SAN JUAN RIVER RECOVERY IMPLEMENTATION PROGRAM

FY2020 ANNUAL REPORT

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

Bureau of Reclamation Agreement Number:
R20PG00102

Principal Investigator:
Melody Saltzgiver and Steven Mussmann, Geneticist and Supervisory Geneticist
US Fish and Wildlife
Southwestern Native Aquatic Resources and Recovery Center
7116 Hatchery Road
Dexter, NM, 88230
Phone: (575) 734-5910
Fax: (575) 734-6130
melody_saltzgiver@fws.gov, steven_mussmann@fws.gov

Executive Summary:

Many fish species have exhibited dramatic population declines due to a myriad of human-induced threats. A majority of the native freshwater fishes of the southwestern United States are imperiled. To compensate for their decline, captive propagation and stocking of hatchery-reared individuals have become common management strategies to sustain many species. Population size is a central parameter for conservation, so monitoring abundance is vital for understanding the conservation status of species and for assessing management efforts. While it is important to look at population size, it is also important to examine how it relates to reproductive success, which must also be evaluated if long-term survival is to be achieved. Specific information about reproductive success is extremely difficult to quantify from field studies, which can make it difficult to relate demographic trends to the recovery of the species. Many studies have demonstrated that genetic monitoring of reintroduction programs assists with demographic viability assessments.

In the San Juan River, a genetic monitoring method examining the effective number of breeders (N_b) has been implemented successfully for the Colorado Pikeminnow and Razorback Sucker from 2009 through 2020. These estimates were calculated from microsatellite data through 2020, when N_b estimates were also calculated from single nucleotide polymorphisms (SNPs) for the first time. In contrast to microsatellites, SNPs are expected to yield more precise estimates. The double digest restriction-site associated DNA (ddRAD) sequencing method of SNP discovery was used for both Colorado Pikeminnow and Razorback Sucker. This method yielded a robust number of SNPs for Razorback Sucker (N = 18,335). The 2020 calculation of N_b for Razorback Sucker using microsatellites was 131.7 while the N_b calculated from SNPs was very similar at 129.1. The 2020 calculation of N_b for Colorado Pikeminnow using microsatellites was 52.1, but the ddRAD method for SNP discovery yielded an order of magnitude fewer SNPs than generated in the Razorbacks (N = 2,227). Consequently, estimates will not be calculated for Colorado Pikeminnow using SNPs until the cause of this discrepancy is determined. However, this was most likely due to the extreme relatedness of the individuals sampled. Future ddRAD implementations for Colorado Pikeminnow will be modified to capture a greater proportion of the genome, thus identifying more SNPs.

Past N_b estimates were compared to adult population size estimates to determine the levels of contribution from the population to reproduction. These comparisons could not be implemented in 2020 due to COVID-19 restrictions hindering field sampling that is used in calculating the adult population estimates. Annual N_b estimates for Colorado Pikeminnow have been variable, ranging from 4 to 58 over 8 different years, with the overall trend showing an increase over time. In fact, 2020 was the second highest recorded (N_b = 52.1). For the years that population sizes could be estimated for Colorado Pikeminnow, the ratio of reproductive efforts had an average of 20.5% (N = 7, range = 4-40%) of the adult population contributing to the larvae collected. Razorback Sucker had N_b estimates ranging from 74 to 233 over 12 years, but the 2020 estimates (microsatellite N_b = 131.7 and SNP N_b = 129.1) were lower than observed in recent years. This was most likely due to COVID-19 restrictions truncating the field sampling season for Razorback Sucker. Despite Razorback Sucker having higher N_b than Colorado Pikeminnow, the mean ratio to estimated population size was only 3.8% (N = 7, range = 2.2-7.6%) of the adult population. These results exhibit different trends for Colorado Pikeminnow and Razorback Sucker. Colorado Pikeminnow shows high annual variability in reproductive output that appears to be slightly increasing over time. Razorback Sucker is relatively consistent, but with low overall contributions to larval cohorts, which suggests a reproductive bottleneck is occurring at spawning or during early life stages. With these trends in mind, it is important to continue estimating N_b so that conservation activities can be assessed in addition to environmental factors affecting reproductive output of these two endangered fish species in the San Juan River.

Tasks:

1. Quantify the effective number of breeders (N_b) for Razorback Sucker and Colorado Pikeminnow collected in the San Juan River in 2020 using both microsatellites and single nucleotide polymorphisms (SNPs).
2. Examine the N_b estimates for Razorback Sucker and Colorado Pikeminnow larvae captured below the Piute Farms Waterfall in 2020.
3. Examine the spatial distribution of full-siblings for both Razorback Sucker and Colorado Pikeminnow in 2020.

Introduction

Many species of fish have exhibited dramatic population declines over the last two centuries due to a myriad of human-induced threats (Bruton 1995; Ricciardi and Rasmussen 1999; Su et al. 2021). This is especially true for native freshwater fishes of the southwestern United States including the Razorback Sucker (*Xyrauchen texanus*) and Colorado Pikeminnow (*Ptychocheilus lucius*) (Miller 1961; Minckley and Deacon 1968; Clarkson et al. 2005). Population declines in the southwest can be attributed to habitat loss and degradation, and the introduction of non-native species (Miller 1961; Minckley and Rinne 1991; Minckley and Marsh 2009). In response, conservation actions have included restriction of harvest, habitat restoration, non-native species control, and artificial propagation and repatriation programs (Clarkson et al. 2005, Rasmussen et al. 2009). Demographic monitoring is a useful tool for understanding the effectiveness of propagation and repatriation programs. It is often a key objective of monitoring programs which are vital for understanding the conservation status of species, the regulating effects of biotic and abiotic factors, and for the assessment of management efforts (Lindenmayer et al. 2020). Monitoring reveals detailed information on reintroduction success and population development (La Haye et al. 2017). Studies have demonstrated that genetic monitoring of

reintroduction programs assists with demographic viability assessments within an adaptive management framework (Hartl and Clark 2007; Sard et al. 2016).

The number of fish in a population is not the only metric that needs to be examined when working toward recovery. Knowing how many of those recruiting fish genetically contribute to the next generation is also vitally important (Miller and Kapuscinski 2003; Le Vay et al. 2007; Rasmussen et al. 2009). Successful recovery requires a significant portion of the restored population to reproduce annually to both increase population sizes and ensure the maintenance of genetic diversity. Therefore, quantifying the relative number of individuals that contribute to annual reproductive efforts allows for informed management decisions to aid in their recovery. Estimating the effective number of breeders (N_b) using genetic analyses is an effective tool for estimating reproductive output. N_b refers to the effective number of breeders in a single breeding season (Waples 2005). For long-lived, highly fecund, iteroparous species with overlapping generations, like Colorado Pikeminnow and Razorback Sucker, N_b is an extremely useful metric for understanding population-level spawning success when fish have well-defined seasonal reproductive bouts (Waples et al. 2013; Waples et al. 2014; Wang et al. 2016).

Monitoring short-term fluctuations in N_b, as well as its ratio to adult census population size (N_b/N_c), provides insight into population demography and informs conservation programs about long-term loss of evolutionary potential in wild populations (Perrier et al. 2016). Smaller ratios can potentially be caused by factors such as nonrandom mating, uneven sex-ratios, variance in individual reproductive success, and fluctuating N_c over generations (Frankham 1995; Kalinowski and Waples 2002; Hedrick 2005). A single cohort N_b estimate can quantify the number of individuals that contributed to a given cohort (Waples et al. 2014). Obtaining N_b estimates for the endangered fishes of the San Juan River will provide a look into how natural, population-level reproduction is affecting progress towards recovery in the San Juan River.

Presently, the most common genetic markers applied for assessing genetic diversity, relatedness, and population structure are single nucleotide polymorphisms (SNPs) and microsatellites (Lemopoulos et al. 2019, Camacho-Sanchez 2020). Both types of markers can be used to estimate N_b, but the more precise estimates are most likely obtained from SNPs in comparison to genetic data microsatellites (Gala et al. 2020, Sunde et al. 2020).

Methods

Sample Collection

Larval fish sampling was intended to be conducted along a 180-mile reach of the San Juan River between its confluence with the Animas River in New Mexico and Piute Farms Waterfall in Utah. Those collections were truncated during the 2020 field season due to COVID-19 restrictions. Colorado Pikeminnow were collected from 14 miles downstream [lake mile (LM) 40 to LM 54] of Piute Farms Waterfall to 77.0 miles upstream of the waterfall near Mexican Hat. Razorback Sucker larval samples were also restricted within two different sampling locations. The first location was from river mile (RM) 166.6 near Public Service Company of New Mexico (PNM) weir up to RM 180.6 near the Animas River confluence. The second collection location for Razorback Sucker spanned from RM 0.0 at Piute Farms Waterfall up to RM 77.0 near Mexican Hat. In 2020, up to 120 Colorado Pikeminnow and Razorback Sucker were also intended to be subsampled from the below-waterfall collections. Only five Colorado Pikeminnow larvae and no Razorback Sucker larvae were collected below the waterfall. There were 120 Colorado Pikeminnow sampled from the larvae collected above the waterfall. All Razorback Sucker larvae collected (N = 177) were sampled for genetic testing except one which was left as a specimen voucher at the Museum of Southwest Biology (MSB), Albuquerque, NM Biology. Tissue was

subsampled from the posterior portion of each larval specimen while the anterior portion was retained at MSB for potential future otolith studies.

