neogobius melanostomus*) IN LOWER MICHIGAN

A.J. Wicks, M. Bowman, and T.E. Dowling

Department of Biological Sciences
Wayne State University

Correspondence
Abby Wicks
5047 Gullen Mall
Wayne State University
Detroit MI 48202
Email: abby.wicks@wayne.edu

ABSTRACT

Round goby (*Neogobius melanostomus*) is an invasive fish present in all five Great Lakes and is becoming increasingly common in their tributaries. Johnny darter (*Etheostoma nigrum*) is a native species that often coexists with *N. melanostomus*. In this work, historic factors are addressed as a source of genomic variation in study populations of these species. To do this, patterns of variation in the mitochondrial gene NADH dehydrogenase subunit 2 (ND2) were characterized for both species throughout Lower Michigan. Populations of *N. melanostomus* and *E. nigrum* were sampled from 17 localities representing both eastern and western basins of Lower Michigan to test the hypothesis that populations differ between the eastern and western basins of the Great Lakes. *Neogobius melanostomus* populations were largely homogenous with no significant differences detected among populations or between the eastern and western basins. Additionally, *N. melanostomus* exhibited no evidence of overarching historical genetic structure, consistent with the recent invasion and rapid expansion of this species. *Etheostoma nigrum* exhibited significant differentiation among local populations; however, similarity among mtDNA haplotypes indicated that differences among populations are recent, suggesting that local forces are a more important factor in shaping patterns of variation than historical factors. Contrary to predictions, there were no significant differences detected between the eastern and western basins of the Great Lakes; however, construction of a neighbor joining tree with F_{st} estimates revealed clustering of populations by basin with some anomalies. These anomalies may be the result of recent stream capture events facilitating gene flow between the two basins.

KEYWORDS biogeography, Great Lakes fishes, freshwater fish, genetic distance, molecular dating

INTRODUCTION

Evolutionary history, geologic history, and ecology of a species contribute to variable patterns of genetic diversity across the landscape (e.g., Avise 1992, Heithaus and Laushman 1997, Leclerc et al. 2008). Quantifying these patterns provides information on the structure and dynamics of populations in the context of landscape features and how historical factors can influence variation in the genome and gene expression. Depending on an organism's dispersal ability, landscape features may function to facilitate gene flow or create barriers that isolate populations. Isolated populations become genetically distinct over time, proceeding on independent evolutionary tracks due to mutation, drift, and differential selection. However, if connections exist that facilitate dispersal, populations can experience continuous gene flow, potentially eliminating differences among them (Slatkin 1987). Species vary in their ability to disperse and the extent to which landscape features facilitate or prevent gene flow depends on the ecology of a species (Hitt and Angermeier 2008). Taxa with more limited habitat preferences may be restricted to suitable areas (Tibbets and Dowling 1996) while a more plastic species may be able to disperse through variable habitats; therefore, patterns of genetic diversity are likely to reflect these ecological differences (Bohonak 1999).

The recent glacial history of the Great Lakes region has influenced population structure and the distribution of genetic diversity in its native fishes. The Laurentian Great Lakes were formed by glacial action during the Pleistocene, approximately 14,000 years ago, created by retreat of the Laurentide ice sheet (reviewed in Bailey and Smith 1981). The Lake Michigan and Lake Erie/Huron basins were formed due to the retreat of separate glacial lobes, each with their own periglacial lakes (Lakes Chicago and Maumee, respectively) and outflow systems. Thus, the eastern and western basins of the Great Lakes were potentially colonized by fishes from distinct

source populations that derived from separate glacial refugia (Bailey and Smith 1981, Lewis et al. 2008). Population structure of fishes in rivers of the two basins often reflects this geologic history as some species from these basins are distinct relative to each other (e.g., Gach 1996, Dowling et al. 1997, Stepien et al. 2009). Ecological factors influence observed patterns as well. Coldwater fishes may maintain populations near the glacial front where more recent connections could result in a more homogenous population colonizing the eastern and western basins of the post-glacial Great Lakes (Bailey and Smith 1981).

Here, population and phylogenetic approaches were used to characterize levels of genetic diversity and population structure within two common Great Lakes fishes, round goby (*Neogobius melanostomus)* and Johnny darter (*Etheostoma nigrum)*, to better understand factors that contribute to patterns of variation in these two species. Next generation sequencing technology has made it possible to study ecological interactions through quantification of patterns of gene expression (Ekblom and Galindo 2011), providing a perspective of how physiological processes can change in varying environments of natural systems. Comparing gene expression patterns could be informative in understanding the distribution of *N. melanostomus* in the Great Lakes region and impacts on interactions with native species like *E. nigrum*; however, knowledge of evolutionary history is important because differences in this history may have a significant influence on local patterns of gene expression. Given the recent ages of these populations, local evolutionary forces are more likely to play a significant role in gene expression patterns for round goby and Johnny darter than would the influence of genetic differences accumulated in older populations.

As a native species to the Great Lakes, *E. nigrum* populations may exhibit patterns that reflect their colonization history as glaciers retreated. As such, populations from the western basin of the Great Lakes may be significantly different from those of the eastern basin, illustrating how

historical processes can influence current patterns of genetic variation. Natural and anthropogenic barriers are also likely to influence structure in *E. nigrum* populations as. *E. nigrum* prefer shallow, slow-moving water; therefore, large waterways may function as barriers to gene flow and could result in differentiation across study rivers (Leidy 1992).

Unlike *E. nigrum*, *N. melanostomus* is a recent invader of the Great Lakes. It was first discovered in the St. Clair River in 1990 (Jude et al. 1992); therefore, this species is not influenced by local geological history. Population genetic studies using microsatellite and mitochondrial DNA markers determined that the Great Lakes were populated by individuals derived from a single source population from a tributary to the Black Sea (Brown and Stepien 2009), and *N. melanostomus* is now present in all five Great Lakes and in many tributaries (Kornis et al. 2012). Analyses of population structure indicate that genetic diversity of the source population is largely maintained throughout the Great Lakes; some limited structure has been identified and is primarily driven by differences in the populations of Saginaw Bay and Lake Ontario (Brown and Stepien 2009). In recently colonized Great Lakes tributaries, *N. melanostomus* populations have generally come to represent their local source populations likely due to consistent propagule pressure through migration and aided by human activity (Bronnehuber et al. 2011, Sard et al. 2019).

As a recent invader, *N. melanostomus* are expected to exhibit limited differences (e.g., numbers of haplotypes and mutations among them) within and among sampled populations. Additionally, it is expected that recently established riverine populations may exhibit only a subset of the variation found in the source lake populations (Brown and Stepien 2008). *Etheostoma nigrum* has resided in this region since glaciation (Bailey and Smith 1981), and given the complex glacial history of the region, there may be differences within and/or among regions (e.g., eastern vs western Great Lakes). As a native species, *E. nigrum* populations are more likely to have

diverged since their colonization as compared to *N. melanostomus*, thus more genetic differentiation is expected among populations of *E. nigrum* than *N. melanostomus*. Because of their propensity to inhabit shallow streams, large waterways are also likely to represent a barrier to gene flow for *E. nigrum* and could lead to divergence among populations. These hypotheses were tested by sequencing the mitochondrial DNA gene NADH dehydrogenase subunit 2 (ND2) for populations of *E. nigrum* and *N. melanostomus* throughout lower Michigan.

METHODS

Sixteen stream localities were selected representing the Lake Michigan, Lake Huron, and Lake Erie watersheds (Figure 1). Fishes were collected by seining, with up to 20 individuals of each species collected at each site. Acronyms used in tables and figures for sampling localities are: Au Sable (AS, 44.503294, -83.793609), Clinton (CL, 42.671619, -83.095602), Crockery Creek (CC, 43.053465, -86.062942), Dowagiac (DG, 42.011859, -85.962827), Kalamazoo (KA, 42.638600, -86.163315), Little Manistee (LM, 44.209024, -86.263904), Lower Rouge (LR, 42.285374, -83.388700), Muskegon (MU, 43.297940, -86.079321), Oqueoc (OQ, 45.456219, -84.087664), Pentwater (PW, 43.769223, -86.424267), Raisin (RN, 41.922478, -83.695259), Red Cedar (RC, 42.698207, -84.404845), Rifle (RF, 44.141451, -84.043657), St. Joseph (SJ, 42.074666, -86.461342), Shiawasee (SH, 42.919861, -83.969519), Stony Creek (SC, 42.023489, -83.419425). In some localities where one or both species was rare, multiple nearby localities and/or sample dates were combined to achieve desired sample size. As both species were not present at every locality, the Dowagiac, Raisin, and Shiawasee include only *E. nigrum*. Crockery Creek, Pentwater, and St. Joseph include only *N. melanostomus*. Individuals were fin clipped, and tissue was stored in 95% ethanol. For DNA isolation, tissue samples were dissolved in a solution of proteinase K and

sodium dodecyl sulfate and purified via one of two methods: phenol/chloroform extraction method or magnetic bead DNA purification (samples collected in 2015-2016 and 2017-2019, respectively). Phenol/chloroform extraction followed Tibbets and Dowling (1996). Magnetic bead purifications followed the manufacturer's (Axygen) protocol with the following changes: 25 ul of lysate was added to 25 ul of AxyPrep, 70% ethanol was used for two washes, and DNA was eluted in 40 ul of water. DNA was assessed for quality and quantity with a Nanodrop Spectrophotometer (ThermoFisher).

The mitochondrial gene NADH dehydrogenase subunit 2 (ND2) was selected because of its relatively high rate of mutation and strict maternal inheritance, allowing for better characterization of more recent events. Sequences were amplified using the primers B2Gila (5' CTCTTAGTGCTTCCTCACA 3') and ASN (5' CGCGTTTAGCTGTTAACTAA 3') for *E. nigrum* and RGND2F (5' AGCATGCCGGTTAAAATCC 3') and RGND2R (5' GGATCCGAGGCCTTCCTGTCT 3') for *N. melanostomus*. The following PCR conditions were used for both species: 94°C for 15 min, 25 cycles of denaturation at 94°C for 1 min, annealing at 58°C for 1 min, and polymerization at 72°C for 2 min with a final annealing step of 72°C for 10 min, and holding at 4°C until long term storage at -20°C. Amplification products were checked for quantity and quality using agarose gel electrophoresis prior to being sent to the Wayne State University's Applied Genomics Technology Center or Eton Bioscience for 2015-2018 and 2019 samples, respectively. Products were sequenced with the same primers using Applied Biosystems DNA Analyzer 3730 sequencer, yielding total of 1047 bp (the entire ND2 gene) for *E. nigrum*. Due to low quality reads at the beginning of the gene for some samples, *N. melanostomus* sequences were trimmed to 1029 bp.

