

****ABSTRACT NOT FOR CITATION WITHOUT AUTHOR PERMISSION.** The title, authors, and abstract for this completion report are provided below. For a copy of the full completion report, please contact the author via e-mail at tkrabben@buffalo.edu. Questions? Contact the GLFC via email at frp@glfc.org.

Genomics of Great Lakes ciscoes (*Coregonus* spp.): identifying genetic differences associated with morphological differences

Trevor J. Krabbenhoft¹, Nathan J. C. Backenstose¹, Daniel L. Yule², Wendylee Stott³, Thomas E. Dowling⁴

¹Department of Biological Sciences, State University of New York at Buffalo

²USGS Great Lakes Science Center, Lake Superior Biological Station, Ashland, Wisconsin

³Ontario Ministry of Northern Development, Mines, Natural Resources and Forestry (OMNDMNRF)

⁴Department of Biological Sciences, Wayne State University, Detroit, Michigan

July 2022

ABSTRACT:

A key aim of the Great Lakes Coregonine Restoration Framework is to resolve coregonine taxonomy through genetics and ecology. We used an array of genetic tools (transcriptomics, long-read amplicon sequencing, low-coverage genome sequencing, and moderate-coverage genome sequencing) to evaluate patterns of molecular variation within and among species of Great Lakes coregonines (ciscoes – Genus *Coregonus*). We focused mainly on Lakes Superior and Nipigon, where extant diversity remains high and opportunities to understand the relationship between genetic, morphological, and ecological variation are afforded. We evaluated patterns of genetic variation in light of morphological and ecological data, some of which was also new in this research. We summarize results of this project in four manuscripts (Appendix I-IV), two of which are published and two of which are in preparation for publication. Appendix I: We (Bernal et al., 2022) evaluated the relationship between transcriptome, geometric morphometric, and stable isotope data, and discovered general concordance among these datasets. We also identified candidate genes that are strongly associated with different phenotypes (species) and may represent causal genetic variants that literally shape some of the morphological variation seen in nature. Appendix II: We validated some of these candidate genes, specifically related to spectral tuning of visual perception and light availability across a depth gradient (Eaton et al. 2021). This study involved development of a novel genetic approach (long-read amplicon sequencing) and, strikingly, found identical genetic variants in ciscoes in a vision gene, rhodopsin, as are found in independently evolved deepwater marine taxa, strengthening our conclusions that we have identified a causal genetic variant under natural selection for differences in vision across *Coregonus* species. This gene likely played an important role in cisco speciation and more importantly for fisheries management – it likely is an important determinant of habitat preferences of different *Coregonus* species or forms. Thus, the understanding of this genetic variation can help us predict which habitats particular individuals/populations are likely to select and how hatchery fish might respond to the light environments we place them in in the wild. Appendix III: Backenstose et al. (in prep) expanded on the transcriptome results with moderate coverage, whole genome sequencing – the first time complete genomes have been sequenced in ciscoes. We confirmed the general population genetic relationships found by Bernal et al. (2022) and also used coalescent analysis of these data to evaluate the historical timing and environmental context for divergence of Great Lakes ciscoes. Our results indicated that contemporary cisco diversity likely originated in Pleistocene pro- or periglacial lakes just before the formation of the Great Lakes basin. This diversification was likely driven by a response to the novel appearance of deepwater, lentic habitats (niches) that differed strongly from the glacial riverine refugia used by ancestral ciscoes. By providing an understanding of the historical origins and maintenance of cisco diversity, this study helps us better contextualize the contemporary restoration and rehabilitation challenges facing fisheries managers, along with future threats such as ongoing climate change and invasive species. It also provides a

much-needed geologic timeline related to the Great Lakes Coregonine Restoration Framework's goal of clarifying cisco taxonomy. Appendix IV: The first three manuscripts necessarily relied on small sample sizes due to cost factors. In Appendix IV, we (Backenstose et al., in prep) employed low-coverage genome sequencing across much larger sample set (277 whole genomes) to provide a more robust understanding of the extent of variation within and between Lakes Superior, Huron, and Nipigon. We also preliminarily combined geometric morphometric analyses with our genome data to better link genetic and morphological variation more broadly than in Bernal et al. (2022). We identified a large-scale genetic break between species found in Lakes Superior and Nipigon, suggesting independent evolution of those faunas (a finding with important implications for fisheries management and the taxonomy of these nominal species). In addition, these low-coverage genomes support previous findings of genetic differentiation among cisco species, but with occasional contemporary gene flow (introgressive hybridization). Finally, preliminary analyses link genetic variation across a depth gradient with eye size, strengthening previous findings from Eaton et al. (2021) and Bernal et al. (2022). Collectively, the four manuscripts resulting from this project provide a detailed picture of the historical origins of Great Lakes coregonines, evolutionary forces driving those differences (in genetics, morphology, and ecology), and the fisheries management, restoration/rehabilitation, and conservation implications of these findings.