Larval collections have been preserved in 95% ethanol since 2009, making those years suitable for genetic sampling. Annual N_b estimates were obtained using 2,248 samples (N = 1,429 Razorback Sucker and N = 819 Colorado Pikeminnow) representing the spatial and temporal distribution of larval sampling efforts in the San Juan River (Figure 1). Larval Razorback Sucker were collected during much of the sampling season for most years except 2020 which was truncated due to COVID-19 restrictions. For Razorback Sucker, early larval stages (e.g., protolarvae to mesolarvae) were still targeted during larval collections under the assumption that these individuals were from recent spawning events. Conversely, Colorado Pikeminnow were only collected later in the sampling season, which made it relatively easy to have those captured individuals reflect the entire seasonal spawning period of adults. To ensure N_b estimates in Colorado Pikeminnow were not artificially low due to limited spatial representation of samples, rare collections were targeted while sites with high larval densities were proportionally sampled for analysis.

Genetic Sampling

Genomic DNA was extracted using NucleoSpin Tissue Kits (Macherey-Nagel, Duren, Germany). Twenty-one microsatellite loci were used for Razorback Sucker (Dowling et al. 2011; Wilson 2012) and twenty-five microsatellite loci were used to amplify Colorado Pikeminnow DNA (Martin et al. 2015). Microsatellite amplification consisted of 10 µl reactions containing the following: 1 µl DNA, 3 µl Qiagen Multiplex Master Mix® (Qiagen, Valencia, CA, USA), 0.2 µl of both forward and reverse primers, and 5.6 µl of nuclease-free water. Forward primers were labeled with one of four fluorescent dyes (6-FAM, PET, NED, VIC; Applied Biosystems, Inc., Foster City, CA, USA). Amplification of all samples consisted of an initial denaturing step of 95°C for 15 min, followed by 35 cycles of 94°C for 60 s, 56°C for 45 s, and 72°C for 60 s, with a final extension of 10 min at 70°C. Amplified products were processed on an ABI 3500xl Genetic Analyzer. Composite genotypes for individual fish were compiled using GeneMapper™ 4.0 software (Applied Biosystems, Inc., Foster City, CA, USA). An independent researcher reviewed all genotypes to ensure accuracy. A second researcher also re-extracted and re-amplified 10% of samples for quality assurance/quality control assessment.

Library preparation for SNP detection followed a double digest restriction site-associated DNA (ddRAD) protocol (Peterson et al. 2012). Restriction digests of up to 1 µg genomic DNA per sample were performed in 50 µl reactions containing 5 µl New England BioLabs CutSmart Buffer and 20 units each of PstI and MspI restriction enzymes. Samples were digested at 37°C for 18-24 hours then purified using Agencourt AMPure XP beads (Beckman Coulter, Inc.). Barcoded samples (25-100 ng DNA each) were pooled in sets of 48 following Illumina adapter ligation, then size-selected at 329-429 bp for Colorado Pikeminnow and 353-403 bp for Razorback Sucker (Bangs et al. 2018; Chafin et al. 2018) using the Pippin Prep System (Sage Science). Size-selected DNA was subjected to 12 cycles of PCR amplification using Phusion high-fidelity DNA polymerase (New England BioLabs) following manufacturer protocols. Quality checks and quantification of final libraries were performed using qPCR (NEBNext® Library Quant Kit for Illumina®; New England BioLabs) and the Agilent 2100 Bioanalyzer. All indexed libraries were pooled in equimolar amounts in sets of four to six per lane and sent for 150 bp single-end sequencing (Nextseq 500, Whitney Genetics lab, Midwest Fisheries Center, Onalaska, WI).

Data Analysis – SNPs

The quality of all ddRAD data was assessed using FastQC v11.7 (Andrews 2010). Components of the STACKS v2.6 (Rochette et al. 2019) pipeline (*process_radtags; denovo_map.pl; populations*) were used for processing raw ddRAD data using the methods of Rochette & Catchen (2017). The Colorado Pikeminnow ddRAD data were first demultiplexed in *process_radtags* to remove sequences with low quality scores (Phred < 10) and uncalled bases, trim adapter sequences, rescue sequences containing sequencing errors in barcodes, and trim sequences to 130 bp to remove low quality bases at the 3' end of the sequences. The cleaned data were assembled in *denovo_map.pl* using default settings. The *populations* module was applied to retain loci appearing in at least 80% of samples, a single SNP per locus, and SNPs with a minimum minor allele frequency of 0.01.

Fastp (Chen et al. 2018) was used to remove low-quality base calls present in the restriction cut site of some Illumina sequences for Razorback Sucker. These low-quality bases were removed from all Razorback Sucker sequences by trimming the first five bases from the 5' end of each sequence. Fastp was also used to remove adapter sequences and polyG tails, trim sequences to a maximum length of 100 bp, and discard sequences <100 bp after trimming. The quality filtering feature was disabled in fastp so that *process_radtags* filtering options could later be applied, similar to those used for Colorado Pikeminnow.

The same components of the STACKS pipeline that were utilized on the Colorado Pikeminnow were used for processing the cleaned Razorback Sucker ddRAD data. The trimmed sequences were run in *process_radtags* to remove sequences with low quality scores (Phred < 10) and uncalled bases, trim adapter sequences, rescue sequences containing sequencing errors in barcodes, and trim sequences to 130 bp to remove low quality bases at the 3' end of the sequences. The option to check for intact RAD cut sites was disabled because these were removed in fastp. Data were assembled using *denovo_map.pl* with default settings. The *populations* module was used to retain loci appearing in at least 50% of individuals for each population with just one SNP per locus, and those appearing in at least two populations because we wanted SNP loci shared among pairs of species to facilitate hybrid detection among Razorback, Bluehead (*Catostomus discobolus*) and Flannelmouth Sucker (*C. latipinnis*). A minimum minor allele frequency of 0.01 was also applied.

AdmixPipe v3.0 (Mussmann et al. 2020) was used to quantify species ancestry for each sample using the program Admixture (Alexander et al. 2009). Genetic clustering values (*K*) ranged from 2 to 4, with 20 replicates performed at each *K*. AdmixPipe results were evaluated in Clumpak (Kopelman et al. 2015) to determine the optimal *K* value and identify hybrid individuals. The *populations* module was then rerun with the hybrid individuals removed and using the aforementioned filtering parameters for Colorado Pikeminnow.

For both species, N_b was determined using NeEstimator v2.1 (Do et al. 2014), which estimates the effective number of breeders using multilocus diploid genotypes from population samples. The number of SNP loci used to calculate N_b for Razorback Sucker was 18,335 and for Colorado Pikeminnow the number of microsatellite loci used was 25. The updated version of the linkage disequilibrium method (Waples 2006) was used for all estimates. The criteria for excluding rare alleles (P_{crit}) with frequency <0.02 was applied (Waples and Do 2010).

All identified SNPs (N = 18,335) were used for pedigree reconstruction in Sequoia v2.3.5 (Huisman 2017). Sequoia accounts for genotyping errors and overlapping or discrete generations, with or without inbreeding and any proportion of genotyped parents. This software uses a maximum likelihood method to estimate relationships among offspring belonging to a single cohort by identifying full and half-sibling families.

Data Analysis - Microsatellites

Departures from Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD) were tested using Genepop version 4.7.5 (Raymond and Rousset 1995). Additionally, alpha (0.05) was adjusted for multiple comparisons using the Benjamini and Yekutieli (2001) method of false discovery rate (Narum 2006). The inclusion of non-target species and/or hybrids has been shown to affect N_b estimates (Schrey et al. 2011), therefore, any potential misidentified larvae were classified using a Bayesian clustering method implemented in Structure v2.3.2 (Pritchard et al. 2000; Falush et al. 2003) for microsatellite data. Structure assigns individuals to a user-specified number of genetic clusters with no *a priori* definition required. Approximately 80 to 160 individuals from each native sucker species (Bluehead Sucker, Flannelmouth Sucker and Razorback Sucker) were included in each Structure run to identify misidentified fish. The admixture model assuming gene flow among populations with correlated allele frequencies was used with twenty replicates performed for each *K* (the number of genetic clusters), with *K* assumed to be between 2 and 4 genetic clusters. For each replicate there was a burn-in of 100,000 Markov chain Monte Carlo (MCMC) iterations followed by 1,000,000 iterations of data collection. Species were identified using the q-values estimated in Structure and individuals were assigned to species at q-value > 0.90. The Structure results were evaluated in Clumpak using the ΔK method (Evanno et al. 2005) to determine the optimal number of genetic clusters.

NewHybrids, a Bayesian assignment method that classifies individuals to hybrid categories, was used to classify Razorback Sucker larvae into discrete hybrid categories (Anderson and Thompson 2002). These categories were 1) non-admixed Flannelmouth Sucker, 2) non-admixed Razorback Sucker, 3) first generation (F1) hybrids, 4) second generation (F2) hybrids, 5) F1 by Flannelmouth Sucker hybrid (Flannelmouth Backcross), and 6) F1 by Razorback Sucker hybrids (Razorback backcross). Non-admixed adults from the San Juan River which had Structure q-values > 0.97 were assigned as reference genetic profiles for Razorback Sucker and Flannelmouth Sucker using the 'z' option. We applied the Jeffreys prior for both π and θ and performed a burn-in of 100,000 MCMC iterations followed by 1,000,000 iterations of data collection. Any individuals identified as a F1, F2, or F2 backcross were removed for N_b estimates.

