Conservation genetics of Zuni bluehead sucker (*Catostomus discobolus yarrowi*) in New Mexico

Thomas F. Turner and Wade D. Wilson

Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico, 87131-0001

Contact: Thomas F. Turner
Museum of Southwestern Biology
MSC 03-2020
1 University of New Mexico
Albuquerque, New Mexico, 87131-0001
505-277-7541 voice
505-277-1361 FAX
turnert@unm.edu

Funding provided by New Mexico Department of Game and Fish
Contract #09-516-0000-00017

Final Report submitted to Stephanie Carman
Conservation Services Division
New Mexico Department of Game and Fish

Executive Summary

The purpose of this study was to conduct genetic analysis of Zuni bluehead sucker (*Catostomus discobolus yarrowi*) in the Zuni River Basin, New Mexico. Fin clips (n = 72) were collected from these localities in 2007 and 2008 by the New Mexico Department of Game and Fish from three localities in New Mexico: Tampico Springs (n = 25), Agua Remora (n = 22), and Rio Nutria (n = 25). Genotypes were determined at four genetic loci that included two protein-encoding mitochondrial genes cytochrome b (Schwemm & Dowling, unpublished) and ND4L (McPhee et al., 2008), a single microsatellite locus Xte4 (Turner et al., 2009), and the nuclear Recombination Activation Gene, RAG1 (Schwemm & Dowling, unpublished). Results from genetic analysis of these four loci are provided to assist conservation and management of this species in New Mexico waters.

Major Findings:

Genetic analysis indicated that individuals from Tampico Springs and Agua Remora exhibited lower levels of genetic diversity than Rio Nutria at all loci studied. Lowered diversity occurs in part because Tampico Springs and Agua Remora exhibited very few alleles/hapotypes that are of *C. plebeius* orgin. However, we recommend against introducing fish from the Rio Nutria intoTampico Springs and Agua Remora until it is clear why *C. plebeius* alleles are lacking in the upstream populations. Our genetic results from Rio Nutria are very similar to those reported by Schwemm and Dowling in that both data sets indicated relatively high diversity and relatively high frequency of alleles that are derived from *C. plebeius*. However, it cannot be determined from this study whether

C. plebeius alleles confer any benefit (in terms of viability) to the Rio Nutria population relative to Tampico Springs and Aqua Remora.

Tampico Springs and Agua Remora fishes are nearly genetically identical at the loci we studied and presumably shared common ancestry more recently that either did with Rio Nutria fishes. Based on these data, Agua Remora could be used as a donor population for Tampico Springs if needed, and vice versa. It is unclear why headwater populations differ genetically from the downstream Rio Nutria population. There could be a substantial barrier to gene flow, or alternatively, *C. plebeius* alleles may be purged by genetic drift or natural selection in the headwater populations.

Introduction

The Zuni bluehead sucker *Catostomus discobolus yarrowi* is hypothesized to have originated as a hybrid between *C. discobolus* and the Rio Grande sucker, *C. plebeius* (Smith et al., 1983). The geographic range of *C. d. yarrowi* is restricted to the Little Colorado River drainage on the western side of the Continental Divide, but *C. plebeius* occurs on the eastern side of the Continental Divide. Introgression is hypothesized to have occurred during the Pleistocene via stream capture between the eastern and western portions of the Divide (Smith et al., 1983). In New Mexico, the current distribution of *C. d. yarrowi* is limited to three localities and roughly two habitat types: Tampico Springs (headwater isolated spring), Agua Remora (headwater isolated spring), and Rio Nutria (canyon-bound river mainstem) with most fish found in the canyon-bound Rio Nutria (Carman, 2008). Tampico Springs and Agua Remora harbor relatively small numbers of fishes and may have suffered demographic bottlenecks over their history. It is also

possible that the Rio Nutria is a core population and that Agua Remora and Tampico are satellite populations that receive migrants periodically from downstream. Our genetic study has two main objectives: (1) to provide a baseline genetic survey of the headwater populations and compare these to the main population of Rio Nutria, and (2) to determine if the headwater populations of Tampico Springs and Agua Remora are genetically similar enough so that they can be used for reestablishment via brood stock being developed.

