

****ABSTRACT NOT FOR CITATION WITHOUT AUTHOR PERMISSION.** The title, authors, and abstract for this completion report are provided below.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.

Cisco (*Coregonus artedi*) reference genome project: an integrative resource for management and restoration of Great Lakes coregonines

Trevor J. Krabbenhoft², Sarah L. Chang², Nathan J.C. Backenstose², Daniel J. MacGuigan², Amanda S. Ackiss³

²109 Cooke Hall, University at Buffalo, Buffalo, NY 14260

³USGS Great Lakes Science Center, 1451 Green Rd., Ann Arbor, MI 48105

March 2024

ABSTRACT:

Management of Great Lakes coregonines (*Coregonus* spp.) requires a diverse suite of datasets for informed and adaptive decision making. A major knowledge gap has been a detailed understanding of coregonine genomes and how they vary within and among species, forms, populations, stocks, lakes, sexes, etc. – until recently. We sequenced and constructed a reference genome assembly of *Coregonus artedi* (Otoonapi, Cisco). We attempted to contextualize previous genetics work by mapping existing genetic marker panels (i.e., GT-seq markers from Ackiss et al., and RAD-seq markers from the cisco genetic linkage map from Blumstein et al.) to this reference assembly, which provides chromosomal coordinates for individual markers and their relation to each other. This spatial genomic information informs the relative heritability and statistical independence of markers for ongoing population genetic analyses being conducted in genetic labs throughout the Great Lakes basin. This resource will be permanently available publicly as a resource for fisheries genetics research and management via online databases on publication of the manuscripts. We then used whole genome resequencing data mapped to the new Cisco genome assembly to study the environmental context and geological timeline of *Coregonus* ecological diversification. These data provide a historic example of when conditions favored a dramatic increase in population sizes and diversity of ecologically differentiated forms – key information for contemporary efforts to achieve the same. Finally, we present the first genome-wide DNA methylation data ever collected (to our knowledge) for Great Lakes cisco species, which may play a role as an initial driver of ecological and morphological divergence, and their maintenance among forms or species in the Great Lakes. DNA methylation data has many potential applications to Great Lakes fisheries management, including markers for age predictions, population differentiation, hatchery versus wild markers, cohort strength analysis, etc., and these initial baseline data will provide a valuable resource for the next wave of studies using this newly available data source. Overall, this study presents a wealth of new information and improved genomic context for existing genetic data, both of which are useful for data-driven fisheries management.