Using Molecular Techniques to Determine Effective Number of Breeders (N_b) for Razorback Sucker and Colorado Pikeminnow in the San Juan River

Fiscal Year 2021 Scope of Work

June 23, 2020

Principle Investigators
Steven Mussmann
Email: steven_mussmann@fws.gov
Office Phone: 575-734-5910 x115
Southwestern Native Aquatic Resources and Recovery Center
7116 Hatchery Road
Dexter, NM 88230

Melody Saltzgiver
Email: melody_saltzgiver@fws.gov
Office Phone: 575-734-5910 x116
Southwestern Native Aquatic Resources and Recovery Center
7116 Hatchery Road
Dexter, NM 88230

United States Department of the Interior
Fish and Wildlife Service
Southwestern Native Aquatic Resources and Recovery Center
P.O. Box 219, 7116 Hatchery Road
Dexter, New Mexico 88230
575-734-5910, 575-734-6130 fax
June 2020

Title: Using Molecular Techniques to Determine Effective Number of Breeders (N_b) for Razorback Sucker and Colorado Pikeminnow in the San Juan River

Principal Investigator: Steve Mussmann and Melody Saltzgiver, Southwestern Native Aquatic Resources and Recovery Center, Dexter, NM

Introduction & Justification

Recent advances in population-level genetic analyses are increasingly helping biologists monitor and adaptively manage the recovery of endangered populations (Hartl & Clark 2007). Recovery plans for many endangered fishes include the production and release of hatchery-reared individuals to augment populations with two major objectives: increase population size and promote genetic diversity (Miller & Kapuscinski, 2003). Survival, reproduction, and recruitment of wild and augmented individuals must occur at a sustainable scale in order to meet these recovery criteria. However, factors that limit success in achieving self-sustaining populations can be difficult to identify. This frequently requires long-term datasets to provide insight into population responses that might hinder recovery. Long-terms datasets include utilizing genetic monitoring to understand complex ecological, demographic, and genetic factors that can impede the establishment of self-sustaining populations (Schwartz et al. 2006; De Barba et al. 2010).

Augmentation of endangered fish populations in the San Juan River using captive-reared Razorback Sucker (*Xyrauchen texanus*) and Colorado Pikeminnow (*Ptychocheilus lucius*) began in the mid-1990s and continues as a recovery action (USFWS 2005; USFWS 2015). Annual monitoring of survival, reproduction, and recruitment of these populations has been supported through the San Juan River Basin Recovery Implementation Program (SJRRIP). Mark-recapture data on PIT tagged individuals have enabled estimation of survival for both stocked species (Franssen & Durst *unpublished data*; Clark et al. 2018), which in turn has prompted additional research investigating ways to increase survival. Larval fish surveys have documented successful reproduction of both species in the river (Farrington et al. 2015), but recruitment to the adult stage is extremely limited. Successful recovery requires a significant portion of the reestablished population to reproduce annually to both increase population sizes and ensure maintenance of genetic diversity. Therefore, quantifying the number of individuals that reproduce annually provides data to make informed management decisions to aid in the reestablishment of self-sustaining populations.

Population-level spawning success (i.e., number of reproducing adults) is innately difficult to quantify from field studies, especially for highly fecund species where few individuals can produce a large number of offspring. Furthermore, there is substantial evidence that shows reproductive output can depend on environmental conditions and age- or size-related factors (Lauer et al. 2005; Lambert 2008). Both temporal and spatial variation in spawning effort has been detected for Razorback Sucker. Adults may vacate a spawning area early in the season