Some individuals included in this study were also included in a separate RNA-seq study (Wicks 2019). For these individuals, ND2 gene sequences were obtained from RNA-seq reads. From the assembly and aligned read files, the SAMtools package (Li 2011) was used to call SNPs and output as a consensus sequence for each individual. Variants were quality filtered for phred score >30 and individuals with ambiguous variant calls were excluded from the analysis.

Sequences were aligned, assembled, and trimmed in Bioedit (Version 7.0.5.3) (Hall et al. 2011). MEGA (Version 5.2.2) (Kumar et al. 2016) was used to produce a maximum likelihood tree and to assign haplotypes. Sequences and haplotype counts for each location were used in ARLEQUIN (version 3.5.1.2) (Excoffier and Lischer 2010) to obtain estimates of genetic diversity (e.g., number of haplotypes, gene diversity) within populations. Variation among populations was assessed and tested for significance using a molecular analysis of variance (AMOVA), also in ARLEQUIN. This approach was used to partition levels of genetic variation within and among river drainages and regions (e.g., eastern vs western basins of the Great Lakes). Haplotype networks were created as median-joining networks with PopArt (Version 1.7) (Leigh and Bryant 2015). Neighbor joining trees of populations were generated with pairwise F_{ST} values in MEGA. For *E. nigrum*, additional ND2 sequences were obtained from NCBI GenBank and included in the maximum likelihood tree of haplotypes, also generated with MEGA with 1000 bootstrap replicates. These included closely related species *E. olmstedi* (EF027210), *E. podostemone* (JQ088571), *E. perlongum* (JQ088568), *E. susanae* (JQ088589), *E. vitreum* (FJ381264), and *E. longimanum* (JQ088552) and co-occurring darter species *E. blennioides* (JQ088546), *E. exile* (EF027194), *and E. caereleum* (JQ088546). Additionally, one *E. nigrum* ND2 sequence

(JQ088561) from an individual collected from the Embarras River, a tributary of the Wabash River (N. Lang, pers. comm.) was obtained from GenBank and included in the analysis.

To provide context for the large divergence among some *E. nigrum* haplotypes, the program *BEAST (found within BEAST 2 v2.7.3, Bouckaert et al. 2014) was used to estimate divergence dates among *E. nigrum* haplotypes as well as other *Etheostoma* species included in the maximum likelihood tree. *Etheostoma* species are not well represented in the fossil record (Bailey and Smith 1981); therefore, previous estimates of molecular divergence in darters that have used Centrarchid fossil records for calibration for divergence dating were used here (e.g., Near et al. 2011, Bossu et al. 2013, Fluker et al. 2014). The substitution rate was estimated by Near et al. (2011) from the mitochondrial cytochrome b (cyt*b*) gene as 8.99×10^{-3}, and this has been applied in recent estimates of molecular divergence in *Etheostoma* species (Echelle et al. 2015, McCall and Fluker 2021, MacGuigan et al. 2023). Although this rate may result in an underestimate of divergence time for ND2 as ND2 evolves at a faster rate than cyt*b* (Mueller 2006), it is the best available rate for estimating divergence time from mitochondrial genes in *Etheostoma* species and was therefore used as the substitution rate in this analysis. *BEAST was run using default options except for the use of a species tree relaxed clock, HKY substitution model, and Yule model as a tree prior. Within the BEAST package, TreeAnnotator was used to generate a maximum clade credibility tree using median node heights and the tree was visualized with FigTree.

RESULTS

Neogobius melanostomus

A total of 247 *N. melanostomus* samples were collected from 12 localities and sequenced for ND2. Five haplotypes were identified (Table 1, Figure 2). Each was separated by one substitution from its closest neighbor except for haplotype C, which differed by two changes. Four variants resulted from changes to third codon positions (Haplotypes A-D) and did not result in amino acid changes. One variant resulted from a mutation in a first codon position and resulted in an amino acid change (Haplotype E).

Haplotype A was found in 96% of all individuals sampled, resulting in very low levels of gene and nucleotide diversity within samples (Table 2, Figure 3). AMOVA was used to quantify levels of genetic variance within and among samples (F_{ST}). The distribution of variation between the lower Great Lakes was assessed by further partitioning variance between tributaries flowing into eastern (Lakes Huron, St. Clair, and Erie) and western lakes (Lake Michigan) (F_{CT}) and among samples within these two groups of drainages (F_{SC}). Statistical assessment failed to identify significant differences among samples (F_{ST} = 0.033, P = 0.106), between the two drainages (F_{CT}= 0.017, P = 0.055), or among samples within these two drainages (F_{SC} = 0.016, P = 0.259). This lack of differentiation was also supported by examination of pairwise estimates of F_{ST} as only three of the comparisons were significant (Table 3). Similarity among samples was examined by clustering samples by F_{ST} using the neighbor-joining method (Figure 4). There was no overarching genotypic structure of *N. melanostomus* populations as samples from different lake basins were intermingled.

198 *Etheostoma nigrum*

199 A total of 207 *E. nigrum* individuals were collected from 13 localities and sequenced for
200 ND2. Twenty-eight haplotypes were identified (Table 2, Figure 5). There were 53 polymorphic
201 sites with 13 of those variants in the first codon position and 40 in the third codon position. Of
202 these, 10 changes resulted in amino acid changes. Most of variants resulted from single base pair
203 changes from the most common haplotype, A. Two haplotypes, AA and AB, differed from
204 haplotype A by 28 and 27 base pair changes, respectively.

205 A maximum likelihood tree was generated to examine haplotypic variation in *E. nigrum*
206 from the lower peninsula of Michigan in phylogenetic context relative to other samples obtained
207 from GenBank (Figure 6). Most haplotypes sampled cluster closely to the most common
208 haplotype, A, and formed a well-supported monophyletic lineage (98% bootstrap value) with
209 limited phylogenetic structure among haplotypes within this group. Haplotypes D, U, V, W, and
210 X form a lineage in the minimum spanning network (Figure 5) and form a monophyletic group in
211 the maximum likelihood tree (Figure 6). This group exhibited a moderately high bootstrap value
212 in the ML analysis (74%) and is found in samples from three separate drainages from both eastern
213 and western Great Lakes (Stony Creek [Erie], Ocqueoc [Huron] and Kalamazoo [Michigan] rivers,
214 Table 2). Another distinct haplotype lineage Y-AA occurred only in the Oqueoc and Little
215 Manistee rivers (Lakes Huron and Michigan drainages, respectively), and this lineage clusters
216 closely to the unresolved large group of *E. nigrum* haplotypes. *Etheostoma podostemone* was the
217 sister taxon to most haplotypes from the Great Lakes sampled here (98% bootstrap value). The
218 two divergent haplotypes AB and AC from Stony Creek (Lake Erie drainage) share their most
219 recent ancestry with *E. nigrum* from the Embarras River (Wabash River drainage) in central
220 Illinois (95% bootstrap value).

Levels of variation within and among samples were used to assess geographic population structure in *E. nigrum* (Figure 7). Haplotype A was the most common, found in more than 40% of all individuals sampled and at every locality except the Kalamazoo River. All other haplotypes were more localized, occurring at a maximum of two localities. Private alleles were also common, with 22 alleles occurring at only single localities; however, every locality harbored more than one haplotype.

Levels of sequence diversity were highly variable among samples (Table 4), with gene diversity and number of pairwise differences ranging from 0.13 – 0.76 and 0.13-9.66 respectively. Exceptional levels of gene and nucleotide diversity were noted in Stony Creek and Little Manistee samples due to the presence of divergent haplotypes discussed above. Stony Creek samples exhibited five haplotypes with a mean number of pairwise differences of 9.67 relative to the other Great Lakes samples, and the Little Manistee River contained five haplotypes with a mean number of pairwise differences of 5.05.

AMOVA was used to partition genetic variance into within and among sample components, identifying significant differences among samples (F_{ST}= 0.457, P < 0.0001, 54.3% of the variation). Further subdivision to assess levels of variation within and among two sets of drainages (Lake Michigan vs Lakes Huron, St Clair, and Erie) indicated that variation was largely attributable to differences among samples within those two drainage groups (F_{SC} = 0.435, P < 0.0001, 41.9% of the variation), not differences among them (F_{CT}= 0.038, P = 0.19, 3.8% of the variation). The lack of differentiation among the drainage groups appears to be driven by a shared haplotype (U) between Stony Creek and the Kalamazoo River. If Stony Creek is removed from the analysis, differences among drainage groups becomes significant (F_{SC} = 0.565, P < 0.0001, F_{CT}= 0.100, P <0.023). However, differences among samples within the groups and within

244 populations explain a much larger proportion of the variation (51.0% and 39.2% respectively) than
245 differences among groups (9.9%)

246 Pairwise estimates of F_{ST} (Table 6) were used to construct a neighbor-joining tree for *E.*
247 *nigrum* population samples, revealing similarities among the locations within the major basins
248 (Figure 8). Samples from Lake Huron River drainages were like each other while those of the other
249 basins exhibited more divergence among populations. Samples from the Muskegon River
250 population, a Lake Michigan drainage, and the Raisin River, a Lake Erie drainage, clustered with
251 samples from the Lake Huron basin instead of more geographically proximate samples because of
252 the high frequency of haplotype A (Table 4, Figure 7). Drainages of Lake St. Clair and Lake Erie
253 were generally intermediate to those from Lakes Huron and Michigan; however, samples from
254 these basins and Lake Michigan have long terminal branches reflecting the high frequency of
255 private alleles at these locations.

256 Estimates of molecular divergence were calculated using *BEAST to better understand the
257 context of highly divergent *E. nigrum* haplotypes. The resulting tree generated with TreeAnnotator
258 and visualized with FigTree is shown in Figure 9. *Etheostoma nigrum* haplotypes from these
259 populations show two major mtDNA lineages (haplotypes A – AA and AB – AC) which diverged
260 approximately 1.8 mya (95% HPD 1.2-2.6 Ma). There are two groups in the main lineage (A – X,
261 Y – AA) which diverged approximately 500,000 years ago (95% HPD 0.25-0.83 Ma). Within this
262 main lineage, divergence between haplotypes range from 70,000 to almost 200,000 years ago. The
263 sister group contains haplotypes Y and AA which are about 90,000 years diverged. The divergent
264 lineage that contains haplotypes AB and AC shares its most recent common ancestry with the *E.*
265 *nigrum* individual from the Embarras river, with estimated divergence time of approximately
266 600,000 years ago.

