

Final Narrative Report

Project Title:

Management unit delineation of lake sturgeon populations based on adaptive genetic diversity

Grantee Organization:

West Virginia University

Project Team:

Amy Welsh, West Virginia University

Eric Normandeau, Université Laval

Louis Bernatchez, Université Laval

Contact Person:

Amy Welsh

West Virginia University

amy.welsh@mail.wvu.edu

Grant Amount:

\$193,942

Start and End Dates:

12/2/2015 – 8/31/2019

Key Search Words:

lake sturgeon, genomic, SNP, management unit

Background/Overview

Lake sturgeon management units in the past have been delineated based on neutral genetic markers (i.e., microsatellites). However, these markers may not reflect the different evolutionary pathways taken by these populations and may thus be poor estimators of their adaptive variation. To alleviate this limitation, we proposed to analyze lake sturgeon throughout the Great Lakes using genomic data (i.e., SNPs) that included both neutral and adaptive variation. Then, adaptive variation was going to be used to group populations based on adaptive similarity, resulting in management units that will help conserve the adaptive potential of our remaining lake sturgeon populations.

No significant changes were made to the originally proposed work. Some spawning locations in the original proposal had to be excluded from analyses due to low-quality sequencing results from poor samples or small sample sizes. Excluded populations due to small sample size were the Michipicoten River, Nipigon River, and Des Prairies River. Excluded populations due to poor quality were the Kaministiquia River, Mississauga River, Lake Nipissing, and Lake Champlain. Populations with poor quality sequences will be reanalyzed in future studies using different techniques. Even after removing these populations, a sufficient geographic scope and coverage was maintained for the adequate assessment of overall population structure.

Outcomes

Key project outcomes include knowledge, training, relationships, and practice. This research project advanced the scientific knowledge of lake sturgeon ecology and the role of genomics in conservation in genomics. Previous analyses of lake sturgeon population structure were based on neutral genetic data, tracking the roles of migration and genetic drift in determining population structure. We now have a better understanding of how adaptive variation is distributed across the landscape. Prior to this study, questions remained regarding the presence of fine-scale population structure in some systems and genetic differences between phenotypes. Our study detected the possible presence of multiple populations within the St. Clair River and the St. Lawrence River. Our study also supported the likely plasticity of migratory phenotypes in the St. Clair River.

The project contributed to the advancement of graduate students focused on Great Lakes fishery issues. Justine Whitaker was a Ph.D. student mentored by Amy Welsh and her research focused on genetic differences within the St. Clair River using microsatellite loci. Although Dr. Whitaker has since graduated, the project provided the opportunity for her further advancement as she helped to analyze and interpret the portions of the project focused on fine-scale population structure and differences between phenotypes. She is currently working on preparing those publications for submission.

The project helped to build new relationships with others in the research community. Dr. Wesley Larson (University of Wisconsin – Stevens Point) is active in fish genomic research. Our work with lake sturgeon genomics has led to a collaboration to expand our project. We will be working with Dr. Larson on a project to develop a SNP panel for future assignment testing and parentage analysis. This new collaboration will use the SNP data generated from this project to identify the most informative SNPs and develop protocols for their analysis.

The findings of this project have action implications for fishery managers. The designated management units provide refinement to those identified in the Genetic Stocking Guidelines (Welsh et al. 2010). This will help to inform source population decisions for stocking. Lake sturgeon have also been proposed for listing under the U.S. Endangered Species Act; our results can help inform the designation of Distinct Population Segments and Designatable Units under the Canadian Species At Risk Act.

Of all the factors listed above, the most important outcome is the advancement in knowledge. With increasing emphasis on the collection of genomic data, the utility of microsatellite data has been questioned. The similar patterns observed between the two types of data in our study demonstrate that previously collected microsatellite data are reflecting various evolutionary processes for population differentiation. However, the increased power and resolution of the SNP data set generated in this project will be a useful baseline for future assignment and parentage studies of lake sturgeon.

