Genetic structure of the benthic amphipod *Diporeia* (Amphipoda: Pontoporeiidae) and its relationship to abundance in Lake Superior

Erik M. Pilgrim¹*, Jill V. Scharold², John A. Darling¹, John R. Kelly²

¹U.S. Environmental Protection Agency, Ecological Exposure Research Division, 26 W. Martin Luther King Dr., Cincinnati, OH, 45268, USA

²U.S. Environmental Protection Agency, Mid-continent Ecology Division, 6201 Congdon Blvd., Duluth, MN, 55804, USA

*corresponding author, email: pilgrim.erik@epa.gov
other author emails: scharold.jill@epa.gov, darling.john@epa.gov, kelly.johnr@epa.gov

Erik M. Pilgrim
26 West Martin Luther King Dr.
Cincinnati, OH 45268
Phone: 513-569-7797
Fax: 513-569-7115
Email: pilgrim.erik@epa.gov
Abstract

The freshwater amphipod *Diporeia* is a crucial part of the food web in the Laurentian Great Lakes, but has faced serious declines correlated with the invasion of zebra mussels (*Dreissena polymorpha*), except in Lake Superior, which has seen an increase in *Diporeia* abundance. Speculation on the mechanisms causing changes in *Diporeia* densities has not considered the possibility of evolutionarily distinct lineages of *Diporeia* within the Lakes. In this study, we use COI DNA sequence data to investigate the evolutionary history of Lake Superior *Diporeia* relative to the other Great Lakes, and consider potential population structuring within Lake Superior based upon depth or geography. Our analyses reveal that Lake Superior *Diporeia* represent a distinct lineage that diverged from populations of the other lakes at least several hundred thousand years ago. F-statistics show that two localities within Lake Superior were significantly differentiated from all other locales, but analysis of molecular variance did not find significant structure based on depth or geography. Genetic diversity within Lake Superior was not correlated with depth, although abundance was significantly negatively correlated with increasing depth.

Keywords: *Diporeia*, Great Lakes, zebra mussels, amphipod
Introduction

The amphipod genus Diporeia Bousfield is restricted to deep, glacial relict lakes in northern North America (Bousfield 1989). In the Laurentian Great Lakes, Diporeia (6-9 mm long as adults) historically has accounted for 60-80% of the benthic biomass (Dermott et al. 2005). Since the 1990s, Diporeia has been in serious decline in the Great Lakes, being virtually extirpated from Lake Erie (Dermott and Kerec 1997), and declining drastically in Lakes Michigan (Nalepa et al. 2006a,b), Ontario (Lozano et al. 2001, Lozano and Scharold 2005, Watkins et al. 2007), and Huron (Nalepa et al. 2003). The reduced abundance is strongly correlated with the invasion first by the zebra mussel (Dreissena polymorpha) (Ward and Ricciardi 2007) and now by the quagga mussel (D. bugensis) (Watkins et al. 2007).

Diporeia have high lipid content (Cavaletto et al. 1996) and consequently are a very important food source for many fishes of the Great Lakes (Pothoven et al. 2001, Pothoven and Vanderploeg 2004). The decline of Diporeia has begun to impact several fishes in the Great Lakes, including both small prey and larger species that are important commercially and recreationally. Alewife (Alosa pseudoharengus), bloater (Coregonus hoyi), and slimy sculpin (Cottus cognatus) have shifted their diets away from Diporeia to other benthos, and the density of these fishes in the Great Lakes seems to be decreasing (Hondorp et al. 2005). Alewife also have exhibited a decrease in weight (Madenjian et al. 2003, Pothoven and Madenjian 2008) and energy density (Madenjian et al. 2006) which in turn may negatively affect the growth of Chinook salmon (Oncorhynchus tshawytscha) (Madenjian et al. 2006), a popular sport fish and important predator of alewife. Lake whitefish (Coregonus clupeaformis) have shown a decrease in the amount of Diporeia in their diets (70% down to 25%), which has led to detrimental changes in the growth patterns of this species, increased age when reaching sexual maturity.
(Pothoven et al. 2001, Pothoven and Madenjian 2008) and reduced egg production (Kratzer et al. 2007). Commercial harvest of whitefish in Lake Ontario declined from 295,000 kg in 1996 to 100,000 kg in 2001 (Hoyle 2005).