Materials and methods

Tissue Collection

Fin clip tissues used in this study were collected by New Mexico Game and Fish (Table 1). Fishes were temporarily incapacitated by electro-fishing, fin clipped, revived, and released at the point of capture. Fin clips were stored in 95% ethanol until extraction of whole genomic DNA. Genomic DNA was isolated using standard phenol/chloroform extraction protocols followed by precipitation in 100% EtOH.

PCR Amplification, Genetic Characterization, and Nucleotide sequencing

A 260 base pair (bp) region of the mtDNA cytochrome b (cyt-b) was amplified via Polymerase Chain Reaction (PCR) using ^{33}P radio-labeled forward primer LD 5'AGA TAA TGC AAC GCT AAC 3' and unlabeled reverse primer HB 5' TGA CCC CTC CTC ACA TTA 3' (Schwemm & Dowling, unpublished). Reactions were carried out in a total volume of 10 μl and contained 0.5 μM of each primer pair, combined with 10-100 ng of template DNA, 1x *Taq* buffer, 2 mM $MgCl_2$, 0.8 mM dNTP, and 0.375 units of *Taq* DNA polymerase (Promega). PCR was conducted on a T-Gradient thermocycler (Whatman

Biometra, Germany) for a total of 33 cycles. The initial cycle was 95° C for 120 sec, 50° C for 30 sec, and 72° C for 90 sec, followed by 32 shorter cycles (95° C for 30 sec, 50° C for 30 sec, and 72° C for 30 sec) with a final extension at 72° C for 10 min. PCR products were then diluted (1:1) in sequencing stop buffer and denatured at 95° C for 3 min and immediately cooled in an ice bath. Six μl of the PCR dilution mixture were loaded on a nondenaturing 5% polyacrylamide gel run which was run for 18 h at room temperature.

Single-stranded conformational polymorphism (SSCP) analysis (Sunnucks et al. 2000) of PCR fragments was employed, where practical, to visualize differences among individuals haplotypes. SSCP screening can significantly reduce the number of haplotypes that require direct nucleotide sequence because this approach can detect nucleotide sequence-level differences among individual haplotypes with high accuracy. Unique haplotypes identified by SSCP where checked by direct nucleotide sequencing as described below. After SSCP, two individuals per scored haplotype were selected for further analysis via nucleotide sequencing.

A 435 bp segment of the ND4L/ND4 mtDNA region (318 bp region of the mtDNA ND4L plus a 117 bp region of the mtDNA ND4) was amplified for all 72 individuals using the primers NAP-2 5' TGG AGC TTC TAC GTG(GA) GCT TT 3' (Arevalo et al., 1994) and ARG-BL 5' CAA GAC CCT TGA TTT CGG CTC 3' (Bielawski and Gold, 1996). Likewise, a 283 bp region of the Recombination Activation Gene (RAG1) was amplified in separate reactions for all 72 individuals with primers RAG1F 5' TGG CGG TCA RCC CTG GAC AA 3' and RAG1R 5' AAT CAC CAA CTA CCT CCA CA 3' (Schwemm & Dowling, unpublished). PCR reactions and cycling conditions for these regions were identical to those used for mtDNA cyt-b.

Both strands of cyt-b and ND4L/ND4 were sequenced using a BigDye 1.1 kit (Applied Biosystems, Darmstadt, Germany) and were edited using Sequencher version 4.0.5 (Gene Codes, Ann Arbor, Michigan). Multiple sequences aligned and adjusted by eye in Se-Al v2.0a11 (http://evolve.zoo.ox.ac.uk).

Microsatellite Amplification

A single microsatellite locus (Xte4, Turner et al., 2009) was amplified via PCR under the following conditions: each reaction contained 1x PCR buffer, 2mM MgCl2, 125µM dNTPS, 0.4µM of each primer and 0.375 units of TAQ polymerase and amplified with the following cycling conditions: one denaturation cycle of 93°C for 3 minutes, 30 cycles of 90°C for 20 sec, 50 °C for 20 sec and 72°C for 30 sec followed by a final extension step of 72°C for 30 minutes. PCR products (1 µL) were mixed with 10 µL of formamide and 0.3 µL of size standard (ABI HD400) and denatured at 93°C for 5 minutes. All samples were run on an ABI 3130 automated DNA sequencer and analyzed using GeneMapper software (Applied Biosystems).