Related Efforts

This project was a standalone effort. The only portion of the project that was not funded by the GLFT was sample collection. All samples used in this study were from previous projects where samples were collected for the purpose of microsatellite data analysis. No other funders have been involved in this project. As mentioned above, this work has resulted in spin-off work, inspiring subsequent related research. We are currently working on a project with W. Larson on the development of a SNP panel for assignment testing and parentage analysis using data generated from this project. This new project is funded by the Great Lakes Fish and Wildlife Restoration Act.

Communication/Publication of Findings

Sharing of the findings of this research has primarily been through oral presentations delivered at regional, national, and international conferences. We are currently working on preparing three manuscripts for submission to peer-reviewed journals. The technical report attached to this report (“A genome-wide approach to management unit delineation of lake sturgeon populations”) details the results of the project. We ask that you restrict access to the materials while we seek publication.

Project Deliverables

Publications:

- Welsh, A.B., Normandeau, E., Scribner, K., and Bernatchez, L. (In prep) Genomics and the delineation of lake sturgeon management units in the Great Lakes.
- Whitaker, J.M., Price, L.E., Boase, J.C., Bernatchez, L., and Welsh, A.B. (In prep) Detecting fine-scale population structure in the age of genomics: a case study of lake sturgeon in the Great Lakes.
- Whitaker, J.M., Hondorp, D., Boase, J.C., and Welsh, A.B. (In prep) Genomic variation between migratory phenotypes in the St. Clair River.

Presentations:

- Welsh, A., Bernatchez, S., Normandeau, E., and Bernatchez, L. (2018) What the age of genomics means for lake sturgeon management. Great Lakes Lake Sturgeon Coordination Meeting, Port Huron, MI
- Welsh, A., Whitaker, J., Price, L., and Boase, J. (2018) Genomics and fine-scale population structure in lake sturgeon. North American Sturgeon and Paddlefish Society conference, Columbia, MO
- Welsh, A., Bernatchez, S., Normandeau, E., and Bernatchez, L. (2017) Defining management units in the age of genomics. International Sturgeon Symposium, Vienna, Austria.

Discussion

As we move further into conservation genomics, management units may get redefined as we become able to access genome-wide levels of variation. Based on the analyses of population structure

using SNP data and the geography of the spawning locations, we propose the identification of seven management units within the Great Lakes (Figure 7). In general, most groupings correspond to geography. Higher levels of distinction were observed within Lake Superior and less genetic differentiation was observed in the lower Great Lakes. Some groupings span multiple lake basins. For example, the Goulais River appears to be a mixture of populations and is more closely related to the Georgian Bay of Lake Huron. This location is at the juncture of Lakes Superior and Huron, likely resulting in a higher level of mixture among spawning populations.

Some management unit delineations differ from the genetic results and represent a more conservative approach to management unit designation. Based on the SNP data, the Menominee River in western Lake Michigan is most similar to the Goulais River in Lake Superior, the Manistee River in eastern Lake Michigan, and the Spanish and St. Clair rivers in Lake Huron. Using only adaptive SNPs, it is more similar to the Spanish and Goulais rivers. Based on the microsatellite data, the Menominee River is most closely related to its neighbor, the Wolf River, and the St. Lawrence River in Lake Ontario. Given these unclear groupings, the Menominee River was designated as its own management unit. The Wolf River consistently grouped with the St. Lawrence River when using the adaptive SNPs and the microsatellite loci. However, given their geographic distance, those two populations were assigned to different management units. Their genetic similarity could be an artifact of sharing the same refuge during the last glaciation (Ferguson and Duckworth 1997). Finally, the St. Lawrence and Grasse rivers are genetically distinct in all analyses. Despite this genetic distinction, they are placed in the same management unit. The Grasse River population had the lowest genetic diversity using both SNPs and microsatellites. This indicates that the population may have lost diversity through genetic drift due to its isolation behind a dam, resulting in genetic differentiation from the St. Lawrence River. Historically, prior to dam construction, the Grasse River was likely connected to the St. Lawrence River. Therefore, reconnection between the two locations should be a management priority and designation within the same management unit reflects that likely historical connection.