The effective number of breeders (N_b) was estimated for each species in NeEstimator v2.1 using the updated version of the linkage disequilibrium method (Waples 2006). Rare alleles (P_{crit}) with frequency <0.02 were excluded from analysis (Waples and Do 2010). Analyses were conducted separately for each year (i.e., cohort) to estimate N_b. In the years when adequate samples were collected below the waterfall, separate N_b estimates were calculated for above the waterfall, below the waterfall, and combined.

Pedant version 1.0 was used to estimate allelic dropout (E_1) and false allele error rates (E_2) (Johnson and Haydon 2007). These error rates are necessary to run Colony version 2.0.4.0 (Jones and 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 incorporating genotyping errors (i.e., E_1 and E_2) and allowing for inferences related to the mating strategy of the organism. Analyses were conducted separately for each year (i.e., cohort) to estimate the number of adults that contributed to at least one offspring, the number of sampled offspring produced by each parent, and the number of parental pairs. For all species, male and female polygamy was assumed and parameter settings (i.e., dioecious, diploid, inbreeding, long run length, full-likelihood with high likelihood precision, no sibship prior, and updated allele frequencies) were maintained across years. Low error rates were estimated at all

loci (dropout rate E_1 ranged 0.01-0.07 and false allele error rate E_2 0.02-0.05) and were provided independently for each locus (results not reported here).

Adult census estimates (i.e., population sizes) for both Razorback Sucker (2011-2016; 2019) and Colorado Pikeminnow (2011-2012; 2014-2019) were obtained in order to place N_b estimates in the context of demographic and population structures (Schleicher 2016; SJRRIP 2017; Schleicher et al. 2019). The ratio of N_b to adult census size provides an estimate of the relative annual reproductive contribution of adults. Adult Colorado Pikeminnow (age-5+; > 417 mm TL) and Razorback Sucker (age-4+; ≥ 400 mm TL) census sizes were estimated using pass-specific detection probabilities from closed population mark-recapture models (SJRRIP 2017). The relative proportion of spawning adults was estimated by comparing the means (and 95% CI) obtained for N_b and adult census estimates.

Results

Colorado Pikeminnow

The number of microsatellite alleles among San Juan River Colorado Pikeminnow cohorts varied from 4 to 29 alleles per locus with an average of 12 alleles per locus across all 25 loci. There were 2,227 SNP loci identified by Stacks using the ddRAD method of discovery. This was a surprisingly low number of SNPs, with similar ddRAD applications recovering over 9,000 loci for other species (Dacosta and Sorenson 2014). Due to the reduced number of SNPs, all calculations for this report were done with microsatellite data.

Overall, mean N_b estimates averaged 31 across years with a range of 4 to 58 (Figure 2). The lowest mean estimate was observed in 2015 (N_b = 4) and the highest in 2019 (N_b = 58). The 2020 N_b estimate was slightly below the 2019 estimate at N_b = 52. Mean adult census estimates within the San Juan River varied among years with as few as 19 individuals estimated in 2012 to as many as 617 in 2018 (Figure 3). Adult census data were not available for 2013 due to limited numbers of recaptures, and were also unavailable for 2020 because of COVID-19 sampling constraints. Large variation around these means was observed across years with upper estimates as high as 1,753 adults in 2018 to as few as seven adults estimated in 2012.

The relative proportion of spawning adult Colorado Pikeminnow in the San Juan River was variable across years. The lowest annual contribution of 4% coincided with the lowest N_b (N_b = 4). Conversely, the highest N_b estimate coincided with the highest adult contribution of 40.8% (N_b = 58). A ratio could not be calculated for 2020. Overall, a mean proportion of 20.5% adult Colorado Pikeminnow contributed to the larvae collected in the San Juan River across all years (Table 1).

The estimated number of spawning Colorado Pikeminnow adults and the number of offspring produced by each parent varied among years (Figure 4). Comparison of these measures provides a way to visualize variance in reproductive success by examining the number of larval fish each parent produced, with high variation in offspring contribution having the ability to lower N_b estimates. Unsurprisingly, years with low N_b estimates coincided with low estimates of spawning adults. Indeed, in 2015 the lowest number of spawning adults was observed with 7 parents. Conversely, the largest number of spawning adults was observed in 2019 (>100), which coincided with the highest N_b estimate. Individual adult contribution of offspring varied across years with some adults contributing as many as 30 of the sampled offspring, while most had limited contribution. This variance in reproductive success within years resulted in a decrease in N_b estimates. Nonetheless, these estimates were similar most years. For example, adult contribution of offspring was relatively even in 2016 and 2019 with similar estimates of N_b and number of spawning adults. Conversely, in 2014 individual contribution of offspring was more variable, which contributed to a lower N_b estimate.

The number of Colorado Pikeminnow parental pairs varied among years (Figure 5 and 6) with estimates of low and high parental pairs coinciding with low or high N_b estimates and number of spawning adults. Across all years, the number of parental pairs was greater than the estimated number of spawning adults, suggesting a polygamous mating strategy for Colorado Pikeminnow. Full-sibling relationships within parental pairs were also reconstructed. All larval collections and full-sibling relationships that were determined in Colony were plotted across river mile in order to visualize the extent of larval drift (Figure 5 and 6). There was evidence of extensive larval drift in multiple years with full-siblings collected up to 100 miles apart. In terms of distribution of larvae, in years with more larvae in the system, fish were more evenly distributed throughout. These results also suggest at least some pairs are spawning relatively high in the system because there is evidence for relatively large drift between full-siblings.

Razorback Sucker

The number of microsatellite alleles among Razorback Sucker larvae varied from 3 to 43 alleles per locus with an average of 22 alleles per locus across all 21 loci. The initial analysis of the larval microsatellite data (N = 171) from Structure and NewHybrids detected 46 hybrid larvae. The hybrids accounted for 26.9% of samples (Table 2). No species misidentifications were detected. The remaining 125 non-admixed Razorback Sucker larvae were used for analysis.

Using the ddRAD method of discovery, a total of 165 Razorback Sucker larvae yielded SNP data. The analysis of these data in AdmixPipe and Clumpak detected 46 hybrid larvae, which accounted for 27.8% of samples (Table 2). The hybrid individuals were removed from all downstream analysis. No species misidentifications were detected through SNP analysis. The same individuals were identified as hybrids using both microsatellites and SNPs (Figure 7). After removing hybrids, 18,335 SNP loci were identified in the remaining Razorback Sucker larvae (N = 119).

Most years (cohorts) prior to 2020 contained misidentified larvae (Table 3) and hybrid individuals (Table 2). Morphological identification was very accurate in most years, but a higher proportion of hybrids (~27%) was included among Razorback Sucker samples in 2020 relative to preceding years. This was approximately a threefold increase compared to the mean proportion of hybrids detected per year (8.9%). Over all the years very few misidentifications were noted, with most (N = 14; 67%) being Bluehead Sucker that were misidentified as Razorback Sucker. There were also Flannelmouth Sucker misidentified as Razorback (N = 7; 33%). The misidentified samples were reclassified as their appropriate species (i.e., Razorback Sucker or Flannelmouth Sucker) for all downstream analysis.

The 2020 N_b estimates for Razorback Sucker were very uniform between microsatellites and SNPs (Table 4). The N_b estimates for Razorback Sucker cohorts were relatively consistent across all years with a mean of 150 and a range of 74-213 (Figure 8). Adult census size showed little among-year variation, with overlapping confidence intervals and similar means (range = 2,796 – 5,411 individuals; Figure 9). The estimated proportion of spawning adults was low with a comparatively consistent mean ratio of 3.8% (Table 5). Furthermore, we observed little variation within the N_b/N_c with the lowest estimate (2.2%) in 2013 and the highest estimate (7.06%) in 2019.

The estimated number of spawning Razorback Sucker adults was relatively similar across years (Figure 10). Like the Colorado Pikeminnow, the lowest and highest N_b estimates coincided with the lowest and highest numbers of spawning adults. Although there was some evidence of variance in reproductive success, results overall indicated relatively even contribution of offspring per parent. This is apparent in the similar measures between the N_b estimate and number of parents that produced

offspring. The number of parental pairs was similar across years, with more parental pairs than estimated number of parents (Figure 11 and 12). Similar to Colorado Pikeminnow, there is evidence of extensive larval drift between full-siblings, but the collection area was constrained in comparison to recent years due to COVID-19 restrictions (Figure 11 and 12). In past years some full-siblings were collected over 120 river miles apart, but in 2020 only two reaches of the river were accessible (RM 180 to 166 and RM 77 to 0). Full-siblings were found in both of those reaches using both microsatellites and SNPs (Table 6). There were siblings that drifted apart within, but not across, those two reaches (Figure 12). No larval sampling was conducted in the San Juan arm of Lake Powell downstream of the Piute Farms Waterfall in 2020 due to COVID-19 restrictions.