The *Diporeia* in Lake Superior have not experienced a similar decline, and recent studies have shown that densities there have not decreased substantially since the 1970s (Auer and Kahn 2004, Scharold et al. 2004, 2008). Lake Superior *Diporeia* are found at increased densities in near-shore environments (30-70 meters deep) (Scharold et al. 2004) in contrast with the distribution of remaining populations in the other Great Lakes where *Diporeia* are relegated to deeper refuges (Watkins et al. 2007). *Dreissena* had invaded Lake Superior by 1989 (O’Neill and Dextrase 1994) but are restricted to a few bays, possibly due to the physical and chemical characteristics of the lake that may have inhibited the expansion of dreissenids (Grigorovich et al. 2003). The *Diporeia* populations in Lake Superior therefore have not experienced the widespread encroachment of *Dreissena* into their habitats as in the other Lakes.

All the hypotheses put forward to explain the losses of *Diporeia* from the lower Great Lakes concurrent with the lack of decline in Lake Superior *Diporeia* rely on external influences (Nalepa et al. 2006a,b), and have thus far neglected the possibility that there may be significant evolutionary and, thus, ecological differences between distinct populations of *Diporeia* in the Great Lakes. Unfortunately, the taxonomy of the genus *Diporeia* currently is not well understood. *Diporeia hoyi* is considered to be the dominant species in the Great Lakes, but four other species of *Diporeia* (including two that are undescribed) have been reported to occur in the region (Bousfield 1989). Species in the genus can be difficult to identify as females dominate the life cycle and are collected predominantly instead of more morphologically distinct males (Bousfield 1989). Given the possibility of evolutionarily distinct lineages of *Diporeia*, variations
in declines may represent differential ecological responses to recent stressors by genetically divergent populations. The clear differences between current patterns of abundance in the Great Lakes may reflect underlying taxonomic differences between Diporeia populations.

The aim of this study was to test the null hypothesis that Diporeia in the Great Lakes represent a single evolutionary lineage. Using DNA sequence data from the mitochondrial cytochrome c oxidase subunit I (COI), we assess the degree of genetic divergence between Diporeia populations in Lake Superior and those in Lakes Huron, Michigan, and Ontario. We also explore in more detail the patterns of genetic variation found in Lake Superior by testing hypotheses regarding the structuring of genetic variation by geography and by depth, and by assessing relationships between genetic diversity and observed patterns of abundance. These results reveal fundamental differences between Diporeia in Lake Superior and populations remaining in the other Lakes; although the differences do not explain declines in other Lakes, the data provide important new background context on Diporeia distribution in and across Lakes.

Materials and methods

Sample Collection

Specimens of Diporeia were collected by multiple ponar grabs at each locality in August 2007. Depth and GPS coordinates were recorded for each collection site. Five of the collection sites came from near shore localities with depths less than 70 m (SN01, SN17, SU04, SU06, and SU22B), and three sites came from deeper, off-shore localities deeper than 150 m (SU10, SU11, SU19) (Table 1). Abundances in number of individuals per square meter were calculated for each collection site. The specimens were removed from the sediments and stored in 95% EtOH for later use in the molecular study. To compare Lake Superior samples to the other lakes,
samples of Diporeia were collected at one site in Lake Huron, two sites in Lake Michigan, and two sites in Lake Ontario (Table 1).