Genetic Data Analysis

Allele frequencies, expected (H_E) and observed heterozygosity (H_O) and genetic divergence (F_{ST}) between all pairs of populations were calculated using GENALEX v6.1 (Peakall and Smouse, 2006). Statistical testing of F_{ST} was done by a bootstrap algorithm implemented in GENALEX. Phylogenetic analysis included selecting an appropriate nucleotide substitution model by likelihood ratio testing as implemented in the program ModelTest 3.7 (Posada and Crandall, 1998). Phylogenetic analyses and genealogical tree construction (minimum evolution, ME) were carried out using PAUP*, version 4.0b10 (Swofford, 2001).

Results and Discussion

Two unique haplotypes were identified each of the two mtDNA regions examined

for 72 *C. d. yarrowi* individuals (Table 2, Appendix A). Of the two cyt-b haplotypes, one was identical to the "F" allele of Schwemm and Dowling (unpublished). In their study, this haplotype was unique to the Canyon de Chelly sublineage of *C. discobolus*. In our study, this allele was present in all populations, and was the only haplotype found in Tampico Springs and Agua Remora (Table 2 and 3). The other cyt-b haplotype "JJ" is similar to the "J" haplotype of Schwemm and Dowling, but differs by a single nucleotide at the 5' end of the sequence (Appendix A). The "J" haplotype was found to be part of the *C. plebeius* lineage (Schwemm and Dowling, unpublished). This haplotype ("JJ") was only found in individuals from Rio Nutria and occurred in 24.0% of the individuals genotyped with the other 76.0% having the "JJ" haplotype (Table 3). The two ND4 haplotypes were "D" and "P" which indicated the species that they were closely related to, *discobolus* (GenBank # EF370432, see McPhee et al., 2008) and *plebeius* (GenBank # EF370431, AY494958-AY494964, see McPhee et al., 2008) respectively (Figure 1). Haplotypes identified as "D" at the mtDNA ND4 gene always had the F haplotype in cyt-b region (consistent with complete linkage of the mtDNA genome, as expected) with "D" occurring in Tampico Springs, Agua Remora, and Rio Nutria. Haplotypes identified as "P" at the mtDNA ND4 gene also always had haplotype "JJ" at cyt-b. These haplotypes only occurred in Rio Nutria (Table 3). This is the first ND4 sequence data reported for *C. d. yarrowi* and cannot be compared directly to Schwemm and Dowling. Evolutionary affinities of these haplotypes were thus checked by verifying their phylogenetic placement compared to *C. discobolus* and *C. plebeius* data presented in McPhee et al. (2008).

Two nuclear regions were also genotyped and sequenced including one

microsatellite (Xte4, Turner et al., 2009; called "2H2" by Schwemm and Dowling). In the 72 individuals genotyped for Xte4, two alleles were found (Table 3 and 4, Appendix A), one at 286 bp and one at 288 bp, which correspond to the "B" and "A" alleles (respectively) of Schwemm and Dowling (unpublished). Allele "C" was not found in this study but was previously found to occur in low frequency (4/112 individuals) in Rio Nutria (Schwemm and Dowling, unpublished). It is possible that this allele was present in the Rio Nutria, but was missed by us because we evaluated fewer fish (n = 25) than Schwemm and Dowling (n =112). Allelic frequencies of "A" (32%) and "B" (68%) in Rio Nutria were lower and higher (respectively) than Schwemm and Dowling (52 and 45%). All three alleles ("A, B, C") are found in *C. discobolus* , but allele B is in high frequency in *C. plebeius*. The headwater population Tampico Springs was fixed for the "A" (288) allele, and Agua Remora nearly fixed (frequency = 95.5%, Table 3) with two individuals having the "B" (286) allele (Appendix A). This can also be seen in the observed heterozygosity (H_O) values with Tampico Springs lacking any heterozygosity (0.000), Agua Remora low (H_O = 0.091) and Rio Nutria being high (H_O = 0.640) (Table 2). Pairwise F_{ST} measures for Xte4 indicated that Tampico Springs and Agua Remora, but both were different from Rio Nutria (Table 4).