Discussion

Population size is often considered a central parameter for conservation; however, monitoring abundance is often problematic and monitoring how that abundance relates to genetic diversity is equally challenging. The maintenance of genetic variation is essential to the survival of natural populations. To prevent the loss of genetic variation in endangered populations, a common practice has been to supplement them with captive-bred individuals. Stocking is very common in the management of endangered fish populations, many of which depend on restocking for their entire existence. The ultimate goal of most conservation efforts is a population that is genetically diverse, with enough abundance for a self-sustaining population to recruit in a natural environment (Leinonen et al. 2020). Estimates of N_b from genetic methods are useful for tracking population trends to assess if factors could be limiting the reproductive potential of a population (Ferchaud et al. 2016, Bacles et al. 2018; Davenport et al. 2020). Long-term larval monitoring has contributed to an increased understanding of reproduction of San Juan River endangered fishes, and, in conjunction with genetic tools, can further elucidate if population-level reproductive success is a factor limiting recruitment of Razorback Sucker and Colorado Pikeminnow.

In the past, N_b and sibling relationships were calculated for both Colorado Pikeminnow and Razorback Sucker using microsatellites. In 2020, these species were genotyped via both microsatellites and SNPs. Both marker types can be used to estimate relatedness and N_b, but SNPs typically provide more precise estimates due to availability of a large number of markers combined with lower genotyping error relative to microsatellites (Galla et al. 2020, Sunde et al. 2020). The ddRAD method of SNP discovery was used for both species. It was highly effective at discovering SNPs in Razorback Sucker (N= 18,335). The SNPs and microsatellites were used to estimate both N_b and full-sibling relationships in Razorbacks, which resulted in very similar outcomes. This result suggests that future Razorback Sucker work can be conducted solely with SNPs since other studies have shown them to be more precise (Gala et al. 2020, Sunde et al. 2020). The initial effort to identify SNPs in Colorado Pikeminnow was not as successful (N= 2,227). The high level of relatedness among the broodstock used to stock the San Juan River had a profound effect on the number of loci that were discovered (Diver et al. 2020). In the future, the ddRAD method for discovering SNPs in Colorado Pikeminnow will be expanded to identify additional SNPs to compensate for the relatedness issue. With a more informative set of SNPs, Colorado Pikeminnow should be able to transition to a SNP-only platform like Razorback Sucker.

The N_b estimates for Colorado Pikeminnow have been variable over the years. Not only do the N_b estimates vary, but so do the adult census sizes (N_c) estimates. Some years N_b estimates were comparable to the N_c with 40% of the mean estimated number of adults contributing to larval cohorts. That varied greatly between years, potentially due to large variability in adult census estimates. These highly variable N_c results are due to few adult Colorado Pikeminnow being contacted during monitoring efforts. Regardless of the large inter-annual variation observed in N_c estimates, there has been an

increase in N_b estimates over the past two years, suggesting that conditions might be improving for Colorado Pikeminnow spawning success in the San Juan River.

The 2020 N_b estimates for Razorback Sucker were very low compared to the previous five years, but this was most likely due to limited collections occurring under COVID-19 restrictions. This resulted in a reduced number of larvae collected in 2020 compared to previous years, which could have negatively affected N_b estimates. Even when excluding the 2020 sample year, recent N_b estimates for Razorback Sucker are surprisingly low given current population size estimates. Annual contribution of adults to each larval cohort was minimal with the highest estimate of 7.6% observed in 2019 and a consistently low average of 3.8% across years. These results suggest that there is some factor limiting Razorback Sucker reproduction in the San Juan River. Determining the limiting factor, whether it be limited access to suitable spawning habitat or high mortality of eggs or early larval stages, is very difficult. Census size estimates are high, so the ability to find a suitable mate should not be a limiting factor, but the high level of hybrids in the 2020 larval samples does raise concerns. In 2018 and 2019 there was an increased number of larvae collected because of the below-waterfall samples from the San Juan arm of Lake Powell, and those parents were estimated in the N_b calculations. They do not suggest separate spawning events below Piute Farms Waterfall since half- and full-siblings were identified between the two sampled groups.

Little work has been done to assess the range-wide genetic status of Colorado Pikeminnow, particularly in regard to N_b, with the San Juan River estimates representing the first available N_b estimates for the species. There have been N_b estimates for Razorback Sucker calculated in the lower Colorado River basin (Lake Mohave Reservoir). Annual population monitoring and genetic studies conducted on Razorback Sucker from Lake Mohave between 1997 and 2010 showed a significant increase in mean N_b estimates (e.g., N_b = 743 [1998]; N_b = 49,984 [2005]) and N_b/N_c (mean 70.8%) over a fourteen-year period (Dowling et al. 2013; Carson et al. 2016). These results suggest management actions, in particular offshore rearing of larvae, have effectively reduced the variance in population-level reproductive success for Razorback Sucker in Lake Mohave (Carson et al. 2016). However, Lake Mohave is a very different system than the San Juan River. It is a very large impoundment with little current moving sediment or larvae downstream, crystal clear waters, no other native suckers present for potential hybridization, and different management activities. Therefore, the high N_b/N_c estimates are likely not directly comparable. Nonetheless, the Lake Mohave results showed management positively affected N_b, suggesting regular N_b monitoring of the San Juan River population could inform managers when management actions could increase population-level reproductive success.

It is expected in natural populations for N_b to be less than N_c (Waples et al., 2013). The factors that cause N_b to be lower than N_c are often attributed to biological processes such as demographic characteristics, habitat availability, or life-history (Palstra and Fraser 2012; Kanno et al. 2015; Bernos et al. 2018). It is critical that the accuracy of estimates for both N_b and N_c are correct when using them for management decisions. In many cases, N_c is notoriously difficult to estimate, especially for species with age structure (Luikart et al. 2010; Luikart et al. 2021). The accuracy of N_b estimates can be hindered when genetic information is scarce or age-specific survival, fecundity, or longevity is unknown, creating a bias (Wang 2016; Luikart et al. 2021). N_b estimates are most accurately determined from a sample size that is close to or two times greater than the real N_b (Wang 2016; Sánchez-Montes et al. 2017; Bacles et al. 2018). Based upon this information, the sampling size for both species appears to be sufficient since the

increased sample sizes in 2018-2019 did not increase N_b estimates. Finally, biases in spatial and temporal sampling can result in lower N_b estimates (Bacles et al. 2018). That was reflected in the dramatically lower N_b seen in 2020 for Razorback Sucker, which had a greatly truncated collection season in comparison to past years. Spatial sampling was constrained between Shiprock, New Mexico and Clay Hills, Utah; thus, if spawning occurred relatively low in the system and a majority of larvae were lost to Lake Powell, or if spawning events occurred between sampling trips and were flushed down river or into sediment filled areas, then N_b estimates might be artificially low.

Acknowledgements

This work could not be possible without access to the long-term larval collections at the Museum of Southwestern Biology. For that, we thank Michael Farrington, Howard Brandenburg, Jen Kennedy, Steve Platania, ASIR, and the many individuals that have contributed to this dataset. We thank the Program Office for providing adult census data and we want to give special thanks to Eliza Gilbert for assisting with larval fish staging. Assistance with the larval collections was provided by Emily DeArmon, along with other staff at the Museum of Southwestern Biology. We are sincerely thankful to current and former staff at Southwestern ARRC for help with QAQC procedures, data analysis, and logistical support, including Casey Weathers, Tracy Diver, Trish Koenig, Manuel Ulibarri, and Wade Wilson. This study was approved by the San Juan River Basin Biology Committee through the San Juan River Basin Recovery Implementation Program and funded under a U.S. Bureau of Reclamation, Salt Lake City Project Office.

Signed:
Melody Saltzgiver and Steve Mussmann
Principal Investigators
5/20/2022

Literature Cited

Alexander D.H., J. Novembre, and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19:1655–1664.

Anderson, E.C. and E.A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, 160:1217-1229.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bacles, C.F.E., C. Bouchard, F. Lange, A. Manicki, C. Tentelie and O. Lepais. 2018. Estimating the effective number of breeders from single parr samples for conservation monitoring of wild populations of Atlantic salmon *Salmo salar*. *Journal of Fish Biology*, 92(3):699-726.

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.

Benjamini, Y. and D. Yekutieli. 2001. The control of false discovery rate under dependency. *The Annals of Statistics*, 29:1165-1188.

Bernos, T.A., M.C. Yates and D.J. Fraser. 2018. Fine-scale differences in genetic and census population size ratios between two stream fishes. *Conservation Genetics*, 19:265–274.

Bruton, M.N. 1995. Have fishes had their chips? The dilemma of threatened fishes. Environmental *Biology of Fishes*, 43:1–27.

Camacho-Sanchez, M., G. Velo-Antón, J.O.Hanson, A. Veríssimo, I. Martínez-Solano, A. Marques, C. Moritz and S. Carvalho. 2020. Comparative assessment of range-wide patterns of genetic diversity and structure with SNPs and microsatellites: A case study with Iberian amphibians. *Ecology and Evolution*, 10:10353– 10363.

Carson, E.W., T.F. Turner, M.J. Saltzgiver, D. Adams, B.R. Kesner, P.C. Marsh, T.J. Pilger and T.E. Dowling. 2016. Retention of ancestral genetic variation across life-stages of an endangered, long-lived iteroparous fish. *Journal of Heredity*, 107(6):567-72.

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.

Chen S, Y. Zho, Y. Chen and J. Gu. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17):884-890.