**Molecular study**

DNA extractions were done with one half to whole vacuum-dried specimens dependent on size of the individual. All specimens were extracted using the DNeasy Tissue Kit from QIAgen following the manufacturer’s protocol. PCR amplification of a 658 bp fragment of the mitochondrial cytochrome c oxidase subunit I gene (COI) was done in a 20 µL volume reaction under the following conditions: standard buffer concentration, 2.25 mM MgCl2, 200 pM dNTPs, 0.25 µM of each primer (standard DNA barcoding primers LCO-1490F and HCO-2198R (Folmer et al. 1994)), ½ unit of QIAgen Taq polymerase, 400 ng BSA, and approximately 20 ng of template DNA. The PCR amplification program was an initial step of 94°C for 150 sec, 35 cycles of 94°C for 30 sec, 46°C for 60 sec, and 72°C for 60 sec, and a final step of 72°C for 10 min. Additional PCR of the internal transcribed spacer regions (ITS1 and ITS2) was performed for 8 individuals (4 from Lake Michigan and 4 from Lake Superior with the primers gc18SF (5’-GGCGTCGTCGTGCTCG-3’) and gc28SR (5’-CCTCACCCCACCTAGTAG-3’)) following the above conditions and program. PCR products were cleaned using the QIAquick PCR kit on a QIAgen BioRobot 3000. Sequencing reactions were done with the ABI Big Dye Terminator Cycle Sequencing Ready Reaction kit following the manufacturer’s protocol. The sequenced products then were purified using the DyeEx 96 Kit from QIAgen, dried, and re-eluted with formamide, and then run on an ABI Prism 3730xl DNA Analyzer. All products were sequenced in both directions and were compiled into single contiguous sequences with Sequencher 4.8 (Gene Codes, Ann Arbor, MI). All sequences have been deposited in GenBank (accession numbers COI: EU761246 to EU761577; ITS: EU807701 to EU807708).
After alignment of the COI exports in Sequencher, the data set was analyzed using MEGA 3.1 (Kumar et al. 2004) to determine genetic distances (K2P model for corrected distances) and conduct neighbor-joining (NJ) cluster analysis, and Nona (v 2)/Winclada (v 0.9.99) (Goloboff 1999, Nixon 1999) were used for a maximum parsimony analysis with 1000 bootstrap pseudoreplicates. A minimum spanning network was generated using Network 4.5 (Fluxus Technology, Suffolk, England). The data set also was analyzed in Arlequin 3.11 (Excoffier et al. 2005) to generate haplotype diversity, nucleotide diversity, and fixation index ($F_{st}$) values. Analyses of molecular variance (AMOVA) were done in Arlequin to determine if either depth or location had an influence on population structure. To test for partitioning of genetic variance by depth, one group was defined containing all populations found at depths less than 100 m (near-shore: SN01, SN17, SU04, SU06, SU22B) and a second including those found at depths greater than 100 m (off-shore: SU10, SU11, SU19). To test for geographic structure, the populations were placed in eastern (SU04, SU06, SU10, SU11) or western (SN01, SN17, SU19, SU22B) groups based on their location relative to Keweenaw Point. To test for recent demographic expansion in the Lake Superior population we generated a distribution of the frequency of pairwise number of nucleotide mismatches between all Lake Superior COI haplotypes using Arlequin. Rapid population expansion results in the accumulation of mutations with minimal loss of lineages, resulting in a distinctive unimodal peak in this mismatch distribution. Steady-state populations, in contrast, exhibit multimodal or “ragged” distributions reflecting equilibrium between mutation accumulation and stochastic loss of lineages. The statistical significance of the observed distribution’s departure from the expectation of unimodality (the “raggedness index”) can be assessed by simulation to test the hypothesis of recent demographic expansion. In addition, since accumulation of mutations increases the mean
of the mismatch distribution (tau), that value can be used to estimate the time in generations since population expansion, given a known per-sequence mutation rate (Rogers and Harpending, 1992, Excoffier et al. 2005). To visually assess genetic relationships between Superior populations, we constructed a multi-dimensional scaling plot of the $F_{st}$ values using SAS (9.1.3).