DISCUSSION

Patterns of genetic variation in *N. melanostomus* and *E. nigrum* were explored using a variable mitochondrial DNA gene to assess the levels of divergence within and among populations. It was predicted that, as a recently introduced species, *N. melanostomus* would show limited divergence among populations and low genetic diversity. As a native species with more limited dispersal, greater diversity within and more divergence among *E. nigrum* populations was predicted. Results were consistent with these expectations. It was also predicted that *E. nigrum* populations of east and west basins of the Great Lakes would be significantly different due to distinct colonization sources, but those from eastern and western lake basins were not significantly different. Despite this result, the distribution of genetic diversity provided insight into the geologic and ecological factors that may have shaped the population structure in *E. nigrum*.

Neogobius melanostomus

ND2 sequences from *N. melanostomus* showed substantially less diversity than *E. nigrum*. Invasion of the Great Lakes by *N. melanostomus* is recent, arriving in ballast water from the Black Sea circa 1990. Previous work has found that, as it expanded its range within the Great Lakes, *N. melanostomus* has retained levels of genetic variation on par with source populations (Brown and Stepien 2009). Population genetic studies of *N. melanostomus* in more recently established stream populations show that *N. melanostomus* is expanding its range without founder effects (Bronnenhuber et al. 2011). Bronnenhuber et al. (2011) also found *N. melanostomus* populations lacked genetic structure. Consistent with this result, there were few significant differences among populations of *N. melanostomus* and no significant geographic structure. Brown and Stepien (2008) found limited structure among *N. melanostomus* populations in the Great Lakes, with this

result primarily driven by differences between populations (Lake Ontario and Saginaw Bay) which were not included in this study. Additional statistically significant structure in these previous studies may also have been driven by low-frequency private alleles in lake populations which may have been missed here due to the small number of samples and their sizes, or that these rare alleles may not yet be present in these more recently established stream populations. The small number of ND2 haplotypes and dominance of a single haplotype identified in *N. melanostomus* is consistent with other previous work examining levels of mtDNA variation in *N. melanostomus* (Stepien and Tumeo 2006, Brown and Stepien 2008).

Neogobius melanostomus has dispersed rapidly throughout the Great Lakes (Kornis et al. 2012), resulting in high levels of gene flow, and the homogeneity and low diversity of samples around the Great Lakes is consistent with expectations. Established populations of *N. melanostomus* are less prone to dispersal and favor residents (Thorlacius et al. 2015); however, even very low levels of migration can provide sufficient gene flow to prevent population divergence at presumably neutral loci like these (Slatkin 1985, Newman and Tallmon 2001). As a high dispersing species not restricted by habitat, migration will likely be a persistent, homogenizing force in *N. melanostomus*. This situation has been aided by human activity as secondary spread of *N. melanostomus* to Michigan's inland waterways has occurred due to the accidental movement of round gobies in live bait used by anglers (Sard et al. 2019). Populations found where barriers such as dams prevent natural migration will likely be distinct from connected populations due to founder effects and genetic drift. However, where migration and human activity are a constant source of gene flow, the lack of structure observed here will likely persist.

313 *Etheostoma nigrum*

314 As a native species, the distribution of *E. nigrum* has been shaped, in part, by the action of
315 glaciers. Ancestors of this species likely colonized the early Great Lakes region more than 10,000
316 years ago as glaciers receded, and the modern lake basins were formed (Bailey and Smith 1981).
317 In addition, *E. nigrum* tends to be habitat restricted, occupying shallow, slow-moving stretches of
318 streams and rivers. Such preferences have been shown to make fishes less likely to move among
319 rivers through downstream dispersal and more likely to diverge (Tibbets and Dowling 1996).
320 *Etheostoma nigrum* also have relatively low fecundity with males guarding nests, characteristics
321 which may limit gene flow (Turner and Trexler 1998, Stepien et al. 2007). This is reflected in
322 results of the AMOVA, where significant differences were mainly driven by variation among
323 individual river drainages within lakes basins, but not between east and west basins as expected,
324 even if the population with the divergent haplotypes (Stony Creek) as removed from the analysis.
325 Westbrook (2012) identified a similar pattern in Wisconsin drainages of the Great Lakes, in which
326 geographic isolation resulted in limited gene flow and significant divergence among *E. nigrum*
327 populations in different drainages.

328 For species limited in their ability to disperse by suitable habitat, isolation of individual
329 populations can lead to divergence. Dowling et al. (2015) highlighted the importance of local
330 forces in shaping population structure among populations of three species of roundtail chub (*Gila*
331 *intermedia*, *G. nigra*, and *G. robusta*), noting high levels of population level divergence that could
332 obscure broader historical factors. It is possible that the east and west basins of the Great Lakes
333 were colonized by distinct populations of *E. nigrum*, but over time evolutionary processes may
334 have driven divergence among local populations such that differences among broader hierarchical
335 categories cannot be detected (Hedrick 1999). Similar patterns have been observed for other

336 species including smallmouth bass (*Micropterus dolomieu*, Stepien et al. 2017), white sucker
337 (*Catostomus commersoni*, Lafontaine and Dodson 1997), and mottled sculpin (*Cottus bairdii,*
338 Homola et al. 2016).

339 Despite isolation due to limited dispersal, more recent shifts in hydrology and geologic
340 features in the region may provide paths of dispersal that facilitate gene flow. These shifts may
341 explain some of the patterns observed in *E. nigrum*. Ray et al. (2006) examined mtDNA variation
342 among rainbow darter (*Etheostoma caeruleum*) populations (an ecologically similar species to *E.*
343 *nigrum*) and found that sampled Wabash River populations grouped with samples from both Lakes
344 Michigan and Erie. Stream capture of parts of the Wabash River (a tributary of the Ohio River) by
345 the Maumee River (Figure 1) may have facilitated gene flow between the two basins, as evidenced
346 by similarity of haplotypes with Stony Creek (Lake Erie drainage) and the Wabash River drainage
347 (Figure 6). *Etheostoma nigrum* populations appear to have followed a similar colonization pattern
348 in which colonization of Lake Erie occurred via the captured portion of the Wabash River and
349 propagated throughout the Lake Erie watershed, reducing structure based on east/west basins. The
350 molecular dating analysis placed the divergence of this Wabash River-derived lineage during the
351 Pleistocene, approximately 1.8 mya.

352 Construction of a neighbor-joining tree using *E. nigrum* population samples also revealed
353 similarities among samples within major basins, especially Lake Huron; however, one Lake
354 Michigan drainage (Muskegon River) grouped with Lake Huron drainages (Figure 8). This
355 grouping appears to be the result of the higher frequency of the "A" haplotype in the Muskegon
356 River, which is rare in the other drainages of the western basin and common in the eastern basin
357 drainages. Given the close proximity of the headwaters of the Muskegon River to the headwaters
358 of several Lake Huron drainages (Figure 1) and the low, marshy topography of the region, it is

possible that connections historically existed and allowed gene flow between basins in this region. During high water events, connections may temporarily exist in the present day.

Despite their recent origin, there is a high level of ND2 diversity within and among these *E. nigrum* populations (F_{ST}= 0.406, P < 0.0001 and mean pairwise differences ranging from 0.125-9.66). Based on *BEAST analysis, emergence of the most recent haplotypes predates formation of the modern Great Lakes (70,000 to 200,000 years ago), indicating that some haplotypes may have been found in different refugia. Additionally, two highly divergent haplotypes grouped closely with geographically distant populations. This would suggest that much of this diversity was introduced from populations maintained along the glacial front, rather than evolving *in situ*. Given that only one *E. nigrum* ND2 haplotype was available in Genbank and already included in the analysis, further exploration of source populations using ND2 was not available at this time. However, many cyt*b* sequences are available and a sample of these from around the Midwest (MacGuigan et al. 2023), along with a subset of cyt*b* sequences from the Rouge and Clinton River individuals were used in a maximum likelihood tree which yields a large clade with minimal structure. Individuals from these Rouge and Clinton haplotypes group with geographically diverse locations including the Ohio River, Wabash River, Kentucky River, and Wisconsin River, including a shared haplotype between the Rouge, Ohio, Wabash, and Embarrass Rivers (data not shown). The minimal structure in cyt*b* haplotypes supports the conclusion from the ND2 analysis that diversity was contributed by multiple source populations along the glacial front.

Additional unexpected patterns were noted in the maximum-likelihood tree of *E. nigrum* haplotypes and related taxa (Figure 6). The clustering of *E. podostemone* within *E. nigrum* haplotypes is inconsistent with current darter phylogenies. Darters are a highly speciose group and their taxonomy is not fully resolved (Near et al. 2011). MacGuigan and Near (2018) characterized

382 the group which includes *E. nigrum* using Next Generation sequencing data, yielding two major
383 lineages, one containing *E. nigrum, E. olmstedi*, and *E. perlongum*, and the other *E. podostemone*
384 and *E. longimanum* (Figure 6B). Results from this analysis (Figure 6A) are not generally consistent
385 with this result as the two lineages identified here include 1) *E. nigrum*, *E. perlongum*, and *E.*
386 *podostemone* and 2) *E. olmstedi* and *E. longimanum*. Differences between these studies may
387 reflect differences mitochondrial inheritance patterns and/or introgression among species
388 compared to nuclear genes. Understanding reasons behind the discordance between mtDNA and
389 nuclear genes requires further study.
390 The position of *E. podostemone* is especially interesting as it is found clustered within a
391 group of *E. nigrum* haplotypes. *Etheostoma podostemone* is localized to a small and isolated region
392 in Virginia and North Carolina, distant from the localities sampled here. Heckmann et al. (2009),
393 using nuclear genes and the mitochondrial cyt*b* gene, found a similar pattern of nesting of *E.*
394 *podostemone* among *E. nigrum* haplotypes from the Mississippi River, Mobile River, and the Great
395 Lakes. MacGuigan et al. (2023) also identified the nesting of *E. podostemone* within *E. nigrum*
396 with mtDNA, while the phylogeny generated from restriction site associated DNA sequencing
397 (RADseq) grouped *E. podostemone* most closely with *E. longimanum*, forming a sister clade to *E.*
398 *nigrum.* This pattern may reflect ancient mitochondrial introgression, a process that has often been
399 identified as an important mechanism in the evolution of *Etheostoma* species (Ray et al. 2008,
400 Bossu and Near 2009, MacGuigan and Near 2018).
401
402 *Conclusions*
403 It was predicted that, as a recently introduced species, *N. melanostomus* populations would
404 show low diversity and limited geographic structure. Results were consistent with this hypothesis.

405 *Neogobius melanostomus* populations were dominated by a single mtDNA haplotype and few
406 statistically significant differences among locations as detected using F-statistic analysis. This
407 result reflects the history of round goby in the Great Lakes, having been recently founded from a
408 single source population, followed by rapid dispersal through the region.