Clarkson R.W., P.C. Marsh, S.E. Stefferud and J.A. Stefferud. 2005. Conflicts between Native Fish and Nonnative Sport Fish Management in the Southwestern United States. *Fisheries*, 30(9):20-27.

DaCosta, J. M. and M. D. Sorenson. 2014. Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PloS one*, 9(9):e106713.

Davenport D, P. Butcher, S. Andreotti, C. Matthee, A. Jones and J. Ovenden. 2020. Effective number of white shark (*Carcharodon carcharias*, Linnaeus) breeders is stable over four successive years in the population adjacent to eastern Australia and New Zealand. *Ecology and Evolution*, 11(1):186-198.

Diver, T., A. Harrison, and W. Wilson. 2020. Genetic evaluation and history of captive broodstock populations of endangered Colorado Pikeminnow (*Ptychocheilus lucius).* U.S. Fish and Wildlife Service, Dexter, New Mexico.

Do, C., R.S. Waples, D. Peel, G. M. Macbeth, B. J. Tillett and J.R. Ovenden. 2014. NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, 14:209-214.

Dowling, T Dowling, T.E., T.F. Turner, E.W. Carson, M.J. Saltzgiver, D. Adams, B. Kesner and P.C. Marsh. 2013. Time-series analysis reveals genetic responses to intensive management of razorback sucker (*Xyrauchen texanus*). *Evolutionary Applications*, 7:339-354.

Evanno, G, S. Regnaut and J. Gould. 2005. Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology*, 14:2611–2620.

Falush, D., M. Stephens and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164:1567-1587.

Ferchaud, A.L., C. Perrier, J. April, C. Hernandez, M. Dionne and L. Bernatchez. 2016. Making sense of the relationships between N_e, N_b and N_c towards defining conservation thresholds in Atlantic salmon (*Salmo salar*). *Heredity*, 117:268–278.

Frankham R. 1995. Effective population size/adult population size ratios in wildlife: a review. *Genetic Resources,* 66:95–107.

Galla, S. J., R. Moraga, L. Brown, S. Cleland, M.P. Hoeppner, R.F. Maloney, A. Richardson, L. Slater, A.W. Santure and T. 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 programs worldwide. *Evolutionary Applications*, 13:991–1008.

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

Hedrick P.N. 2005. Large variance in reproductive success and the Ne/N ratio. *Evolution,* 59:1596–1599.

Huisman, J. 2017. Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, 17(5):1009-1024.

Johnson, P.C.D. and D.T. Haydon. 2007. Software for quantifying and simulating microsatellite genotyping error. *Bioinformatics and Biology Insights*, 1:71-75.

Jones, O.R. and J. Wang. 2010. Colony: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources,* 10:551-555.

Kalinowski S.T. and R.S. Waples. 2002. Relationship of effective to census size in fluctuating populations. *Conservation Biology*, 16:129–136.

Kanno, Y., B.H. Letcher, N.P. Hitt, D.A. Boughton, J.E. Wofford and E.F. Zipkin. 2015. Seasonal weather patterns drive population vital rates and persistence in a stream fish. *Global Change Biology*, 21(5):1856-1870.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg and I. Mayrose. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5):1179-1191.

La Haye, M.J.J., T.E. Reiners, R. Raedts, V. Verbist and H. P. Koelewijn. 2017. Genetic monitoring to evaluate reintroduction attempts of a highly endangered rodent. *Conservation Genetics*, 18:877–892.

Leinonen, T., J. Piironen, M.L. Koljonen, J. Koskiniemi and A. Kause. 2020. Restored river habitat provides a natural spawning area for a critically endangered landlocked Atlantic salmon population. *PLoS One,* 15(5):e0232723.

Lemopoulos, A., J.M. Prokkola., S. Uusi-Heikkilä, A. Vasemägi, A. Huusko, P. Hyvärinen, M. 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(4):2106–2120.

Le Vay L., G.R. Carvalho, E.T. Quinitio, J.H. Lebata, U.N. Ut and H. Fushimi. 2007. Quality of hatchery-reared juveniles for marine fisheries stock enhancement. *Aquaculture,* 268:169–180.

Lindenmayer, D., J. Woinarski, S. Legge, D. Southwel, T. Lavery, N. Robinson, B. Scheele and B. Wintle. 2020. A checklist of attributes for effective monitoring of threatened species and threatened ecosystems. *Journal of Environmental Management*, 262:110312.

Luikart, G., N. Ryman, D.A. Tallmon, M.K. Schwartz and F.W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11:355-373.

Luikart, G, T. Antao, B.K. Hand, C.C. Muhlfeld, M.C. Boyer, T. Cosart, B. Trethewey, R. Al-Chockhachy and R.S. Waples. 2021 Detecting population declines via monitoring the effective number of breeders (N_b). *Molecular Ecology Resources*, 21(2):379-393.

Martin, R.M., M.L. Robinson and W.D. Wilson. 2015. Isolation and characterization of twenty-five novel microsatellite loci in Colorado Pikeminnow, *Ptychocheilus lucius*, with cross-species amplification for eight other cyprinids. *Conservation Genetic Resources,* 7:113.

Miller R.R. 1961. Man and the changing fish fauna of the American Southwest. *Papers of the Michigan Academy of Science, Arts and Letters*, XLVI: 365-404.

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

Minckley, W.L. and J.E. Deacon. 1968. Southwestern fishes and the enigma of "endangered species". *Science*, 159:1424-1433.

Minckley, W.L. and J.N. Rinne. 1991. Native fishes of the arid lands: a dwindling resource of the desert southwest. General Technical Report RM-206. USDA Forest Service, Rocky Mountain Forest Range Experiment Station, Fort Collins, Colorado.

Minckley, W.L. and P.C. Marsh. 2009. Inland fishes of the greater Southwest: chronicle of a vanishing biota. University of Arizona Press Tucson.

Mussmann, S.M., M.R. Douglas, T.K. Chafin and M.E. Douglas. 2020. ADMIXPIPE: population analyses in ADMIXTURE for non-model organisms. *BMC Bioinformatics*, 21:337.

Narum, S. 2006. Beyond Bonferroni: less conservative analyses for conservation genetics. Conservation *Genetics*, 7:783-787.

Palstra, F.P. and D.J. Fraser. 2012. Effective/census population size ratio estimation; a compendium and appraisal. Ecology and Evolution 2(9):2357-2365.

Perrier, C., J. April, G. Cote, L. Bernatchez and M. Dionne. 2016. Effective number of breeders in relation to census size as management tools for Atlantic salmon conservation in a context of stocked populations. *Conservation Genetics*, 17:31–44.

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):1–11.

Pritchard, J.K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155:945-959.

Rasmussen J.E., M.C. Belk and S.L. Peck. 2009. Endangered species augmentation: a case study of alternative rearing methods. *Endangered Species Research*, 8:225-232.

Raymond, M. and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact test and ecumenicism. *Journal of Heredity*, 86:248-249.

Ricciardi A. and J. E. Rasmussen. 1999. Extinction rates of North American freshwater fauna. *Conservation Biology*, 13:1220–1222.

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-Colon and J.M. Catchen. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21):4737-4754.

Sánchez-Montes, G., J. Wang, A.H. Ariño, J.L. Vizmanos and I. Martinez-Solano. 2017. Reliable effective number of breeders/adult census size ratios in seasonal-breeding species: Opportunity for integrative demographic inferences based on capture-mark-recapture data and multilocus genotypes. *Ecology and Evolution*, 23:10301-10314.

San Juan River Basin Recovery Implementation Program (SJRRIP). 2017. Population abundance estimates for Colorado Pikeminnow and Razorback Sucker in the San Juan River. U.S. Fish and Wildlife Service, New Mexico Ecological Services Field Office, Albuquerque, New Mexico.

Sard, N., M. Johnson, D. Jacobson, M. Hogansen, K. O'Malley and M. Banks. 2016. Genetic monitoring guides adaptive management of a migratory fish reintroduction program. *Animal Conservation*, 19:570-577.

Schleicher, B.J. 2016. Long term monitoring of sub-adult and adult large-bodied fishes in the San Juan River: 2016. U.S. Fish and Wildlife Service, Grand Junction Fish and Wildlife Conservation Office, Grand Junction, Colorado.

Schleicher, B.J., B. Duran and B. Hines. 2019. San Juan River Demographic Monitoring 2019. U.S. Fish and Wildlife Service, Grand Junction Fish and Wildlife Conservation Office, Grand Junction, Colorado.

Schrey, A.W and E. Heist. 2011. Hybridization between pallid sturgeon *Scaphirhynchus albus* and shovelnose sturgeon *Scaphirhynchus platorynchus*. *Journal of Fish Biology*, 79:1828-1850.

Su, G., M. Logez, J. Xu, S. Tao, S. Villéger and S. Brosse. 2021. Human impacts on global freshwater fish biodiversity. Science 371:835–838.

Sunde, J., Y. Yildirim, P. Tibblin and A. Forsman. 2020. Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics*, 11:218.

Wang, J. 2016. A comparison of single-sample estimators of effective population sizes from genetic marker data. *Molecular Ecology*, 25:4692-4711.

Wang, J., E. Santiago and A. Caballero. 2016. Prediction and estimation of effective population size. *Heredity*, 117(4):193–206.

Waples R. S. 2005. Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, 14:3335–3352.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics*, 7:167-184.

Waples, R.S., G. Luikart, J.R. Faulkner and D.A. Tallmon. 2013. Simple life history traits 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.