Low levels of fine-scale population structure were detected at the two large-river systems in this study. The St. Clair and St. Lawrence rivers are the largest systems in the Great Lakes. The St. Lawrence River has multiple dams along it that could have potentially resulted in genetic differentiation among locations. A telemetry study in the St. Clair River documented little movement between lake sturgeon that occupied the Detroit River and those that occupied the St. Clair River (Kessel et al. 2018), suggesting the possibility of genetically distinct populations. Given the large size of both systems, the existence of multiple spawning locations, and the potential barriers on the St. Lawrence River, each river may contain multiple populations, which should be further explored. Previous analyses on a fine geographic scale have generated mixed results. Most studies have been conducted within the Hudson Bay drainage, where lake sturgeon are primarily river residents. Within-river population structure has been identified in lake sturgeon in the Winnipeg and Ottawa Rivers (McDougall et al. 2017; Wozney et al. 2011), but fine-scale analysis of lake sturgeon in the Namakan River revealed the presence of only a single population (Welsh and McLeod 2010).

Limited fine-scale population structure also was detected based on migration phenotype. Five distinct migration behaviors were identified in the St. Clair River system based on telemetry data, two of which were year-round river residency and migratory (spending majority of their time in Lake Huron or Erie) (Kessel et al. 2018). A previous study was not able to detect any differences between the groups using morphology or microsatellite loci (Whitaker et al. 2018). However, Whitaker et al. (2018) were

able to detect epigenetic differences between the groups, suggesting that those phenotypic differences may be plastic and do not have a strong genetic basis.

The overall patterns detected with the genome-wide SNP data and the original microsatellite data were similar, despite microsatellites being criticized for inadequate representation of the full genome and being selectively neutral. Surprisingly, the greatest corroboration of the microsatellite data was with the putatively adaptive SNPs, where similar population structure patterns were observed. The high genetic distinction of the Grasse and Bad rivers was observed using both the adaptive SNPs and the neutral microsatellites. The unexpected grouping of the St. Lawrence River with the Wolf River was also observed using in both data sets. This pattern of outlying loci producing the same population structure as neutral loci was also observed in Atlantic salmon (*Salmo salar*) (Moore et al. 2014), suggesting that microsatellite loci may be doing better at tracking adaptive variation than previously expected.

The greatest advantage to the use of SNPs in future lake sturgeon genetic research is the increased power of assignment testing. Assignment accuracy greatly increased when using the total SNP data set (84.4%) compared to the 12 microsatellite loci (33.3%). This accuracy will likely further improve through exclusive use of the most highly-differentiating SNPs, which may differ according to region. For example, the Manistee River had poor assignment accuracy using the total or neutral SNP data set, but improved when only the adaptive SNPs were used. Assignment testing has been frequently pursued in lake sturgeon studies to determine movements of lake sturgeon during non-spawning periods (e.g., Homola et al. 2010, Marranca et al. 2015, Welsh et al. 2017, 2019) and the current suite of microsatellite loci has often provided limited power and resolution. Use of SNPs will greatly improve the power of both assignment testing and family reconstruction and this study has generated the necessary baseline information for future studies.

As conservation *genetics* moves further into conservation *genomics*, it is critical that we reconsider how we define evolutionarily significant units (ESUs) and management units (MUs), which can translate into the designation of Distinct Population Segments (DPSs) under the U.S. Endangered Species Act and Designatable Units under the Canadian Species at Risk Act. When neutral data were being collected using microsatellite loci, it was hoped that genetic differences correlated to adaptive differences that were evolutionarily significant. With genomic data that include both neutral and potentially adaptive variation, we can more directly assess adaptive differences between populations. Funk et al. (2012) propose an approach where ESUs are first identified using both neutral and adaptive SNPs, MUs are then identified using only neutral SNPs, and the last step is the identification of adaptive groups using SNP outliers. However, this approach is relegating the variation with the greatest evolutionary significance to a category below MUs. Instead, to preserve adaptive variation between populations, the first step should be the identification of ESUs using adaptive loci, followed by the further delineation of MUs using neutral variation. This suggestion will more likely result in the conservation of the adaptive potential of remaining populations, one of the primary goals of conservation genetics.