and return to spawn again later that year (Marsh et al. 2015), with individuals visiting multiple sites during the same spawning period (Modde & Irving 1998). Such reproductive strategies further compound the difficulty in determining individual contribution to cohorts over a reproductive season; however, this question can be addressed by estimating the effective number of breeders (N_b) using genetic analyses. Colorado Pikeminnow and Razorback Sucker have defined seasonal reproductive bouts, thus making N_b an extremely useful metric for understanding population-level spawning success despite these species having life history traits (i.e., long-lived, highly fecund, and iteroparous with overlapping generations) that typically confound similar analyses (Waples et al. 2013; Waples et al. 2014). Single cohort N_b estimates can reliably quantify the number of individuals that contribute to a given cohort (Waples et al. 2014). Obtaining annual N_b estimates for the endangered fishes of the San Juan River provides insight into how management activities (i.e., increasing passage, managing flows) affect population-level reproduction, thus, improving the opportunity to manage the San Juan River to favor spawning success.

Objectives:

1. Continue N_b monitoring for both Razorback Sucker (N = 120) and Colorado Pikeminnow (N = 120) collected in the mainstem San Juan River during 2021.

Methods

Larval fish surveys are conducted annually along a 140-mile section of the San Juan River between Shiprock, NM, and Clay Hills, UT. Up to 120 larval samples per species representing the spatial and temporal distribution of these sampling efforts will be examined to calculate N_b estimates. Tissue subsamples will be taken from the posterior portion of Razorback Sucker and Colorado Pikeminnow larvae collected during 2021. The anterior portion of all specimens will be saved for otolith studies. Rare collections will be targeted while sites with high larval densities will be proportionally reflected in samples in order to ensure N_b estimates are not artificially lowered due to limited spatial representation of samples. Larval Razorback Sucker are collected during much of the sampling season. Early larval stages (e.g., protolarvae to mesolarvae) will be targeted throughout larval collections under the assumption that these individuals are from recent spawning events; thus ensuring sampling is representative of the temporal spawning season. Conversely, Colorado Pikeminnow are collected later in the sampling season, making it relatively easy to ensure captured individuals reflect the entire seasonal spawning period of adults. Samples collected below the waterfall will be used to evaluate larval emigration or adult contribution downstream of the waterfall.

Previous N_b estimates have been obtained using multi-locus microsatellite markers (Diver & Wilson 2018); however, the field of population genetics has shifted to utilize massively parallel “next-generation sequencing” (NGS) methods. This technology provides a cost-effective and efficient means for sampling large quantities of genetic data from individuals through “reduced representation” genomic sequencing (Campbell et al. 2018). These methods sequence a small fraction of an organism’s genome to identify thousands of single nucleotide polymorphisms (SNPs) that are analogous to microsatellite markers for conservation genetic applications. However, SNPs quantify an order of magnitude more loci compared to microsatellites (i.e., SNPs = hundreds or thousands of loci, microsatellites = 10 – 30 loci) which will functionally increase our resolution of genetic variation among individual genomes. Application of SNP-based methods for determining relatedness and N_b estimates has increased as microsatellite-based studies have reduced in number (Flanagan & Jones 2019). SNPs have been successful in reconstructing multi-generational pedigrees in the absence of known parental relationships (Levine et al. 2019). Comparative studies have found that SNP-based methods provide improved

confidence intervals and equal or greater power for determining relationships among individuals (Thrasher et al. 2018; Lemopoulos et al. 2019; Galla et al. 2020). SNP data also promote other analyses of interest in aquatic conservation, including, but not limited to, the assessment of hybridization in complex systems (Chafin et al. 2019). Finally, SNPs offer benefits over traditional microsatellite methods due to lower error rates and broader genome coverage (Smouse 2010; Hauser et al. 2011). Therefore, N_b estimates for 2021 will continue to be collected using these methods.