**Results**

We were able to successfully sequence a 658 base pair (bp) fragment of COI for 235 individuals from Lakes Superior, 74 from Lake Michigan, 13 from Lake Ontario, and 10 from Lake Huron. Within Lake Superior we found 59 different haplotypes, and within the other three Great Lakes we found 37 haplotypes. Multiple haplotypes were found at each locality, ranging from 3 to 16 different haplotypes at the Lake Superior sites, and from 4 to 15 at the sites in Lakes Huron, Michigan, and Ontario (Table 1). In Lake Superior haplotype diversity ($H_e$) ranged from 0.4571 (site SU10) to 0.9048 (site SU22B) with an overall haplotype diversity of 0.8447 for the entire lake (Table 1). The near-shore, shallow sites had higher average haplotype diversity (mean = 0.7842) than the off-shore, deep sites (mean = 0.6146), but a t-test of the data did not show these differences to be significant (P = 0.157). An analysis of correlation between $H_e$ and abundance was not significant (P = 0.138), however, a strong negative correlation between abundance and depth was found (P= 0.009). For Lakes Huron, Michigan, and Ontario haplotype diversity ranged from 0.7500 to 1.0 with an overall diversity of 0.9156, which did not differ significantly from that found in Lake Superior (t-test P = 0.0567). Nucleotide diversity ranged from 0.000753 to 0.003860 (mean = 0.002417) in Lake Superior and from 0.001411 to 0.006293 (mean = 0.003773) in the other Great Lakes, and a t-test did not find the means significantly different (P = 0.184).
None of the COI haplotypes found in Lake Superior were found in the other Great Lakes (Fig. 1). Both cluster (NJ) and maximum parsimony analysis of the COI haplotypes across the four Great Lakes revealed a distinct separation between the Lake Superior haplotypes and those of the other Great Lakes with high bootstrap support (100% and 97%, respectively; trees not shown). Six fixed point differences were found between the Lake Superior haplotypes and the haplotypes of the other lakes (Table 2). The mean genetic distance observed between individuals in Lake Superior was 0.31% (range: 0 to 1.54%). Genetic distances between individuals of Lake Superior and individuals from the other Great Lakes had a mean of 1.69% (range: 0.92 to 2.64%).

Amplification of the ITS regions was successful for all 8 specimens attempted. The ITS1 was relatively short at 284 bp, but the ITS2 was considerably longer at an estimated 946 bp. Seven of the individuals (3 from Lake Michigan and 4 from Lake Superior) had identical sequences for both loci. The other individual from Lake Michigan differed by only a single base pair in the ITS2 and its ITS1 sequence was identical to the other 7 specimens.

The most common COI haplotype (SUP h01) in Lake Superior was found at all 8 sites and occurred in 85 (36.0%) individuals. This haplotype was rare at some sites (e.g. 11.1% of the population at SU19) and common at others (e.g. 73.3% at SU10). Of the 59 COI haplotypes found in Lake Superior, only 10 haplotypes were shared between multiple sites (Fig. 2). The other 49 haplotypes were unique to a given locality with 38 of these haplotypes found only in single individuals. The second most common haplotype (SUP h02—34 individuals) was found in individuals at SN01, SN17, SU19, and SU22B, all western sites. Most of the haplotypes (46 out of 58) differ from SUP h01 by 1 or 2 bp, and the overall average genetic distance between Lake
Superior haplotypes was 0.58%. Mismatch distribution analysis (Fig. 3) of all Lake Superior individuals revealed a value of 1.184 for \( \tau \), and a raggedness index of 0.028 (\( P = 0.99 \)).

A comparison of F\textsubscript{st} values between localities suggested that some population structure may be present in Lake Superior. Two western, near-shore populations (SN17 and SU22B) were significantly differentiated from all the eastern populations and also from SU19, the only western, off-shore population (Table 3). The populations SU04 (eastern, near-shore) and SU19 (western, off-shore) had significant F\textsubscript{st} values when compared against all other populations. A multi-dimensional scaling plot incorporating all F\textsubscript{st} values between populations shows populations SU04 and SU19 as potential outliers to a cluster of the remaining populations (Fig. 4). Despite AMOVA results indicating significant genetic differentiation between individual collection sites (consistent with FST estimates), overall differentiation between western and eastern regions of Lake Superior was not significant (\( P = 0.0704 \)). Similarly, we found no evidence to support the hypothesis of genetic differentiation between near- and off-shore sites in Superior (AMOVA partitioned by depth, \( P = 0.462 \)) (Table 4).

**Discussion**

The lack of shared COI haplotypes between populations of *Diporeia* in Lake Superior and populations in Lakes Huron, Michigan, and Ontario renders these populations reciprocally monophyletic and strongly suggests absence of significant gene flow between *Diporeia* of Lake Superior and the other Great Lakes. Although the *Diporeia* in Lake Superior appear to have diverged evolutionarily from populations in the other Great Lakes, our data are currently not strong enough to support consideration of the Lake Superior *Diporeia* as a distinct species. The minimum interpopulation genetic distance is only 0.92% for COI, which is considerably smaller than the maximum intrapopulation variation found in Lake Superior (1.54%). Furthermore, the
identical sequences of the ITS1 and ITS2 regions (two loci that are often considered to show species-specific differences (Pilgrim and Pitts 2006, Pilgrim and von Dohlen 2007)) that occur in the different lake populations are also suggestive that the *Diporeia* in all the Great Lakes have not diverged enough to show differences in these nuclear loci. More extensive sampling of individuals from Lakes Huron, Michigan, and Ontario would be necessary to resolve whether these lineages constitute separate species. Our current sampling is not consistent, however, with the hypothesis that as many as four different *Diporeia* species (Bousfield 1989) occur in these lakes.