Waples, R.S. and C. Do. 2010. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3:244-262.

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

Wilson, W.D. 2012. Razorback sucker broodstock evaluation and genetic monitoring, 2012 annual report. Report to Lower Colorado River Multi-Species Conservation Office, Bureau of Reclamation, Interagency Acquisition R11PG30006.

Table 1.
Mean effective number of breeder (N_b) and adult census (N_c) estimates for Colorado Pikeminnow (*Ptychocheilus lucius*) calculated with microsatellites. Parentheses contain 95% confidence intervals. The estimated proportion of spawning adults (N_b/N_c) was calculated from mean N_b values. Areas with dashes indicate data were not available for those years.

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
N_b	10 (6 - 15)	~	~	26 (22 - 31)	4 (3 - 6)	44 (37 - 53)	25 (21 - 29)	27 (19 - 40)	58 (52 - 65)	52.1 (49 - 55)
N_c	81 (8 - 1125)	19 (7 - 65)	~	67 (21 - 282)	100 (35 - 333)	133 (31 - 652)	242 (167 - 381)	617 (244 - 1753)	142 (47-186)	~
N_b/N_c	12.3	~	~	38.8	4.0	33.1	10.3	4.4	40.8	~

Table 2.
Total number of F1 hybrids, F2 hybrids, Razorback Sucker (*Xyrauchen texanus*) backcrosses and Flannelmouth Sucker (*Catostomus latipinnis*) backcrosses identified among Razorback Sucker larvae per year from the San Juan River.

Year	F1 hybrids	F2 hybrids	Razorback Sucker Backcross	Flannelmouth Sucker Backcross	Mean Percentage Removed
2009	2	2	3	0	5.8%
2010	0	9	3	0	7.5%
2011	1	3	4	0	5.8%
2012	2	7	0	0	7.5%
2013	2	10	8	1	9.0%
2014	0	7	3	9	7.6%
2015	3	5	11	3	9.2%
2016	1	9	2	3	5.8%
2017	0	4	7	1	5.0%
2018	7	14	11	5	12.3%
2019	0	6	1	1	4.8%
2020 SNPs	33	8	4	0	27.8%
2020 Microsatellites	33	8	4	0	26.9%
Total	51	84	57	23	8.9%

Table 3.
Number of misidentified larvae each year calculated with microsatellites (2009-2020) and SNPs (2020). The top headings, Razorback Sucker (*Xyrauchen texanus*), indicate morphological identification of each larva. Subheadings indicate the number of fish inferred to be a different species by genetic assignment in the program NewHybrids. The heading 'Misidentified Flannelmouth Sucker' indicates the number of Flannelmouth Sucker (*Catostomus latipinnis*) per year that were misidentified as Razorback Sucker and the heading 'Misidentified Bluehead Sucker' indicates the number of Bluehead Sucker (*Catostomus discobolus*) per year that were misidentified as Razorback Sucker based upon morphology.

	Razorback Sucker	
Year	Misidentified Flannelmouth Sucker	Misidentified Bluehead Sucker
2009	0	0
2010	0	0
2011	3	1
2012	0	0
2013	0	5
2014	0	3
2015	1	1
2016	1	2
2017	2	0
2018	0	1
2019	0	1
2020	0	0
Total	7	14

Table 4.
The effective number of breeders (N_b) for Razorback Sucker (*Xyrauchen texanus*) larvae calculated using NeEstimator for both microsatellites and single nucleotide polymorphisms (SNPs) for 2020. Not all of larvae evaluated with microsatellites generated SNPs due to failed ddRAD reactions or DNA quality issues.

	Microsatellites	**SNPs**
# Larvae	125	119
N_b	131.7	129.1
Confidence Intervals	121.9 - 142.8	126.0 - 132.4

Table 5.
Mean effective number of breeder (N_b) and adult census (N_c) estimates for Razorback Sucker (*Xyrauchen texanus*) calculated with microsatellites (2009-2019) and SNPs (2020). Parentheses contain 95% confidence intervals. The estimated proportion of spawning adults (N_b/N_c) was calculated from mean N_b values. Areas with dashes indicate data were not available for those years.

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
N_b	74 (60 - 94)	88 (70- 112)	130 (98 - 183)	144 (107 - 211)	99 (72 - 149)	115 (82 - 176)	174 (132 - 247)	184 (138 - 265)	233 (170 - 368)	196 (146 - 283)	213 (169 - 278)	129 (126 - 132)
N_c	~	~	4154 (2554 - 6818)	4915 (3687 - 6585)	4558 (3257 - 6413)	3818 (2915 - 5023)	5411 (3923 - 7508)	3826 (3099 - 4752)	~	~	2796 (2461 - 3210)	~
N_b/N_c	~	~	3.1	2.9	2.2	3	3.2	4.8	~	~	7.6	~

Table 6.
A comparison of microsatellites and single nucleotide polymorphisms (SNPs) for full-sibling relationships between Razorback Sucker (*Xyrauchen texanus*) larvae collected from the San Juan River in 2020. The microsatellite pairs were generated in Colony v2.0.4.0 and the SNP pairs were generated in Sequoia v2.3.5. Pairs marked with (*) refer to sibling pairs not detected with microsatellites; the mark (^) refers to sibling pairs not detected with SNPs; and the mark (~) refers to samples that the SNP reaction failed to amplify.

Microsatellite Full-sibling Pairs (21 loci)		**Single Nucleotide Polymorphisms Full-sibling Pairs (18335 Loci)**	
OffspringID1	OffspringID2	OffspringID1	OffspringID2
XtexW20SJR_213	XtexW20SJR_276	XtexW20SJR_213	XtexW20SJR_276
XtexW20SJR_236	XtexW20SJR_242	XtexW20SJR_236	XtexW20SJR_242
XtexW20SJR_238	XtexW20SJR_240	XtexW20SJR_238	XtexW20SJR_240
XtexW20SJR_293	XtexW20SJR_294	XtexW20SJR_293	XtexW20SJR_294
XtexW20SJR_338	XtexW20SJR_339	XtexW20SJR_338	XtexW20SJR_339
XtexW20SJR_343	XtexW20SJR_367	XtexW20SJR_343	XtexW20SJR_367
		XtexW20SJR_345	XtexW20SJR_371*
XtexW20SJR_349	XtexW20SJR_370	XtexW20SJR_349	XtexW20SJR_370
XtexW20SJR_350	XtexW20SJR_369	XtexW20SJR_350	XtexW20SJR_369
XtexW20SJR_351	XtexW20SJR_366	XtexW20SJR_351	XtexW20SJR_366
XtexW20SJR_354	XtexW20SJR_362	XtexW20SJR_354	XtexW20SJR_362
XtexW20SJR_354	XtexW20SJR_355	XtexW20SJR_354	XtexW20SJR_355
		XtexW20SJR_355	XtexW20SJR_362*
XtexW20SJR_356	XtexW20SJR_365	XtexW20SJR_356	XtexW20SJR_365
XtexW20SJR_358	XtexW20SJR_371^		
XtexW20SJR_359	XtexW20SJR_363~		
XtexW20SJR_359	XtexW20SJR_364~		
XtexW20SJR_363	XtexW20SJR_364	XtexW20SJR_363	XtexW20SJR_364
XtexW20SJR_368	XtexW20SJR_372	XtexW20SJR_368	XtexW20SJR_372
XtexW20SJR_377	XtexW20SJR_382	XtexW20SJR_377	XtexW20SJR_382

