Using genomics to delineate stock structure and create a standardized genetic resource for Great Lakes walleye

Peter T. Euclide[2], Wesley Larson[3], Emily Latch[4], Brian Dixon[5], Matthew Faust[6], Loren Miller[7], Kim Scribner[8], Wendylee Stott[9], Chris Wilson[9]

[2] Department of Forestry and Natural Resources, Purdue University

[3]USGS Wisconsin Cooperative Fishery Research Unit, University of Wisconsin-Stevens Point

[5]Department of Biology, University of Waterloo

[6]Ohio Department of Natural Resources

[7]Minnesota Department of Natural Resources

[8]Department of Fish & Wildlife, Michigan State University

[9]Ontario Ministry of Northern Development, Mines, Natural Resources and Forestry, Aquatic Research Section, Trent University

July 2022

ABSTRACT:
Genetic and genomic resources are being developed at a rapid pace, offering powerful tools that can help protect and sustain ecologically important fish populations and the valued fisheries that they support. Herein, we discuss our recent genetic/genomic project to develop new molecular tools for researching Walleye Sander vitreus biology and management in the Great Lakes. Starting in 2019, we used restriction site associated DNA sequencing to census genome-wide genetic diversity of 45 walleye collected from spawning sites representing each of the Great Lakes. The resulting data set informed the design of a 99,636-bait Rapture panel to facilitate more efficient high-density genome sequencing of walleye. Using this panel, a much broader census of genetic diversity was conducted on 1,289 walleye collected from 29 spawning sites throughout the Great Lakes. Genetic markers from this broad data set were selected to create a genotyping-in-thousands (GTSeq) panel capable of discriminating walleye spawning sites and conducting accurate kinship analysis. We tested 5 alternative marker selection scenarios, targeting SNPs and microhaplotypes with high allele frequency differentiation (high FST) and high genetic diversity (high observed heterozygosity). The best performing scenario included a set of 450 high-FST and 150 high heterozygosity makers; this set was used to create a robust amplicon sequencing GTSeq panel.

Following the removal of 100 markers during PCR and sequencing optimization, the final GTSeq panel contained 500 variants (303 SNP and 197 microhaplotype markers) with an average observed heterozygosity of 0.38. Simulations of genetic stock identification (GSI) conducted separately for each lake indicated that the final GTSeq panel should be capable of identifying the local spawning stock of walleye with >80% accuracy in most cases. Loss of GSI accuracy occurred when stocks were in close geographic proximity or in areas with established patterns of gene flow, such as Lake Erie. Simulations of pairwise kinship indicated that the final GTSeq panel should provide sufficient resolution to identify parent-offspring and full-sibling relationships with low false-positive rates (< 1 X 10 -11), and to a lesser extent half-sibling relationship (false-positive rate < 0.03). Genotypes generated independently for the same individuals in different laboratories were >94% congruent, showing standardized data production among laboratories even when the depth-of-coverage was relatively low (16.9X). Early analysis of GTSeq data from 633 walleye collected from important managed walleye stocks around the Great Lakes (Green Bay, Saginaw Bay, and the West Basin of Lake Erie) further indicated that the GTSeq panel provides consistent and statistically powerful population genetic data. Together these results indicate that the GTSeq panel we developed is a powerful, robust, and broadly applicable molecular tool that can be used to inform the important and multijurisdictional managed walleye stocks in the Great Lakes.