Whether treated as a single or separate species, the genetic distance between the populations of Lake Superior and the other lakes suggests significant evolutionary divergence between the two lineages. An estimate of time of divergence based on a widely utilized COI mutation rate (1.4%/MYA) in crustaceans (Knowlton and Weigt 1998) using the minimum distance between haplotypes (0.92%, representative of the six fixed mutational differences between lineages) places the split between Lake Superior *Diporeia* and the other Great Lakes lineage at least 650,000 years ago (Pleistocene). Using more conservative COI mutation rates (0.19% to 0.55%/MYA) advocated by other authors (Schön et al. 1998, de Bruyn 2005) places the divergence at 1.67 MYA (Pleistocene) to 4.84 MYA (Pliocene). Although some consider dating nodes based on estimated mutations rates rather than fossil evidence to be problematic (Heads 2005), all the estimates here place the split between Lake Superior *Diporeia* and the other Great Lakes at least several hundred thousand years before the most recent formation of the Laurentian Great Lakes (10,000-15,000 years ago). More concrete estimates of the divergence between the Great Lakes populations would be better addressed with a phylogeny of the entire genus. The unimodal mismatch distribution (Fig. 3) for the Lake Superior lineage is indicative of
a rapid demographic expansion in the *Diporeia* lineage currently inhabiting that lake, and using
the estimates of mutation rate for COI above, this expansion can be placed between 400,000 and
3.1 million years ago. These estimates are consistent with a scenario of rapid population
expansion in the Lake Superior lineage following evolutionary divergence from the lineage
founding populations in Huron, Michigan, and Ontario, with both events greatly pre-dating the
formation of the Laurentian Great Lakes. These estimated divergence times also are comparable
to divergence estimates found in freshwater fishes such as white sucker (*Catastomus
commersoni*) (Lafontaine and Dodson 1997), brown bullhead (*Ameiurus nebulosus*) (Murdoch
and Hebert 1997), lake trout (*Salvelinus namaycush*) (Wilson and Hebert 1998), and banded
killifish (*Fundulus diaphanus*) (April and Turgeon 2006), all of which have had their
evolutionary histories and geographic distributions influenced by past North American glacial
events.