Genomic DNA will be extracted from tissues following standard protocols used at Southwestern ARRC. NGS data will be prepared using double digest Restriction-Site Associated DNA (ddRAD) libraries (Peterson et al. 2012). Restriction digest of 1µg genomic DNA/sample will be performed in 50µl reactions containing 5µl New England BioLabs CutSmart Buffer and 20 units each PstI and MspI. Samples will be digested at 37°C for 18-24 hours then purified using Agencourt AMPure XP beads (Beckman Coulter, Inc.). Barcoded samples (100 ng DNA each) will be pooled in sets of 48 following Illumina adapter ligation, then size-selected using the Pippin Prep System (Sage Science) to retrieve DNA fragments between 350 and 400 bp in length for Razorback Sucker and 325 to 425 for Colorado Pikeminnow (Bangs et al. 2018; Chafin et al. 2018). Size-selected DNA will be subjected to 12 cycles of PCR amplification using Phusion high-fidelity DNA polymerase (New England Bioscience), according to manufacturer protocols. Four indexed libraries (192 samples) will be pooled per lane for 100bp single-end sequencing on an Illumina HiSeq 4000 (University of Oregon Genomics & Cell Characterization Core Facility). Data will be de-multiplexed and filtered in STACKS 2 (Rochette et al. 2019) to discard reads with uncalled bases or low Phred quality scores (<10), while simultaneously attempting to recover those reads with ambiguous barcodes (=1 mismatched nucleotide). The *de novo* assembly of ddRAD loci will be accomplished in STACKS 2 (Rochette et al. 2019) with clustering parameters being determined by the methods of Rochette & Catchen (2017). Only loci appearing in 90% of individuals will be retained for analysis. Data will be filtered to retain one SNP per ddRAD locus. A second researcher will perform a 10% quality assurance/quality control of samples to ensure accuracy.

SNPs will be evaluated for linkage disequilibrium and Hardy-Weinberg Equilibrium (Raymond & Rousset 1995; Purcell et al. 2007). N_b will be estimated using the sibship-assignment (SA) method in COLONY version 2.0.4.0 (Jones & Wang 2010). This software uses a maximum likelihood method to estimate relationships among offspring belonging to a single cohort by identifying full and half-sibling families while considering genotyping error rates and allowing for inferences related to the mating strategy of the organism. Analysis will be conducted separately for each year to estimate N_b, the number of adults that contributed at least one offspring, number of sampled offspring produced by each parent, and the number of parental pairs. For both species, male and female polygamy will be assumed and parameter settings (i.e., dioecious, diploid, inbreeding, medium run length, full-likelihood with medium likelihood precision, no sibship prior, and updated allele frequencies) will be maintained across years.

Schedule:

Completion of genetic analysis	March 30, 2022
Final Report	June 30, 2022

Deliverables:

Dissemination of the results will include a final report and presentation of project results at the San Juan Researcher's meeting.

FY21 - Detailed Spending Plan

1. PERSONNEL	
A. Laboratory Work	
1 Bio/Geneticist (GS-11-2; 200 hours – 2 pay periods) @ $44.58/hr	$8,916.84
B. Data Analyses and Report Writing	
1 Bio/Geneticist (GS-11-3; 320 hours - 4 pay periods) @ $46.12/hr	$14,758.75
C. Travel	
2 Trips to San Juan Meeting (4 days each; with travel) @ $664.50/trip	$1,329.00
1 Trip to Museum of Southwestern Biology to obtain larval samples (4 days; including travel) @ $576.50/trip	$576.50
Subtotal Personnel	**$25,581.09**
2. MATERIALS/SUPPLIES/SEQUENCING	
A. Extractions	$1,274.00
B. Library Prep	$3,991.00
C. Library Quantification/QAQC	$2,280.00
D. Sequencing	$3,112.00
E. Other (tubes, tips, etc.)	$1,795.00
Subtotal Supplies	**$12,452.00**
Southwestern ARRC Utilities	
-Electrical, (approx. 4,259 KW/h @ 0.34569 per KW/h) =	$1,472.29
Total	**$39,505.38**
Administrative and Overhead Costs Regional Office @ 3%	$1,185.16
Project Total FY2021	**$40,690.54**

Literature Cited

Bangs, M.R., M.R. Douglas, S.M. Mussmann, and M.E. Douglas. 2018. Unraveling historical introgression and resolving phylogenetic discord within *Catostomus* (Osteichthyes: Catostomidae). BMC Evolutionary Biology 18:86.