409 This is the first study to examine population genetics of *E. nigrum* in the Great Lakes
410 region, with population genetic structure of *E. nigrum* in Lower Michigan offering insight into
411 the historic processes that shaped the geographic distribution of the species. Native *E. nigrum* is
412 more likely to exhibit geographic differences among and higher levels of genetic diversity within
413 populations than *N. melanostomus*, reflecting its past geological and evolutionary history. This
414 study identified differences among populations within basins as a significant source of variation
415 with limited divergence among regions. Contrary to predictions based on findings of previous
416 studies (Gach 1996, Dowling et al. 1997, Stepien et al. 2009), there were not significant
417 differences between the eastern and western basins of the Great Lakes; however, multiple
418 mtDNA lineages found within *E. nigrum* populations suggests that multiple source populations
419 colonized the region as the glaciers receded, and historic routes of gene exchange among basins
420 may have reduced the level of genetic differences between them. Given this pattern was
421 consistent with that of an ecologically similar species (*E. caeruleum*) it may follow for other
422 similar species as well. Future research could expand to a broader geographical range with
423 markers, allowing for greater historical perspective. This also informs future work comparing
424 gene expression differences in *E. nigrum* and *N. melanostomus* populations, allowing
425 evolutionary history to be considered as a source of variation in patterns of gene expression
426 within and between these species.
427

AUTHOR CONTRIBUTIONS

AJW and **TED** conceptualized and designed the work; **AJW**, **MB**, and **TED** contributed to sample collection; **AJW** and **MB** contributed to lab work and data processing; **AJW** completed data analysis; **AJW**, **MB**, and **TED** contributed to drafting, editing, and approving the manuscript.

ACKNOWLEDGEMENTS

Funding for this work was provided by Wayne State University to TED. Fieldwork and collections were undertaken with permission from the Michigan Department of Natural Resources, under scientific collector permits issued to T.E. Dowling and C.A. Krabbenhoft. Collection and handling of specimens were in accordance with Wayne State University Institutional Animal Care and Use Committee (IACUC) protocol 18-02-0553. We greatly appreciate the contribution of C.A. Krabbenhoft, R. Roose, K. Kargol, K. Pollard, B. Dodds, and J. Nitz to field sampling. We also thank Friends of the Rouge, in particular S. Patrella, R. Muller, and P. Kukulski for assistance in site selection and field sampling.

CONFLICT OF INTEREST

The authors declare no competing interest.

DATA ACCESSABILITY STATEMENT

All ND2 haplotype sequences for *E. nigrum* and *N. melanostomus* have been submitted to GenBank under accession numbers OR777617-OR777644 and OR795530-OR795534, respectively.

REFERENCES

Avise, J.C. (1992). Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* 62-76.

Bailey, R.M. and Smith, G.R. (1981). Origin and geography of the fish fauna of the Laurentian Great Lakes basin. *Canadian Journal of Fisheries and Aquatic Sciences 38(12)*: 1539-1561.

Bossu, C.M. and Near, T.J. (2009). Gene trees reveal repeated instances of mitochondrial DNA introgression in orangethroat darters (Percidae*: Etheostoma*). *Systematic biology 58(1)*: 114-129.

Bronnenhuber, J.E., Dufour, B.A., Higgs, D.M. and Heath, D.D. (2011). Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round goby, *Neogobius melanostomus*, in Great Lakes tributaries. *Molecular Ecology 20(9)*: 1845-1859.

Brown, J.E. and Stepien, C.A. (2008). Ancient divisions, recent expansions: phylogeography and population genetics of the round goby *Apollonia melanostoma*. *Molecular Ecology 17(11)*: 2598-2615.

Brown, J.E. and Stepien, C.A. (2009). Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. *Molecular Ecology 18(1)*: 64-79.

Bohonak, A.J. (1999). Dispersal, gene flow, and population structure. *The Quarterly Review of Biology 74(1)*: 21-45.

Bossu, C.M., Beaulieu, J.M., Ceas, P.A. and Near, T.J. (2013). Explicit tests of paleodrainage connections of southeastern North America and the historical biogeography of Orangethroat Darters (Percidae: *Etheostoma*: C). *Molecular Ecology 22(21)*: 5397-5417.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. and Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10(4) p.e1003537.

Dowling T.E., Anderson C.D., Marsh P.C., and Rosenberg M.S. (2015). Population Structure in the Roundtail Chub (*Gila robusta* Complex) of the Gila River Basin as Determined by Microsatellites: Evolutionary and Conservation Implications. *PLoS ONE* 10(10): e0139832.

Dowling, T.E., Broughton, R.E. and DeMarais, B.D. (1997). Significant role for historical effects in the evolution of reproductive isolation: evidence from patterns of introgression between the cyprinid fishes, *Luxilus cornutus* and *Luxilus chrysocephalus*. *Evolution 51(5)*:1574-1583.

Echelle, A.A., Schwemm, M.R., Lang, N.J., Baker, J.S., Wood, R.M., Near, T.J. and Fisher, W.L. (2015). Molecular systematics of the Least Darter (Percidae: *Etheostoma*

microperca): historical biogeography and conservation implications. *Copeia 103(1)*: 87-98.

Excoffier, L. and Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources 10(3)*: 564-567.

Fluker, B.L., Kuhajda, B.R. and Harris, P.M. (2014). The influence of life-history strategy on genetic differentiation and lineage divergence in darters (Percidae: Etheostomatinae). *Evolution 68(11)*: 3199-3216.

Gach, M. (1996). Geographic variation in mitochondrial DNA and biogeography of *Culaea inconstans* (Gasterosteidae). *Copeia* 3: 563-575.

Hall, T., Biosciences, I. and Carlsbad, C. (2011). BioEdit: an important software for molecular biology. *GERF Bull Biosci 2(1)*: 60-61.

Heckman, K.L., Near, T.J. and Alonzo, S.H. (2009). Phylogenetic relationships among *Boleosoma* darter species (Percidae: *Etheostoma*). *Molecular Phylogenetics and Evolution 53(1)*: 249-257.

Heithaus, M.R. and Laushman, R.H. (1997). Genetic variation and conservation of stream fishes: influence of ecology, life history, and water quality. *Canadian Journal of Fisheries and Aquatic Sciences 54(8)*: 1822-1836.

Hitt, N.P. and Angermeier, P.L. (2008). Evidence for fish dispersal from spatial analysis of stream network topology. *Journal of the North American Benthological Society 27(2)*: 304-320.

Homola, J.J., Ruetz III, C.R., Kohler, S.L. and Thum, R.A. (2016). Complex postglacial recolonization inferred from population genetic structure of mottled sculpin *Cottus*

511 *bairdii* in tributaries of eastern Lake Michigan, USA. *Journal of Fish Biology 89(5)*:
512 2234-2250.
513 Jude, D.J., Reider, R.H. and Smith, G.R. (1992). Establishment of Gobiidae in the Great Lakes
514 basin. *Canadian Journal of Fisheries and Aquatic Sciences 49(2)*: 416-421.
515 Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018). MEGA X: molecular
516 evolutionary genetics analysis across computing platforms. *Molecular Biology and*
517 *Evolution 35(6)*: 1547-1549.
518 Kornis, M.S., N. Mercado-Silva, and M.J. VanderZanden. (2012). Twenty years of invasion: a
519 review of round goby *Neogobius melanostomus* biology, spread and ecological
520 implications. *Journal of Fish Biology 80(2)*: 235-285
521 Lafontaine, P. and Dodson, J.J. (1997). Intraspecific genetic structure of white sucker
522 (*Catostomus commersoni*) in northeastern North America as revealed by mitochondrial
523 DNA polymorphism. *Canadian Journal of Fisheries and Aquatic Sciences 54(3)*: 555-
524 565.
525 Leclerc, E., Mailhot, Y., Mingelbier, M. and Bernatchez, L. (2008). The landscape genetics of
526 yellow perch (Perca flavescens) in a large fluvial ecosystem. *Molecular ecology 17*(7):
527 1702-1717.
528 Leidy, R.A. (1992). Microhabitat selection by the Johnny darter, *Etheostoma nigrum* Rafinesque,
529 in a Wyoming stream. The Great Basin Naturalist: 68-74.
530 Leigh, J.W. and Bryant, D. (2015). Popart: full-feature software for haplotype network
531 construction. *Methods in Ecology and Evolution 6(9)*: 1110-1116.

Lewis, C.F.M., Karrow, P.F., Blasco, S.M., McCarthy, F.M., King, J.W., Moore Jr, T.C. and Rea, D.K. (2008). Evolution of lakes in the Huron basin: deglaciation to present. *Aquatic Ecosystem Health & Management 11(2)*: 27-136.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics: 27(21)*: 2987-93.

MacGuigan, D.J. and Near, T.J. (2018). Phylogenomic Signatures of Ancient Introgression in a Rogue Lineage of Darters (Teleostei: Percidae*)*. *Systematic Biology 68(2)*: 329-346.

MacGuigan, D.J., Orr, O.D. and Near, T.J. (2023). Phylogeography, hybridization, and species discovery in the Etheostoma nigrum complex (Percidae: Etheostoma: Boleosoma). *Molecular Phylogenetics and Evolution 178*: 107645.

McCall, B.L. and Fluker, B.L. (2021). Conservation genetic assessment of the paleback darter, Etheostoma pallididorsum, a narrowly distributed endemic in the Ouachita Highlands, Arkansas, USA. Aquatic Conservation: *Marine and Freshwater Ecosystems 31(7)*: 1817-1830.

Mueller, R.L. (2006). Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology* 55(2): 289-300.

Near T.J., Bossu C.M., Bradburd G.S., Carlson R.L., Harrington R.C., Hollingsworth P.R., Keck B.P., and Etnier D.A. (2011). Phylogeny and temporal diversification of darters (Percidae: Etheostomatinae). *Systematic Biology 60(5)*: 565–595.

Newman, D. and Tallmon, D.A. (2001). Experimental evidence for beneficial fitness effects of gene flow in recently isolated populations. *Conservation Biology 15(4)*: 1054-1063.

555 Ray, J.M., Wood, R.M. and Simons, A.M. (2006). Phylogeography and post-glacial colonization
556 patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). *Journal of*
557 *Biogeography 33(9)*: 1550-1558.
558 Sard, N., Robinson, J., Kanefsky, J., Herbst, S. and Scribner, K. (2019). Coalescent models
559 characterize sources and demographic history of recent round goby colonization of Great
560 Lakes and inland waters. *Evolutionary Applications 12*(5): 1034-1049.
561 Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*
562 *236(4803)*: 787-792.
563 Slarkin, M. (1985). Gene flow in natural populations. Annual review of ecology and systematics
564 16(1): 393-430.
565 Stepien, C.A., Murphy, D.J., Lohner, R.N., Sepulveda-Villet, O.J. and Haponski, A.E. (2009).
566 Signatures of vicariance, postglacial dispersal and spawning philopatry: population
567 genetics of the walleye *Sander vitreus*. *Molecular Ecology 18(16)*: 3411-3428.
568 Stepien, CA, Karsiotis, S.I. Sullivan, T.J. and Klymus, K.E. 2017. Population genetic structure
569 and comparative diversity of smallmouth bass *Micropterus dolomieu*: congruent patterns
570 from two genomes. *Journal of Fish Biology 90(5)*: 2125-2147.
571 Stepien, C. A., D. J. Murphy, and R. M. Strange. 2007. Broad-to fine-scale population genetic
572 patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great
573 Lakes and beyond: an interplay of behaviour and geography. *Molecular Ecology 16(8)*:
574 1605–1624.
575 Stepien, C.A. and Tumeo, M.A. 2006. Invasion genetics of Ponto-Caspian gobies in the Great
576 Lakes: a 'cryptic'species, absence of founder effects, and comparative risk analysis.
577 *Biological Invasions 8*: 61-78.