The segregation of the Lake Superior *Diporeia* from the populations of the other Great
Lakes is consistent with population studies of fish species distributed throughout the lakes. The
walleye (*Stizostedion vitreum*) of Lake Superior were distinct from lakes Michigan, St. Clair,
Erie, and Ontario based on mtDNA control region haplotypes (Stepien and Faber 1998).
Microsatellite data show that populations of smallmouth bass (*Micropterus dolomieu*) in Lake
Superior have not experienced gene flow with the other Great Lakes (Stepien et al. 2007). The
distinctness of the Lake Superior populations of smallmouth bass and walleye are also congruent
with patterns found in yellow perch (*Perca flavescens*) and brown bullhead (*Ameiurus
nebulosus*) (Stepien et al. 2007). The dispersal capabilities of these fish species very likely is
much greater than that of *Diporeia*, and therefore the lack of gene flow in this amphipod between
Lake Superior and the other lakes should be expected.
Although the abundance of *Diporeia* outside Lake Superior has declined drastically, the comparable haplotype diversity between Lake Superior and the other Great Lakes (except for Lake Erie where *Diporeia* have been virtually extirpated) suggests that the decline has not resulted in a genetic bottleneck for the remaining populations. Within Lake Superior, the near-shore localities exhibit higher genetic diversity based both on a higher percentage of individuals with unique haplotypes (22.5% to 16.1%; see fig. 2) and on overall haplotype diversity (0.7842 vs. 0.6146), but these differences in diversity were not significant. The widely disparate abundances in *Diporeia* seen here (Table 1) and in previous studies (Scharold et al. 2004, 2008) between near-shore and off-shore collection sites originally led us to investigate genetic population structure in the lake. No significant population structure was found between shallow and deep collection sites and this suggests that gene flow regularly occurs between near-shore and off-shore habitats. Comparisons of eastern vs. western collection localities also did not show significant structure. An eastern, near-shore population (SU04) and a western, off-shore population (SU19) were both shown to be significantly different from all other populations based on F_{st} values, implying that some population structuring is present in the lake. Although these two populations are outliers, they do suggest that population structuring exists in *Diporeia* of Lake Superior, but that neither of the broad hypotheses tested here (east vs. west or near-shore vs. off-shore) is a sufficient explanation for that structure. Dispersal in *Diporeia*, however, is male biased (Bousfield 1989), and as a consequence, the maternally inherited COI haplotypes may underestimate gene flow among populations within the lake. More comprehensive sampling and the use of nuclear markers may provide a more complete picture of genetic structure within Lake Superior.
The populations of *Diporeia* in Lake Superior, especially the near-shore habitats, have not suffered a decline correlated with zebra mussel invasion (Scharold et al. 2004, 2008) as have the *Diporeia* of the other Laurentian Great Lakes. In fact, the abundance of *Diporeia* in Lake Superior seems to have increased 5- to 8-fold over levels seen in the early 1970s (pre-*Dreissena* invasion), possibly due to a decrease in pollution and the rebound of the lake trout, which feeds on the predators of *Diporeia* (Scharold et al. 2004, 2008). The data in this study show that *Diporeia* of Lake Superior have diverged from the populations of the other lakes at least several hundred thousand years ago, and the distinct evolutionary history of Lake Superior *Diporeia* should be taken into account in any work that seeks to explain the current status of *Diporeia* within Lake Superior. That the zebra mussel occurs in Lake Superior is not in question, but it does not appear to be fully established in the lake (O’Neill and Dextrase 1994, Grigorovich et al. 2003, Scharold et al. 2004) because known populations are not consistently found year-to-year (Grigorovich et al. 2003), possibly due to Lake Superior’s physical and chemical characteristics such as depth, temperature, and nutrient content which may approach the habitat limits of zebra mussels (Grigorovich et al. 2003). Certainly, the divergent evolutionary history of the Lake Superior *Diporeia* could be one factor in the current high population abundances in Lake Superior, but to assume that the genetic differences are the only explanation is not prudent, especially considering the invasion of Lake Superior by zebra mussels has not been as severe as in the other Great Lakes. This is further emphasized by the fact that in the Finger Lakes of New York, *Diporeia* populations are not declining despite the presence of high densities of dreissenids (Nalepa et al. 2006b), although they belong to the same genetic lineage as the *Diporeia* found in Huron, Michigan and Ontario (data not shown). The Lake Superior populations and the lower Great Lakes populations each could be considered unique evolutionary lineages warranting
management as the loss of Diporeia in Lakes Huron, Michigan, and Ontario would not be mitigated by the lack of decline in Lake Superior Diporeia. The Diporeia of Lake Superior may face new challenges in the recent introduction of the quagga mussel (D. bugensis) (Vanderploeg et al. 2002, Grigorovich et al. 2008), or by expansions in the ranges of invasive fish such as ruffe (Gymnocephalus crenuus) (Bauer et al. 2007) or the round goby (Neogobius melanostomus).

Acknowledgements

We would like to thank Jim Watkins for his extensive help during the collection of these samples, as well as Tim Corry, Sam Miller, and Mario Picinich and the crew of the Lake Guardian for their help with sample collection. We also thank Kenneth Oswald and two anonymous reviewers for providing helpful comments on earlier versions of this manuscript.
References


Murdoch, M.H. and Hebert, P.D.N. 1997. Mitochondrial DNA evidence of distinct glacial