Campbell, E.O., B.M.T. Brunet, J.R. Dupuis, F.A.H. Sperling. 2018. Would an RRS by any other name sound as RAD? Methods in Ecology and Evolution 9:1920-1927.

Chafin, T.K., B.T. Martin, S.M. Mussmann, M.R. Douglas, and M.E. Douglas. 2018. FRAGMATIC: in silico locus prediction and its utility in optimizing ddRADseq projects. Conservation Genetics Resources 10(3):325-328.

Chafin, T.K., M.R. Douglas, B.T. Martin, and M.E. Douglas. 2019. Hybridization drives genetic erosion in sympatric desert fishes of western North America. Heredity 123:759-773.

Clark, S.R., M.M. Conner, S.L. Durst and N.R. Franssen. 2018. Age-Specific Estimates Indicate Potential Deleterious Capture Effects and Low survival of Stocked Juvenile Colorado Pikeminnow. North American Journal of Fisheries Management. 38(5): 1059-1074.

De Barba, M., L.P. Waits, E.O. Garton, P. Genovesi, E. Randi, A. Mustoni, and C. Groff. 2010. The power of genetic monitoring for studying demography, ecology, and genetics of a reintroduced brown bear population. Molecular Ecology 19:3938-3951.

Diver, T.A. and W.D. Wilson. 2018. Using Molecular Techniques to Determine Effective Number of Breeders (N_b) for Razorback Sucker and Colorado Pikeminnow in the San Juan River. San Juan River Basin Recovery Implementation Program, Albuquerque, New Mexico.

Farrington, M.A., R.K. Dudley, J.L. Kennedy, S.P. Platania, and G.C. White. 2015. Colorado Pikeminnow and Razorback Sucker larval fish survey in the San Juan River during 2014.

Flanagan, S.P. and A.G. Jones. 2019. The future of parentage analysis: From microsatellites to SNPs and beyond. Molecular Ecology 28:544-567.

Galla, S.J., R. Moraga, L. Brown, S. Cleland, M.P. Hoeppner, R.F. Maloney, A. Richardson, L. Slater, A.W. Santure, and T.E. Steeves. 2020. A comparison of pedigree, genetic, and genomic estimates of relatedness for informing pairing decisions in two critically endangered birds: Implications for conservation breeding programmes worldwide. Evolutionary Applications. 00:1-18.

Hartl, D.L. and A.G. Clark. 2007. Principles of Population Genetics. Sinauer Associates, Inc., Sunderland, MA.

Hauser, L., M. Baird, R.A. Hilborn, L.W. Seeb and J.E. Seeb. 2011. An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. Molecular Ecology Resources. 11: 150-161.

Jones, O.R. and J. Wang. 2010. COLONY: a program for parentage and sibship interference from multilocus genotype data. Molecular Ecology Resources. 10(3): 551-555.

Lambert, Y. 2008. Why should we closely monitor fecundity in marine fish populations? Journal of Northwest Atlantic Fishery Science. 41:93-106.

Lauer, T.E., Shroyer, S.M., Kilpatrick, J.M., McComish, T.S., Allen, P.J. 2005. Yellow perch length–fecundity and length–egg size relationships in Indiana waters of Lake Michigan. North American Journal of Fisheries Management. 25:791-796.

Lemopoulos, A., J.M. Prokkola, S. Uusi-Heikkilä, A. Vasemägi, A. Huusko, P. Hyvärinen, M-L. Kolijonen, J. Koskiniemi, and A. Vainikka. 2019. Comparing RADseq and microsatellites for estimating genetic diversity and relatedness – implications for brown trout conservation. Ecology and Evolution 9:2106-2120.

Levine, B.A., M.R. Douglas, A.A. Yackel-Adams, B. Lardner, R.N. Reed, J.A. Savidge, and M.E. Douglas. 2019. Genomic pedigree reconstruction identifies predictors of mating and reporoductive success in an invasive vertebrate. Ecology and Evolution 9:11863-11877.