578 Thorlacius, M., Hellström, G. and Brodin, T. 2015. Behavioral dependent dispersal in the
579 invasive round goby *Neogobius melanostomus* depends on population age. *Current*
580 *Zoology 61(3)*: 529-542.
581 Tibbets, C. A., & Dowling, T. E. 1996. Effects of intrinsic and extrinsic factors on population
582 fragmentation in three North American minnows (Teleostei: Cyprinidae). *Evolution*
583 *50(3)*: 1280-1292.
584 Turner, T.F. and Trexler, J.C. 1998. Ecological and historical associations of gene flow in darters
585 (Teleostei: Percidae). *Evolution 52(6)*: 1781-1801.
586 Westbrook, L.J. (2012). Genetic Structure of Rock Bass and Johnny darter: Implications for
587 Gamefish Management in Wisconsin. Master's Thesis, University of Wisconsin-Stevens
588 Point, College of Natural Resources.

589
590
591

TABLES AND FIGURES

Table 1. Sample size (N) and haplotype counts (identified as letters) for samples of *N. melanostomus.* Complete locality data is provided in the Methods section.

			Haplotype				
Major Drainage	**Locality**	**N**	**A**	**B**	**C**	**D**	**E**
Lake Erie	Stony Creek	30	30				
	Lower Rouge	30	30				
Lake St. Clair	Clinton	19	19				
Lake Huron	Rifle	4	4				
	Au Sable	20	19	1			
	Oqueoc	40	39		1		
Lake Michigan	Manistee Lake	5	5				
	Pentwater Lake	25	24		1		
	Muskegon	20	17				3
	Crockery Creek	20	19		1		
	Kalamazoo	14	14				
	St Joseph	20	17		2	1	
Total		247	237	1	5	1	3

Table 2. Sample size (N) and haplotype counts (identified as letters) for samples of *E. nigrum*. Complete locality data is provided in the methods section.

			Haplotype																											
Major Basin	**Locality**	**N**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**AA**	**AB**	**AC**
Lake Erie	Stony Creek	17	6									2											6						1	2
	Raisin	13	10						2			1																		
	Lower Rouge	21	5																			16								
Lake St. Clair	Clinton	25	1											7						17										
Lake Huron	Shiawasee	14	10				2	2																						
	Rifle	23	17					2		2					2															
	Au Sable	13	11	1	1																									
	Oqueoc	11	10																									1		
Lake Michigan	Little Manistee	11	2										2													2	2	3		
	Muskegon	12	6								1					1	2	1	1											
	Red Cedar	15	1																14											
	Kalamazoo	16				4																	8	2	2					
	Dowagiac	16	1																		15									
Total		207	80	1	1	4	2	4	2	2	1	3	2	7	2	1	2	1	15	17	15	16	14	2	2	2	2	4	1	2

Table 3. Estimates of molecular diversity in samples of *N. melanostomus* for each of the drainages sampled. N and N_h refer to the number of individuals and number of haplotypes per sample, respectively. Standard deviations are provided for each estimate.

Drainage	N	N_h	Gene Diversity	Mean No. of Pairwise Differences	Nucleotide Diversity
Stony Creek	30	1	0.0000 +/- 0.0000	0.000000 +/- 0.000000	0.000000 +/- 0.000000
Lower Rouge	30	1	0.0000 +/- 0.0000	0.000000 +/- 0.000000	0.000000 +/- 0.000000
Clinton	19	1	0.0000 +/- 0.0000	0.000000 +/- 0.000000	0.000000 +/- 0.000000
Rifle	4	1	0.0000 +/- 0.0000	0.000000 +/- 0.000000	0.000000 +/- 0.000000
Au Sable	20	2	0.1000 +/- 0.0880	0.100000 +/- 0.177536	0.000097 +/- 0.000193
Oqueoc	40	2	0.0500 +/- 0.0469	0.150000 +/- 0.216477	0.000146 +/- 0.000234
Manistee Lake	5	1	0.0000 +/- 0.0000	0.000000 +/- 0.000000	0.000000 +/- 0.000000
Pentwater Lake	25	2	0.0800 +/- 0.0722	0.240000 +/- 0.284615	0.000233 +/- 0.000308
Muskegon	20	2	0.2684 +/- 0.1133	0.268421 +/- 0.305890	0.000261 +/- 0.000332
Crockery Creek	20	2	0.1000 +/- 0.0880	0.300000 +/- 0.326275	0.000292 +/- 0.000354
Kalmazoo	14	1	0.0000 +/- 0.0000	0.000000 +/- 0.000000	0.000000 +/- 0.000000
St. Joseph	20	3	0.2789 +/- 0.1235	0.647368 +/- 0.523752	0.000629 +/- 0.000568

Table 4. Pairwise estimates of F_{ST} for populations of *N. melanostomus*. Significant values are underlined. See methods section for locality abbreviations.

	SC	CC	KZ	LM	MU	PW	AS	CL	LR	OQ	RF
CC	-0.0187										
KZ	0.04749	-0.019									
LM	-0.0477	-0.1047	0								
MU	<u>0.08421</u>	0.05263	0.07447	-0.0241							
PW	0.00095	-0.0459	-0.0255	-0.107	0.05935						
AS	0.06579	0	-0.019	-0.1047	0.07895	-0.004					
CL	0.07109	-0.0026	0	0	0.10065	-0.0115	-0.0026				
LR	0.11047	0.02104	0	0	0.14537	0.00744	0.02104	0			
OQ	-0.0462	-0.0289	-0.0326	-0.1095	<u>0.08392</u>	-0.0297	-0.0048	-0.021	-0.0074		
RF	0.07989	-0.1377	0	0	-0.0559	-0.1377	-0.1377	0	0	-0.1416	
SC	0.11047	0.02104	0	0	<u>0.14537</u>	0.00744	0.02104	0	0	-0.0074	0

Table 6. Estimates of molecular diversity for each sample of *E. nigrum*. N and N_h refer to the number of individuals and number of haplotypes per sample, respectively. Standard deviations are provided for each estimate.

Drainage	N	N_h	Gene Diversity	Mean No. of Pairwise Differences	Nucleotide Diversity
Stoney Creek	17	5	0.7647 +/- 0.0657	9.661765 +/- 4.659285	0.009219 +/- 0.004976
Raisin	13	3	0.4103 +/- 0.1539	0.435897 +/- 0.416881	0.000416 +/- 0.000447
Lower Rouge	21	2	0.3810 +/- 0.1005	0.380952 +/- 0.375176	0.000364 +/- 0.000400
Clinton	25	3	0.4767 +/- 0.0855	0.873333 +/- 0.634421	0.000833 +/- 0.000675
Shiawasee	14	3	0.4835 +/- 0.1425	0.527473 +/- 0.467361	0.000503 +/- 0.000501
Rifle	23	4	0.4506 +/- 0.1208	0.498024 +/- 0.440699	0.000475 +/- 0.000469
Au Sable	13	3	0.2949 +/- 0.1558	0.461538 +/- 0.431834	0.000440 +/- 0.000463
Oqueoc	11	2	0.1818 +/- 0.1436	1.272727 +/- 0.864354	0.001214 +/- 0.000930
Little Manistee	11	5	0.8727 +/- 0.0593	5.054545 +/- 2.657675	0.004823 +/- 0.002861
Muskegon	12	6	0.7576 +/- 0.1221	1.151515 +/- 0.798643	0.001099 +/- 0.000858
Red Cedar	15	2	0.1333 +/- 0.1123	0.133333 +/- 0.209858	0.000127 +/- 0.000225
Kalamazoo	16	4	0.7000 +/- 0.0896	1.433333 +/- 0.921163	0.001368 +/- 0.000985
Dowagiac	16	2	0.1250 +/- 0.1064	0.125000 +/- 0.202014	0.000119 +/- 0.000216

Table 7. Pairwise estimates of F_{ST} for *E. nigrum* populations. Significant values are underlined. See methods section for locality abbreviations.

	RN	LR	CL	SH	RF	AS	OQ	SC	LM	MU	RC	KZ
LR	0.59300											
CL	0.42047	0.62722										
SH	0.06667	0.57570	0.41789									
RF	0.04725	0.56864	0.43515	0.01816								
AS	0.02778	0.58206	0.41189	0.04210	0.02262							
OQ	0.02268	0.46580	0.35012	0.03354	0.03996	0.00709						
SC	0.15878	0.28477	0.28823	0.16683	0.22339	0.15921	0.10720					
LM	0.41412	0.55534	0.53071	0.42083	0.49298	0.41102	0.25814	0.11338				
MU	0.12001	0.50833	0.39239	0.12551	0.13964	0.10536	0.07027	0.15679	0.38089			
RC	0.76347	0.83729	0.69670	0.73357	0.71050	0.75303	0.59165	0.27691	0.55047	0.58243		
KZ	0.76165	0.81794	0.77040	0.75823	0.78376	0.75884	0.67100	0.19774	0.41001	0.71184	0.83311	
DG	0.77181	0.84190	0.70357	0.74241	0.71805	0.76167	0.60429	0.28644	0.56226	0.63595	0.93103	0.8381