<table>
<thead>
<tr>
<th>Site</th>
<th>Locality</th>
<th>Depth</th>
<th>Abundance (m⁻²)</th>
<th>No. Individuals</th>
<th>No. Haplotypes</th>
<th>Haplotype diversity</th>
<th>Nucleotide diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lake Superior</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN01</td>
<td>46.75378N, 91.91447W</td>
<td>31m</td>
<td>3569</td>
<td>32</td>
<td>13</td>
<td>0.8085 ± 0.0605</td>
<td>0.002831 ± 0.001852</td>
</tr>
<tr>
<td>SN17</td>
<td>46.79446N, 91.56222W</td>
<td>32m</td>
<td>1449</td>
<td>62</td>
<td>16</td>
<td>0.8112 ± 0.0345</td>
<td>0.002854 ± 0.001834</td>
</tr>
<tr>
<td>SU04</td>
<td>46.52617N, 84.83017W</td>
<td>69m</td>
<td>2254</td>
<td>28</td>
<td>8</td>
<td>0.7487 ± 0.0702</td>
<td>0.003349 ± 0.002124</td>
</tr>
<tr>
<td>SU06</td>
<td>47.10950N, 88.21517W</td>
<td>68m</td>
<td>2043</td>
<td>23</td>
<td>10</td>
<td>0.6403 ± 0.1165</td>
<td>0.001694 ± 0.001279</td>
</tr>
<tr>
<td>SU10</td>
<td>47.51416N, 87.54617W</td>
<td>157m</td>
<td>13</td>
<td>15</td>
<td>3</td>
<td>0.4571 ± 0.1406</td>
<td>0.000753 ± 0.000763</td>
</tr>
<tr>
<td>SU11</td>
<td>47.37033N, 85.96500W</td>
<td>208m</td>
<td>290</td>
<td>29</td>
<td>9</td>
<td>0.7266 ± 0.0859</td>
<td>0.002313 ± 0.001592</td>
</tr>
<tr>
<td>SU19</td>
<td>47.13333N, 90.85400W</td>
<td>191m</td>
<td>51</td>
<td>18</td>
<td>4</td>
<td>0.6601 ± 0.0781</td>
<td>0.001679 ± 0.001287</td>
</tr>
<tr>
<td>SU22B</td>
<td>46.80000N, 91.74967W</td>
<td>55m</td>
<td>1785</td>
<td>28</td>
<td>14</td>
<td>0.9048 ± 0.0341</td>
<td>0.003860 ± 0.002381</td>
</tr>
<tr>
<td><strong>Lake Huron</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HU96B</td>
<td>44.58317N, 81.50000W</td>
<td>49m</td>
<td>n/a</td>
<td>10</td>
<td>8</td>
<td>0.9556 ± 0.0594</td>
<td>0.004712 ± 0.003038</td>
</tr>
<tr>
<td><strong>Lake Michigan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI11</td>
<td>42.38267N, 87.00017W</td>
<td>130m</td>
<td>n/a</td>
<td>43</td>
<td>15</td>
<td>0.7553 ± 0.0683</td>
<td>0.003272 ± 0.002057</td>
</tr>
<tr>
<td>MI40</td>
<td>44.75984N, 86.96684W</td>
<td>173m</td>
<td>n/a</td>
<td>31</td>
<td>14</td>
<td>0.8968 ± 0.0322</td>
<td>0.003177 ± 0.002029</td>
</tr>
<tr>
<td><strong>Lake Ontario</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ON41</td>
<td>43.71667N, 78.02700W</td>
<td>135m</td>
<td>n/a</td>
<td>8</td>
<td>4</td>
<td>0.7500 ± 0.1391</td>
<td>0.001411 ± 0.001230</td>
</tr>
<tr>
<td>ON55M</td>
<td>43.44333N, 77.43834W</td>
<td>201m</td>
<td>n/a</td>
<td>5</td>
<td>5</td>
<td>1.0000 ± 0.1265</td>
<td>0.006293 ± 0.004457</td>
</tr>
</tbody>
</table>
Table 2. Fixed differences in COI haplotypes between Lake Superior and the other Great Lakes. Numbers are relative to the start of the COI fragment analyzed in the current study.

<table>
<thead>
<tr>
<th></th>
<th>202</th>
<th>265</th>
<th>421</th>
<th>458</th>
<th>496</th>
<th>589</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>other Great Lakes</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>T</td>
</tr>
</tbody>
</table>
Table 3. Pairwise $F_{st}$ values between sites in Lake Superior, with statistical support. Significant differences are shown in bold.