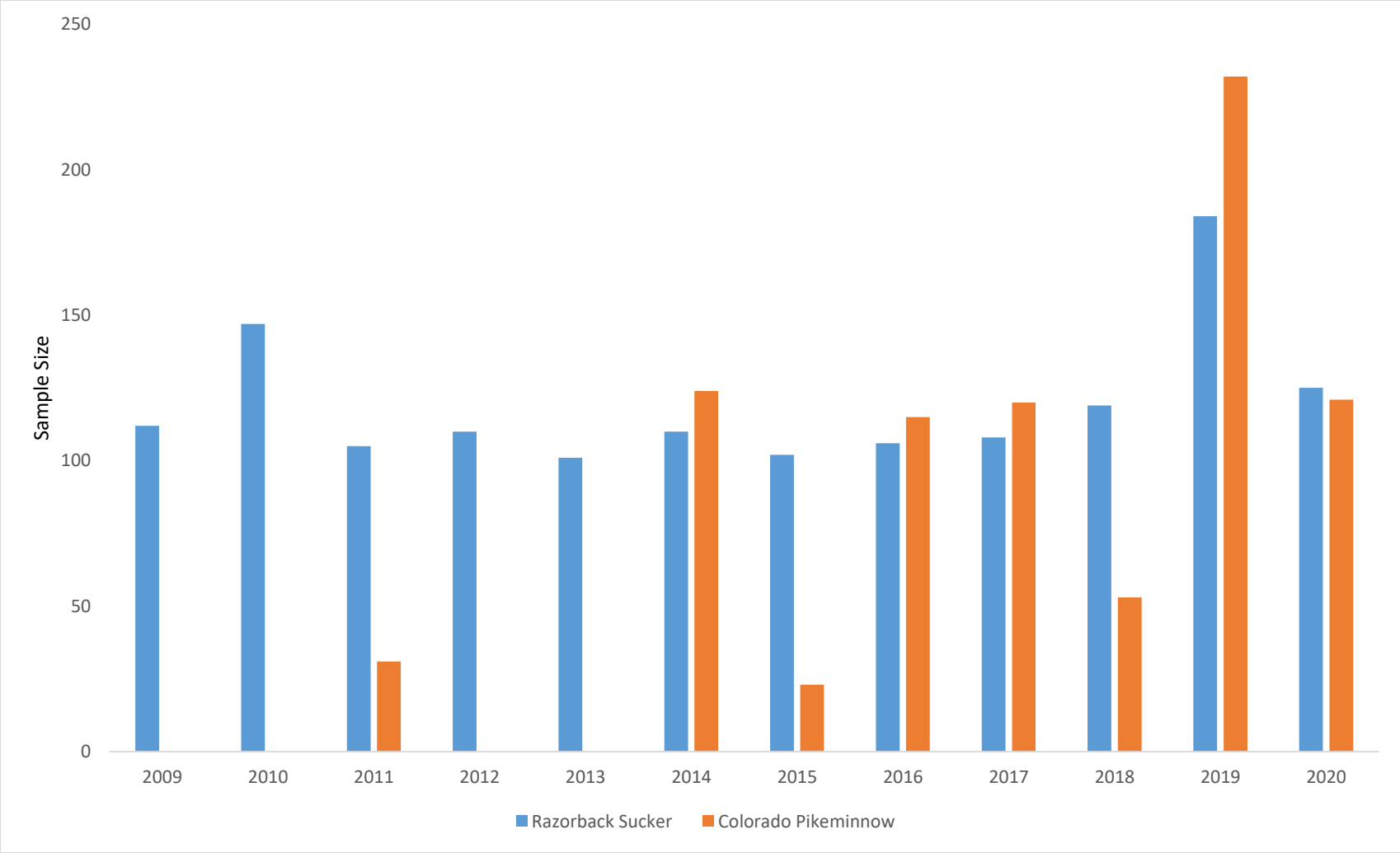

Figure 1.
Sample sizes used for effective number of breeder (N_b) estimates for 2009-2019. Razorback Sucker (*Xyrauchen texanus*) are represented in light blue and Colorado Pikeminnow (*Ptychocheilus lucius*) in orange.

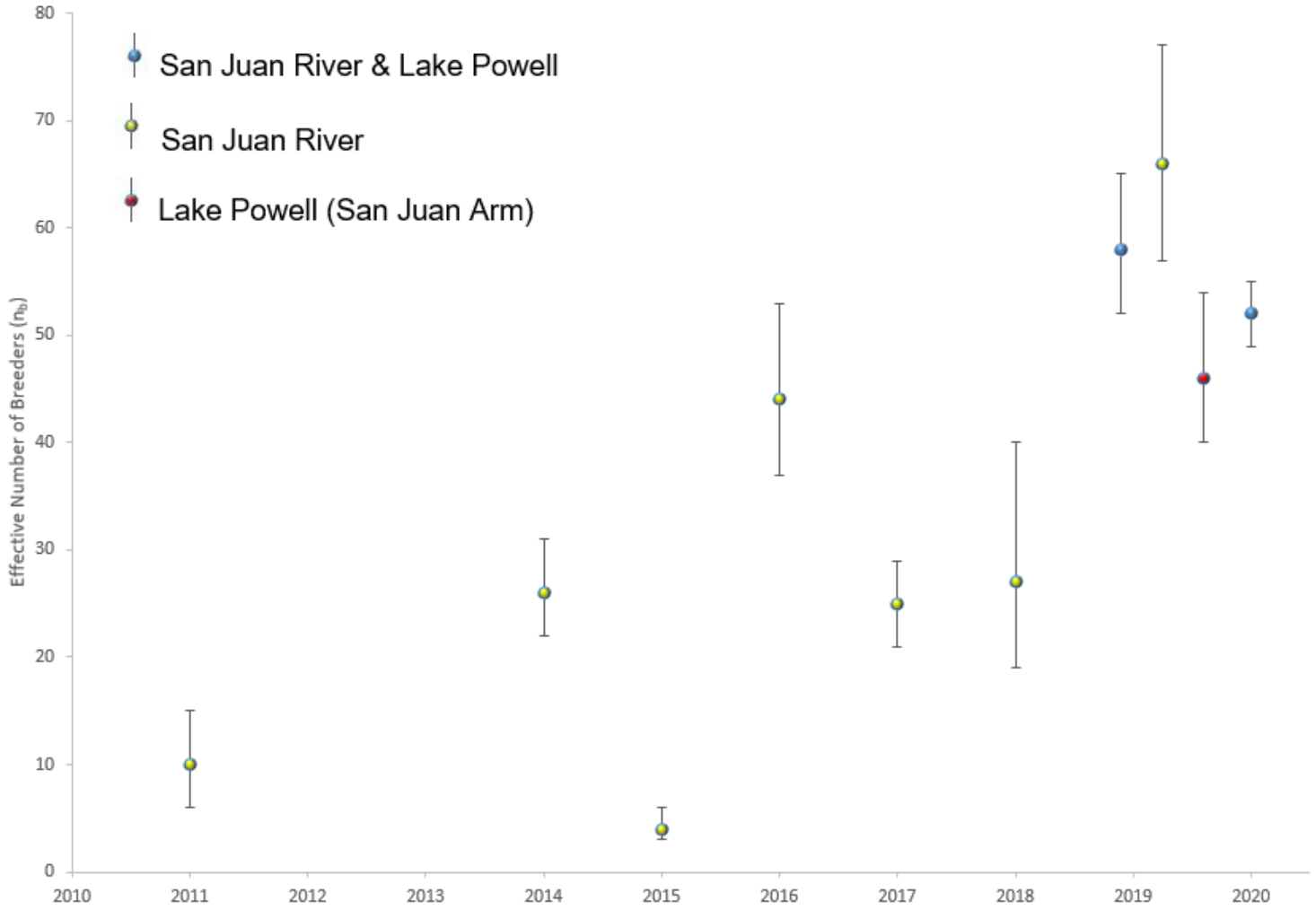

Figure 2.
Effective number of breeder (N_b) estimates for Colorado Pikeminnow (*Ptychocheilus lucius*) calculated with microsatellites. Circles represent mean estimates with vertical bars showing 95% confidence intervals. The 2019 samples are broken down by below the waterfall (Lake Powell), above the waterfall (San Juan River) and both combined.

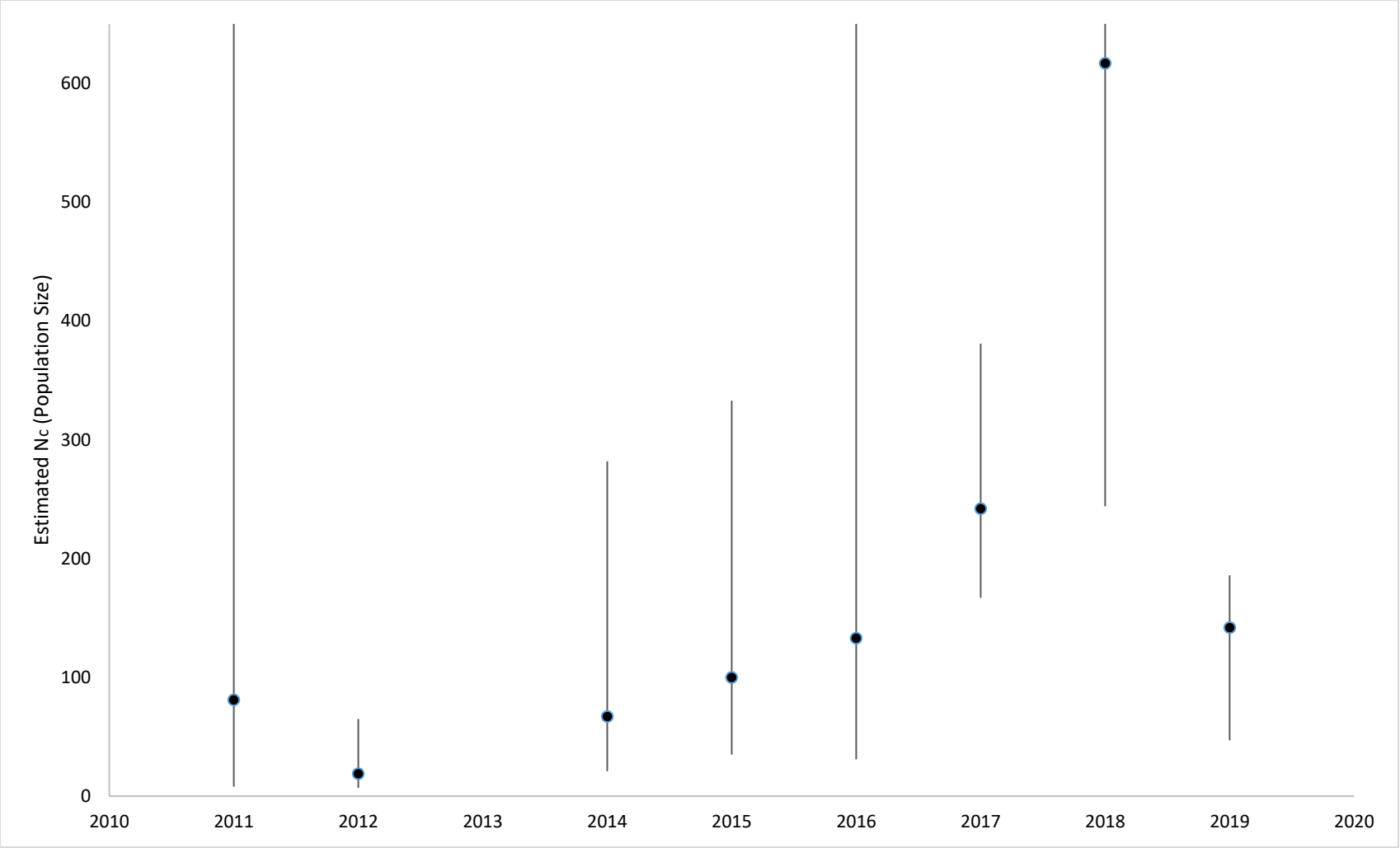

Figure 3.
Estimated adult census sizes (N_c) for Colorado Pikeminnow (*Ptychocheilus lucius*) in the San Juan River. Circles represent mean adult census sizes and the vertical bar represent 95% confidence intervals.