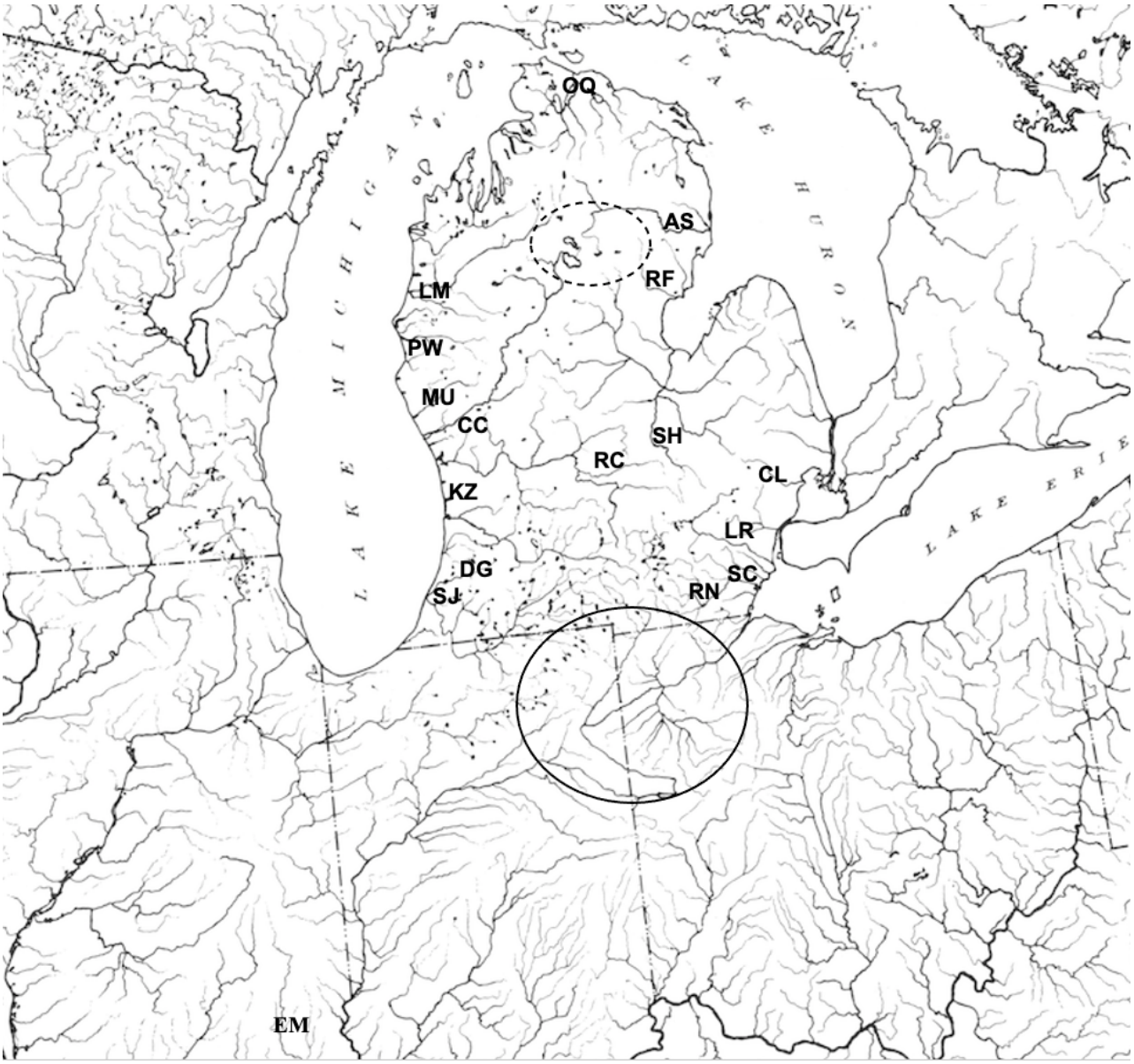

Figure 1. Map of collection localities for *N. melanostomus* and *E. nigrum*. Acronyms for sampling localities are: Au Sable (AS), Clinton (CL), Crockery Creek (CC), Dowagiac (DG), Kalamazoo (KA), Little Manistee (LM), Lower Rouge (LR), Muskegon (MU), Oqueoc (OQ), Pentwater (PW), Raisin (RN), Red Cedar (RC), Rifle (RF), St. Joseph (SJ), Shiawasee (SH), Stony Creek (SC). Precise location information is provided in the Methods section. A sequence of *E. nigrum* from the Embarras River (EM) in Illinois was obtained from GenBank. The dotted circle is the location of the headwaters of the Muskegon River and the Au Sable River. The solid circle shows the region where the Maumee River captured a portion of the Wabash River. Base map is reprinted from the Fish Division drainage map, University of Michigan Museum of Zoology, under a CC BY license, with permission from University of Michigan Museum of Zoology, original copyright 1972.

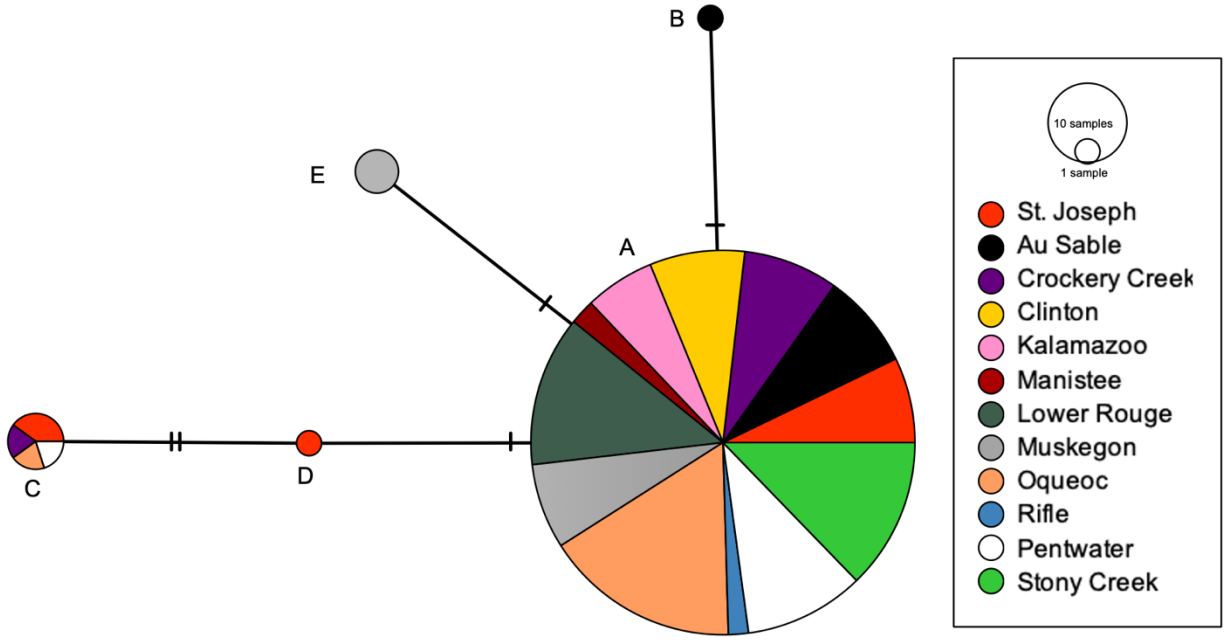

Figure 2. Median-joining network of *N. melanostomus* haplotypes. Pies show proportional representation of haplotypes by locality (indicated by different colors), with size of the circle reflecting the number of individuals. Mutations between haplotypes are indicated by vertical lines.

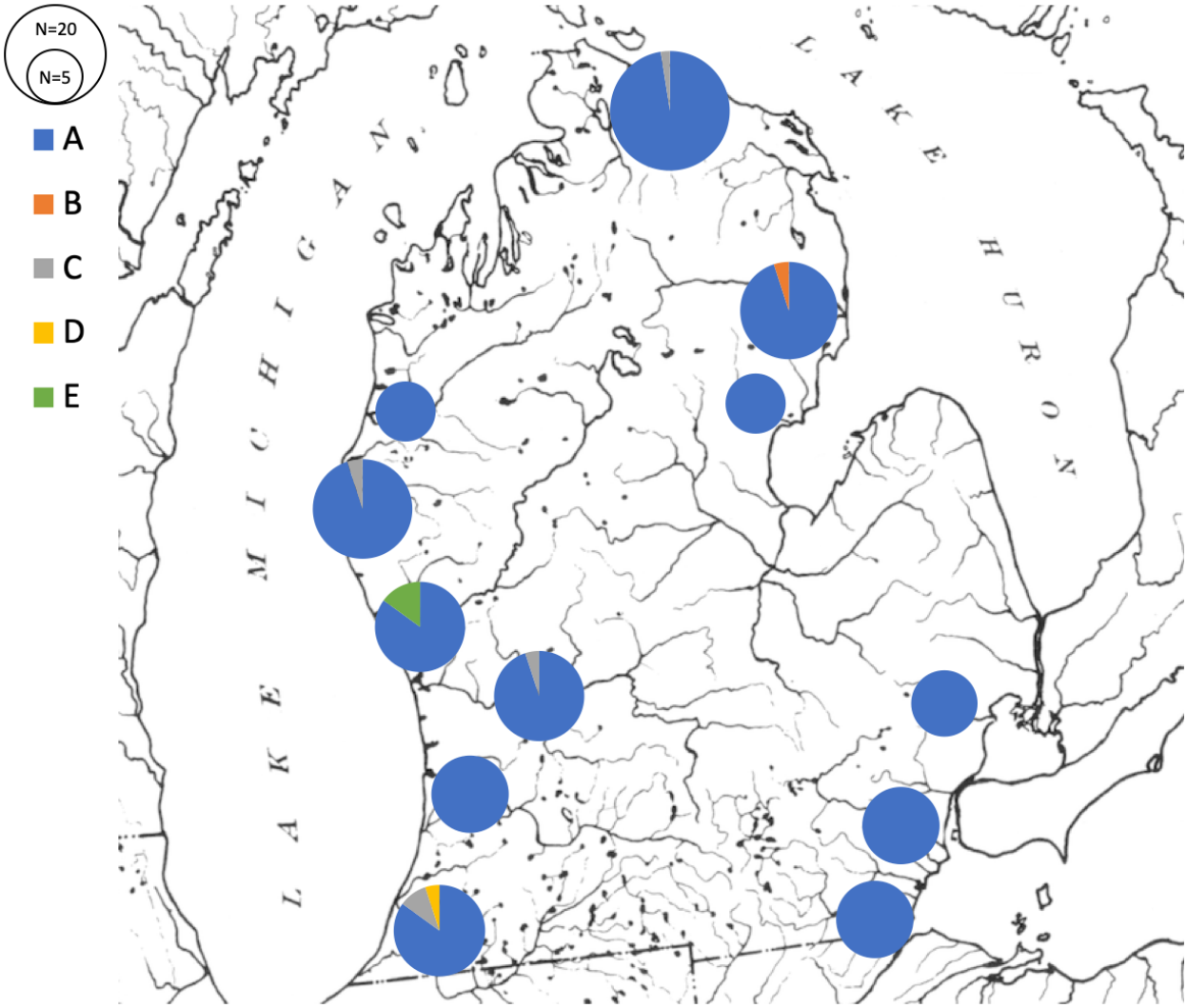

Figure 3. Distribution of haplotypes among sampling localities for *N. melanostomus*. The size of the circle reflects sample sizes, and the haplotype frequency is reflected by size of different colored slices. Haplotype colors are identified in the legend at the upper left.