<table>
<thead>
<tr>
<th></th>
<th>SN01</th>
<th>SN17</th>
<th>SU04</th>
<th>SU06</th>
<th>SU10</th>
<th>SU11</th>
<th>SU19</th>
</tr>
</thead>
<tbody>
<tr>
<td>SN17</td>
<td>$F_{st}$</td>
<td>0.00510</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>0.28829</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU04</td>
<td>$F_{st}$</td>
<td>0.17296</td>
<td><strong>0.18152</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU06</td>
<td>$F_{st}$</td>
<td>0.03561</td>
<td><strong>0.10025</strong></td>
<td><strong>0.23498</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU10</td>
<td>$F_{st}$</td>
<td>0.08256</td>
<td><strong>0.15052</strong></td>
<td><strong>0.30710</strong></td>
<td>0.00474</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>0.05405</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td>0.27027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU11</td>
<td>$F_{st}$</td>
<td>0.02804</td>
<td><strong>0.07905</strong></td>
<td><strong>0.19823</strong></td>
<td>0.00189</td>
<td>0.03013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td>0.30631</td>
<td>0.10811</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SU19</td>
<td>$F_{st}$</td>
<td>0.17127</td>
<td><strong>0.13849</strong></td>
<td><strong>0.28071</strong></td>
<td><strong>0.30311</strong></td>
<td><strong>0.38689</strong></td>
<td><strong>0.26154</strong></td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
</tr>
<tr>
<td>SU22B</td>
<td>$F_{st}$</td>
<td>0.02384</td>
<td>0.03517</td>
<td><strong>0.14717</strong></td>
<td><strong>0.10882</strong></td>
<td>0.17073</td>
<td>0.08291</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>0.09009</td>
<td>0.05405</td>
<td>&lt;&lt;0.0001</td>
<td><strong>0.00901</strong></td>
<td>&lt;&lt;0.0001</td>
<td>&lt;&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Eastern vs. Western</td>
<td>Near Shore vs. Off Shore</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variance Components</td>
<td>Percentage of Variation</td>
<td>Fixation Indices</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Among Groups</strong></td>
<td>0.01798</td>
<td>4.11</td>
<td>0.0411</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Populations within Groups</td>
<td>0.04261</td>
<td>9.73</td>
<td>0.1015*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Populations</td>
<td>0.37719</td>
<td>86.16</td>
<td>0.1384*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Among Groups</strong></td>
<td>-0.00447</td>
<td>-1.05</td>
<td>-0.0105</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Populations within Groups</td>
<td>0.05486</td>
<td>12.83</td>
<td>0.1270*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Populations</td>
<td>0.37719</td>
<td>88.21</td>
<td>0.1179*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P-value < 0.05
Figure Captions:

Fig. 1. Minimum-spanning network of *Diporeia* COI haplotypes and their North American Great Lakes of origin (E: Lake Erie; H: Lake Huron; M: Lake Michigan; O: Lake Ontario; S: Lake Superior). The top left boxed cluster is a group of haplotypes found only in Lake Superior. The top right boxed cluster is a group of haplotypes found in Lakes Huron, Michigan, and Ontario. The size of each circle is proportional to the number of individuals that had that haplotype. Unsampled/missing intermediate haplotypes are marked with black squares. The colored portions of each circle in the network correspond to the colored box of each locality. The thicker black line between the two networks highlights the six fixed base pair differences between the *Diporeia* haplotypes of Lake Superior and those of Lakes Huron, Michigan, and Ontario.

Fig. 2. Haplotype distribution among populations in Lake Superior. The dark blue areas of each pie chart denote the proportion of the most common haplotype at each locality. The white areas of the pie charts denote haplotypes that are unique to that population. The other color patterns show haplotypes shared among at least two populations.

Fig. 3. Observed mismatch distribution for Lake Superior COI haplotypes (columns) plotted with the simulated expectation (dotted line) based on the assumption of rapid demographic expansion followed by stable large population size. The raggedness statistic estimates departure of the observed distribution from the null model expectation, and its significance is indicated in the figure.
Fig. 4. Multi-dimensional scaling plot based on the $F_{st}$ values between each population in Lake Superior. Gray square: eastern, near-shore; black square: eastern, off-shore; gray circle: western, near-shore; black circle: western, off-shore. Six of the localities form a cluster around the lower left quadrant. Population SU19 (a western, off-shore population) and Population SU04 (an eastern, near-shore population) are significantly different from all other populations based on $F_{st}$ values.
Figure 1.
Figure 2.
Figure 3.

 Tau = 1.168

Raggedness index = 0.0279

P-value = 0.99