Marsh, P.C., T.E. Dowling, B.R. Kesner, T.F. Turner, and W.L. Minckley. 2015. Conservation to stem imminent extinction: The fight to save Razorback Sucker *Xyrauchen texanus* in Lake Mohave and its implications for species recovery. Copeia. 103: 141-156.

Miller, L.M. and A.R. Kapuscinski. 2003. Genetic Guidelines for Hatchery Supplementation Programs. Pages 329-355 *in* E.M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Modde, T. and D.B. Irving. Use of multiple spawning sites and season movement by Razorback Sucker in the Middle Green River, Utah. North American Journal of Fisheries Management. 18:318-326.

Peterson, B.K., J.N. Weber, E.H. Kay, H.S. Fisher and H.E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE. 7(5): e37135.

Purcell, S., B. Neale, K. Todd-Brown, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. de Bakker, M.J. Daly, and P.C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics. 81(3): 559-575.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact test and ecumenicism. Journal of Heredity 86:248-249. Available: http://genepop.curtin.edu.au/. (April 2012).

Rochette, N.C and J.M. Catchen. 2017. Deriving genotypes from RAD-seq short-read data using Stacks. Nature Protocols. 12: 2640.

Rochette, N.C., A.G. Rivera-Colón, and J.M. Catchen. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular Ecology 28:4737-4754.

Schwartz, M.K., G. Luikart, and R.S. Waples. 2010. Genetic monitoring as a promising tool for conservation and management. Trends in Ecology and Evolution 22(1):25-33.

Smouse, P.E. 2010. How many SNPs are enough? Molecular Ecology 19(7):1265-6.

Thrasher, D.J., B.G. Butcher, L. Campagna, M.S. Webster, and I.J. Lovette. 2018. Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: a proof of concept in a highly promiscuous bird. Molecular Ecology Resources. 18(5): 953-965.

U.S. Fish and Wildlife Service. 2005. An augmentation plan for Razorback Sucker in the San Juan River. U.S. Fish and Wildlife Service, Colorado River Fishery Project, Grand Junction, Colorado.

U.S. Fish and Wildlife Service. 2015. San Juan River Razorback Sucker *Xyrauchen texansu* & Colorado Pikeminnow *Ptychocheilus Lucius* population augmentation: 2014. U.S. Fish and Wildlife Service, New Mexico Fish and Wildlife Conservation Office, Albuquerque, New Mexico.

Waples, R.S., G. Luikart, J.R. Faulkner, and D.A. Tallmon. 2013. Simple life history trains explain key effective population size ratios across diverse taxa. Proceedings of the Royal Society B: Biological Sciences. 280: 20131339.

Waples, R.S., T. Antao, and G. Luikart. 2014. Effects of overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. Genetics. 197: 769-780.

Responses to Comments:

Bill Miller, Southern Ute Indian Tribe, BC member

How can the technical aspects of this SOW be improved?

This SOW is a continuation of the previous Nb work completed for the San Juan River. Can this method be applied to the 2019 genetic material if that material has not been processed? It seems the newer methodology may provide more definitive result.

Response:

Thank you for the comments. We are currently funded through the San Juan program to analyze the 2019 samples using microsatellite data. However, we are funded in FY2020 to analyze larval samples collected during 2020 to provide a direct comparison between the two data types. This will be done to assess impacts of the two genetic methods on estimation of Nb in the San Juan River, which will provide additional context to this study and others as we continue to monitor the reproductive contribution of endangered fishes in this system.

Program Office

How can the technical aspects of this SOW be improved?

It appears ASIR will not be sampling below the waterfall in FY2021. If this is the case, then you will likely need to reduce the number of samples you propose to process by half.

Response:

We have revised the SOW and associated budget to reflect this by removing the proposed below-waterfall sampling for FY2021.