Three nuclear RAG alleles were found to occur in the three populations surveyed (Table 2 and 3, Appendix A) and correspond to the "A" "B" and "C" alleles of Schwemm and Dowling (unpublished). Again, the headwater populations of Tampico Springs and Agua Remora had lower diversity than Rio Nutria and only had the "A" and "B" alleles, allele "C" was lacking in these populations. However, all three alleles were found to occur in Rio Nutria (Table 2 and 3, Appendix A**)**. The "A" and "B" alleles were found to

be part of the *C. discobolus* lineage and was lacking in the *C. plebeius* lineage, however the "C" allele was found to occur in both the *C. discobolus* and *C. plebeius* lineages. Rio Nutria allelic frequencies for all three alleles ("A" = 26.0%, "B" = 38% and "C" = 36%) were similar to those of Schwemm and Dowling (A = 24%, B = 34% and C = 37%). The fourth allele found by Schwemm and Dowling ("D") was not found in the current survey, but had a low frequency (0.08%) in their study and may not have been detected in our study, again due to a lowered sample size. Pairwise F_{ST} measures for RAG indicate that allele frequencies of Tampico Springs and Agua Remora are not significantly different from one another, but that both are significantly different from Rio Nutria (Table 4).

In summary, analysis of the four genetic markers (two mitochondrial, one nuclear coding gene and one nuclear DNA microsatellite) indicates that individuals from Tampico Springs and Agua Remora exhibit substantially lowered genetic diversity relative to Rio Nutria. This is because Tampico Springs and Agua Remora have few alleles/hapotypes that are of *C. plebeius* origin. We identified fish with identical genotypes at these four loci in all three populations. Because fin clips were taken and fish released back into the wild, it is possible that some individuals were genotyped more than once. It is probably more likely that genetically identical individuals for these four slowly evolving loci are present in the population due to small population size and genetic drift. Comparisons of allele frequencies indicated that Tampico Springs and Agua Remora are nearly identical (at these genetic loci) and presumably share common ancestry more recently that either does with Rio Nutria. When compared to the broader data set of Schwemm and Dowling, nearly all the alleles in the headwater populations (Tampico Springs and Agua Remora) were present in *C. discobolus* lineages. *C. plebeius*

alleles were either present in very low frequency or absent completely. Based on this data set, Agua Remora could be used as a donor population for Tampico Springs if needed, and vice versa, but it is not clear whether the headwater populations have recently exchanged migrants with Rio Nutria.

Our results, combined with those of Schwemm and Dowling, suggest a complex mosaic of processes, including multiple stream capture and cross-basin gene flow events, hybridization, and persistence of small, genetically bottlenecked populations shaped the evolutionary history of *C. d. yarrowi*. For the New Mexico conservation effort, future work should seek to understand why Tampico and Agua Remora populations contain little evidence for recent introgression with *C. plebeius.* One possibility is that there is a significant barrier to upstream gene flow from the Rio Nutria canyon-bound population below the confluence of the Tampico Draw and upper Rio Nutria. Alternatively, *plebeius* alleles may be purged via genetic drift (especially because plebeius alleles occur in low frequency in the Rio Nutria) or natural selection in Tampico and Agua Remora populations following gene flow from the Rio Nutria canyon-bound areas downstream. Routine genetic monitoring may help distinguish these possibilities, but additional ecological studies of fish movement may be required to understand patterns of connectivity among these populations. For example, otolith microchemistry or mark-recapture approaches could verify migration between headwater localities, or downstream reaches and upstream headwater springs.

Literature Cited

Arevalo, E, SK Davis and JW Sites, Jr. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Syst. Biol. 43: 387-418.

Bielawski, JP and JR Gold. 1987. Uneven synonymous substitution rates within and between two protein-coding mitochondrial genes. Mol. Biol. Evol. 13: 889–892.

Carman, S. 2008. Zuni bluehead sucker monitoring and conservation efforts 2008. New Mexico Department of Game and Fish.

McPhee, MV, MJ Osborne, and TF Turner. 2008. Genetic diversity, population structure and demographic history of the Rio Grande sucker, *Catostomus* (*Pantosteus*) *plebeius,* in New Mexico. Copeia 2008:191-199.

Peakall R, PE Smouse. 2006. GenAlEx 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6: 288-295.

Posada D, KA Crandall .1998. Modeltest: Testing the model of DNA substitution. Bioinformatics 14: 817-818.