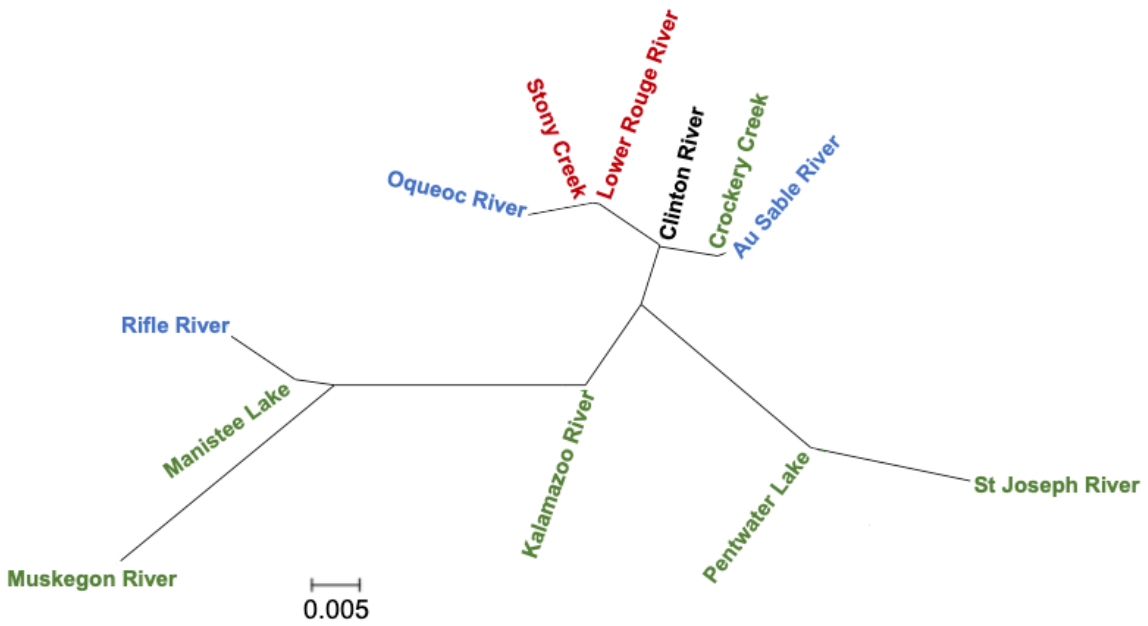

Figure 4. Neighbor-joining tree based of pairwise F_{ST} estimates for *N. melanostomus*. Color indicates major drainage: Black, Lake St. Clair; Green - Lake Michigan; Red - Lake Erie; Blue, Lake Huron.

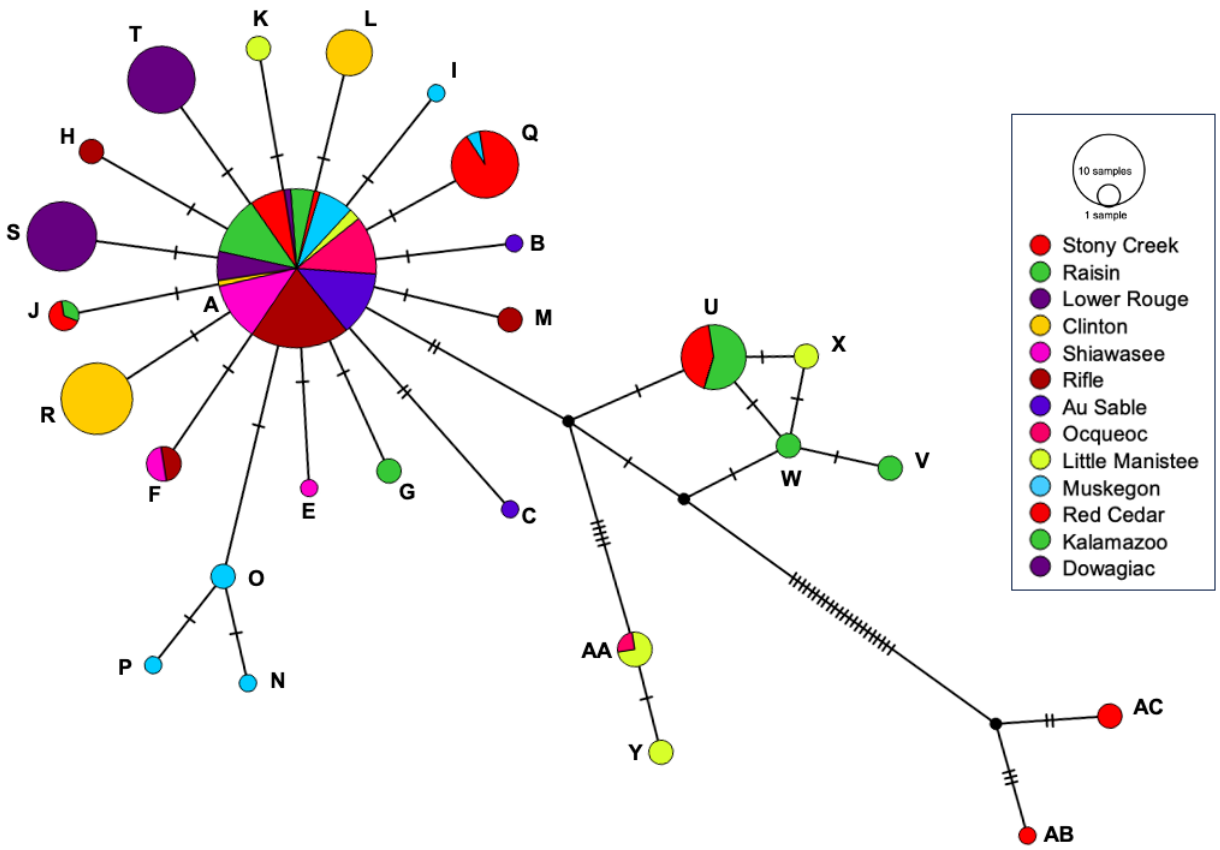

Figure 5. Median-joining network of *E. nigrum* haplotypes. Pies show proportional representation of haplotypes by locality, and size of the circle reflects number of individuals with that haplotype. Mutations between haplotypes are indicated by vertical lines, and the frequency of each haplotype is reflected by size of different colored slices, which are identified in the legend at the right.

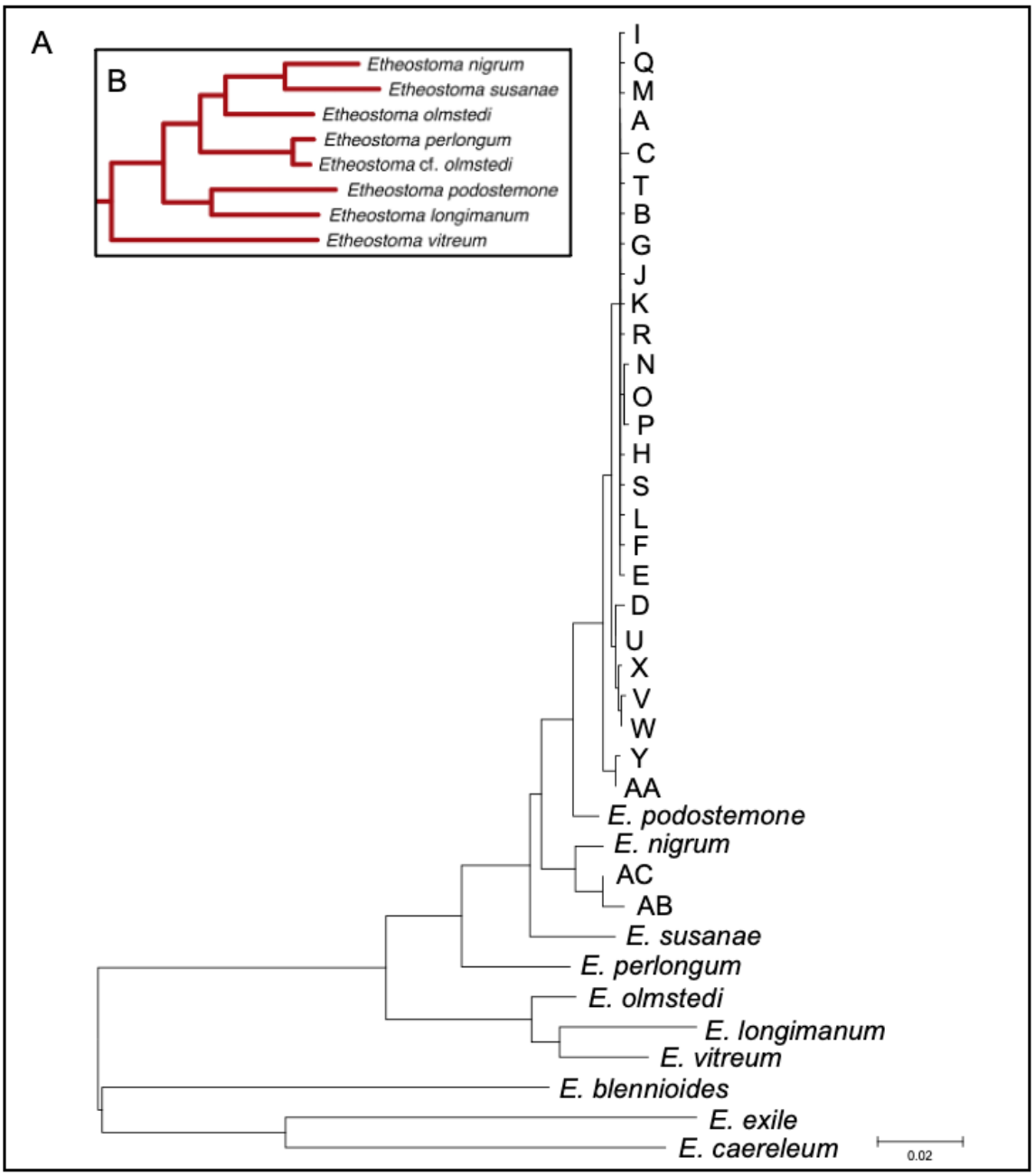

Figure 6. A) Maximum likelihood tree of *E. nigrum* haplotypes with bootstrap values indicated. The tree includes closely related species *E. olmstedi, E. podostemone*, and *E. perlongum*, *E. susanae, E. vitreum*, *E. longimanum* and more distantly related but co-occurring darter species *E. blennioides, E. exile*, and *E. caereleum*. The additional *E. nigrum* individual obtained from GenBank was sampled from the Wabash River drainage. B) Phylogeny based on single nucleotide polymorphisms modified from MacGuigan and Near (2018).

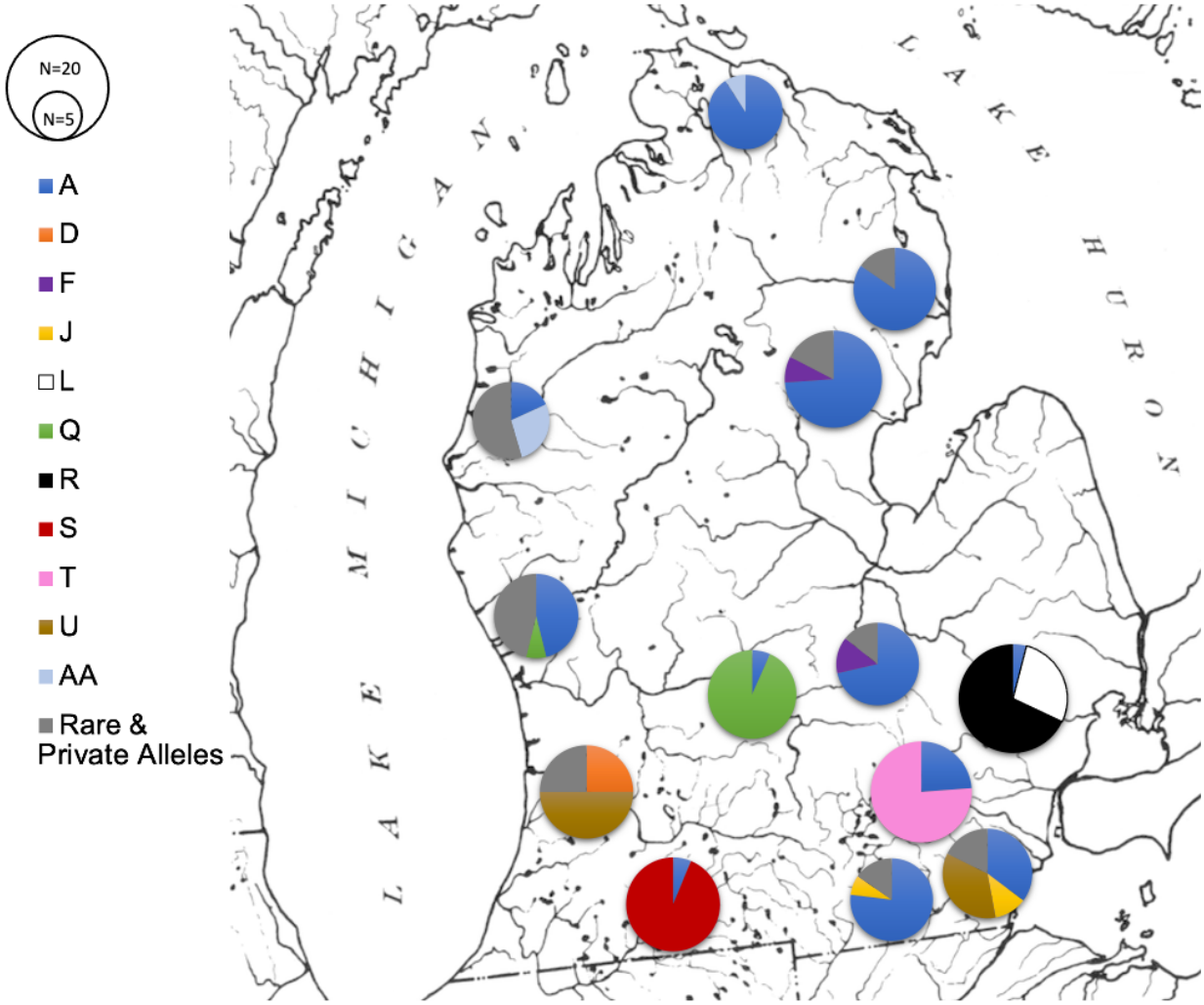

Figure 7. Distribution of haplotypes among sampling localities for *E. nigrum*. The size of the circle reflects sample size. The frequency of each haplotype is reflected by size of different colored slices, which are identified in the legend at the left. For ease of presentation, alleles that are both rare and private were collapsed into one category. Precise allele counts are available in Table 2.

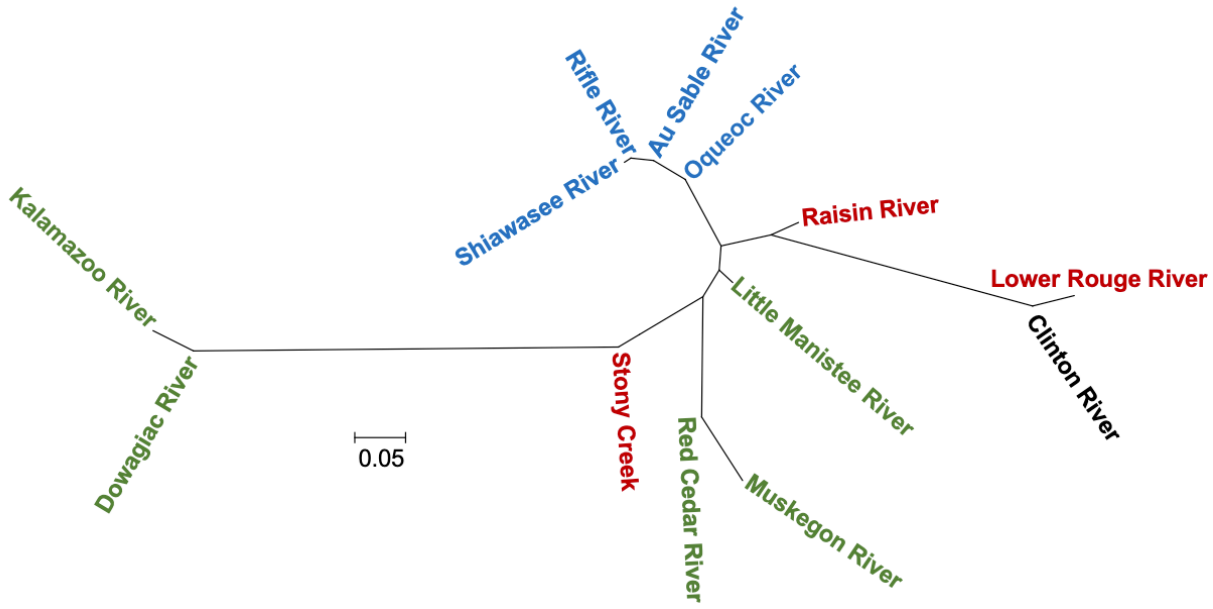

Figure 8. Neighbor-joining tree based on pairwise F_{ST} estimates for *E. nigrum*. Label color indicates major basin: Black - Lake St. Clair, Green - Lake Michigan, Red - Lake Erie, Blue -Lake Huron.

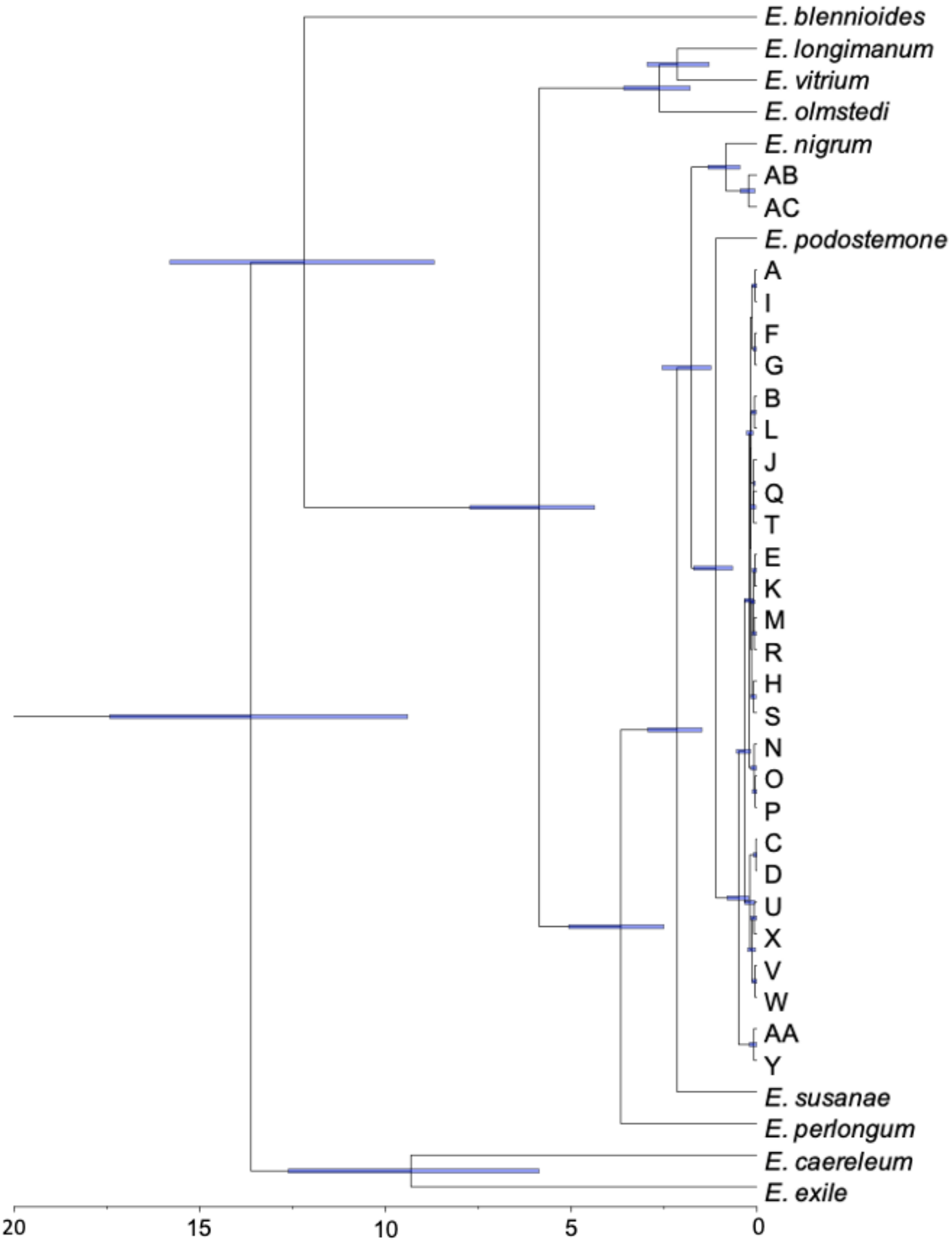

Figure 9. Time tree generated by molecular divergence analysis with *BEAST for *E. nigrum* haplotypes including related *Etheostoma* species. Blue bars represent the 95% HPD Scale is shown in millions of years before present.