Smith, GR, JG Hall, RK Koehn and DJ Innes. 1983. Taxonomic relationships of the Zuni mountain sucker, *Catostomus discobolus yarrowi*. Copeia 183: 37-48

Sunnucks P, ACC Wilson, LB Beheregaray, K Zenger, J French, AC Taylor. 2000. SSCP is not so difficult: The application and utility of single-stranded conformational polymorphism in evolutionary biology and molecular ecology. Mol. Ecol. 9: 1699-1710.

Swofford DL. 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other

Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Turner, TF, TE Dowling, MV McPhee, C Secor, RE Broughton, and JR Gold. 2009. Polymorphic microsatellite primers for the endangered sucker, *Xyrauchen texanus* (Catostomidae), are useful for hybridization studies of other catostomids. Conserv. Genet. 10: 551-553. DOI 10.1007/s10592-008-9563-8.

Wilson, WD, and TF Turner. 2009. Phylogenetic analysis of the Pacific cutthroat trout (*Oncorhynchus clarki* ssp.: Salmonidae) based on partial mtDNA ND4 sequences: A closer look at the highly fragmented inland species. Mol. Phy. Evol. 52: 406-415.

Table 1. Fin clips of Zuni bluehead sucker (*Catostomus discobolus yarrow*), obtained by Stephanie Carman of the New Mexico Department of Game and Fish, and formally accessioned into the fish collection at the Museum of Southwestern Biology. All fin clip samples (total = 72) reported in this table were characterized for genetic diversity as indicated in the body of the report. Individual genotype data taken at different times were pooled by locality prior to statistical analysis.

Fin Clips	Date Collected	State	County	Locality	UTM
25	7/2/2007	New Mexico	McKinley	Tampico Springs at Silva	12S 726366E, 3913453N
12	8/19/2007	New Mexico	McKinley	Agua Remora at sect 26 FS	12S 727222E, 3912095N
17	8/20/2007	New Mexico	McKinley	Rio Nutria, below confluence with Tampico Draw	12S 724161E, 3908546N
10	8/25/2008	New Mexico	McKinley	Agua Remora at sect 26 FS	12S 727222E, 3912095N
8	8/28/2008	New Mexico	McKinley	Rio Nutria, below confluence with Tampico Draw	12S 724161E, 3908546N

Table 2. Genetic summary for the three populations Tampico Springs (TS), Agua Remora (AR) and Rio Nutria (RN). Abbreviations are N_A = number of alleles, H_O = observed gene diversity (heterozygosity), H_E = expected heterozygosity.

Locus	Statistic	TS (25)	AR (22)	RN (25)
Cyt b	N_A	1	1	2
ND4L/ND4	N_A	1	1	2
RAG1	N_A	2	2	3
	H_O	0.360	0.455	0.800
	H_E	0.385	0.463	0.685
Xte4	N_A	1	2	2
	H_O	0.000	0.091	0.640
	H_E	0.000	0.087	0.435

Table 3. Allelic frequencies for the three *C. d. yarrowi* populations Tampico Springs (TS), Agua Remora (AR) and Rio Nutria (RN). Microsatellite alleles A and B are 2H2 from Schwemm and Dowling (unpublished) and correspond to the current Xte4 designation of alleles 286 and 288.

Locus	Allele	TS (25)	AR (22)	RN (25)
Cyt b	F	1.000	1.000	0.760
	JJ	0.000	0.000	0.240
ND4L/ND4	D	1.000	1.000	0.760
	P	0.000	0.000	0.240
Xte4 (2H2)	A (288)	1.000	0.955	0.700
	B (286)	0.000	0.045	0.300
RAG1	A	0.260	0.364	0.260
	B	0.740	0.636	0.380
	C	0.000	0.000	0.360

Table 4. Pairwise F_{st} measures between the three *C. b. yarrowi* populations Tampico Springs (TS), Agua Remora (AR) and Rio Nutria (RN). Values above diagonal indicate *p* values (* indicates significant differences).

Gene	Population			
Xte4		**TS**	**AR**	**RN**
	TS	-	0.460	0.010*
	AR	0.002	-	0.010*
	RN	0.184	0.137	-
RAG1		**TS**	**AR**	**RN**
	TS	-	0.210	0.010*
	AR	0.028	-	0.010*
	RN	0.313	0.207	-

Figure 1. Unrooted phylogenetic tree (minimum evolution, Tamura-Nei + I + Γ) based on a 435 bp segment of the mtDNA ND4L/ND4.