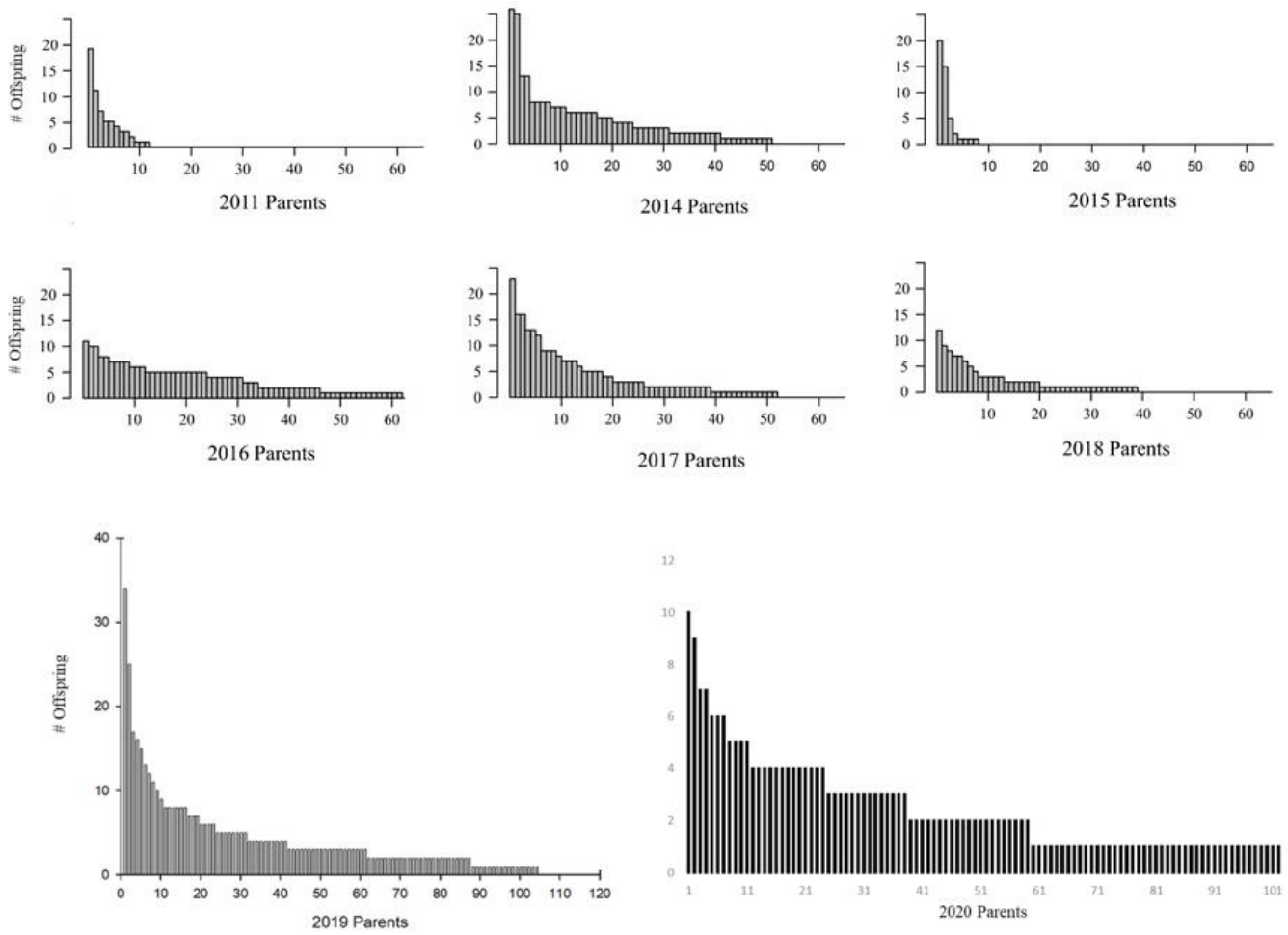

Figure 4.
Frequency of offspring contributed by individual Colorado Pikeminnow (*Ptychocheilus lucius*) parents from the San Juan River (2011 and 2014 – 2020) calculated with microsatellites.

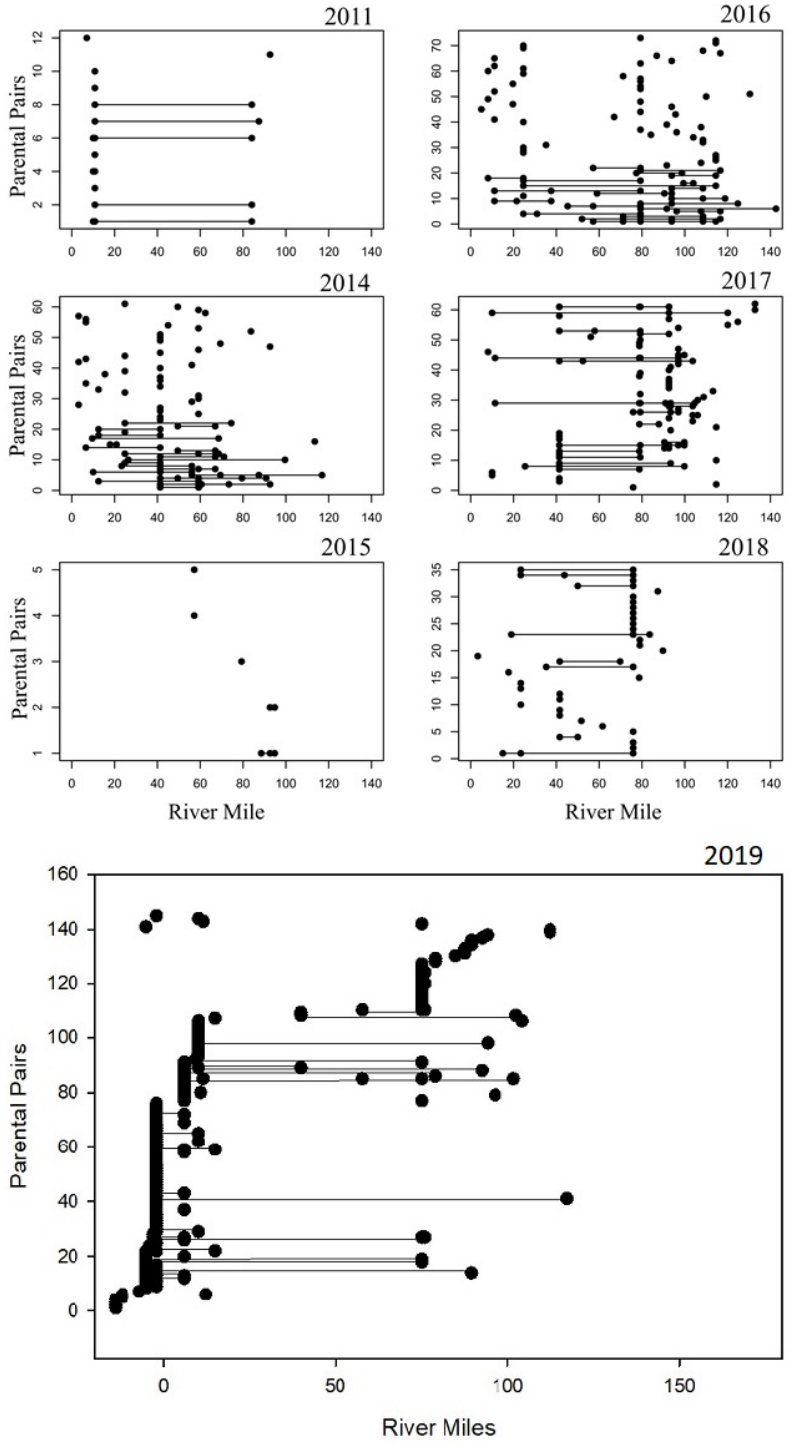

Figure 5.
Estimated number of parental pairs for Colorado Pikeminnow (*Ptychocheilus lucius*) and the spatial distribution of larval samples calculated with microsatellites. Each circle represents a larval Colorado Pikeminnow with lines connecting each circle indicating full-sibling relatedness. Spatial distribution of samples is shown by plotting larvae according to the river mile at which they were collected (X axis).

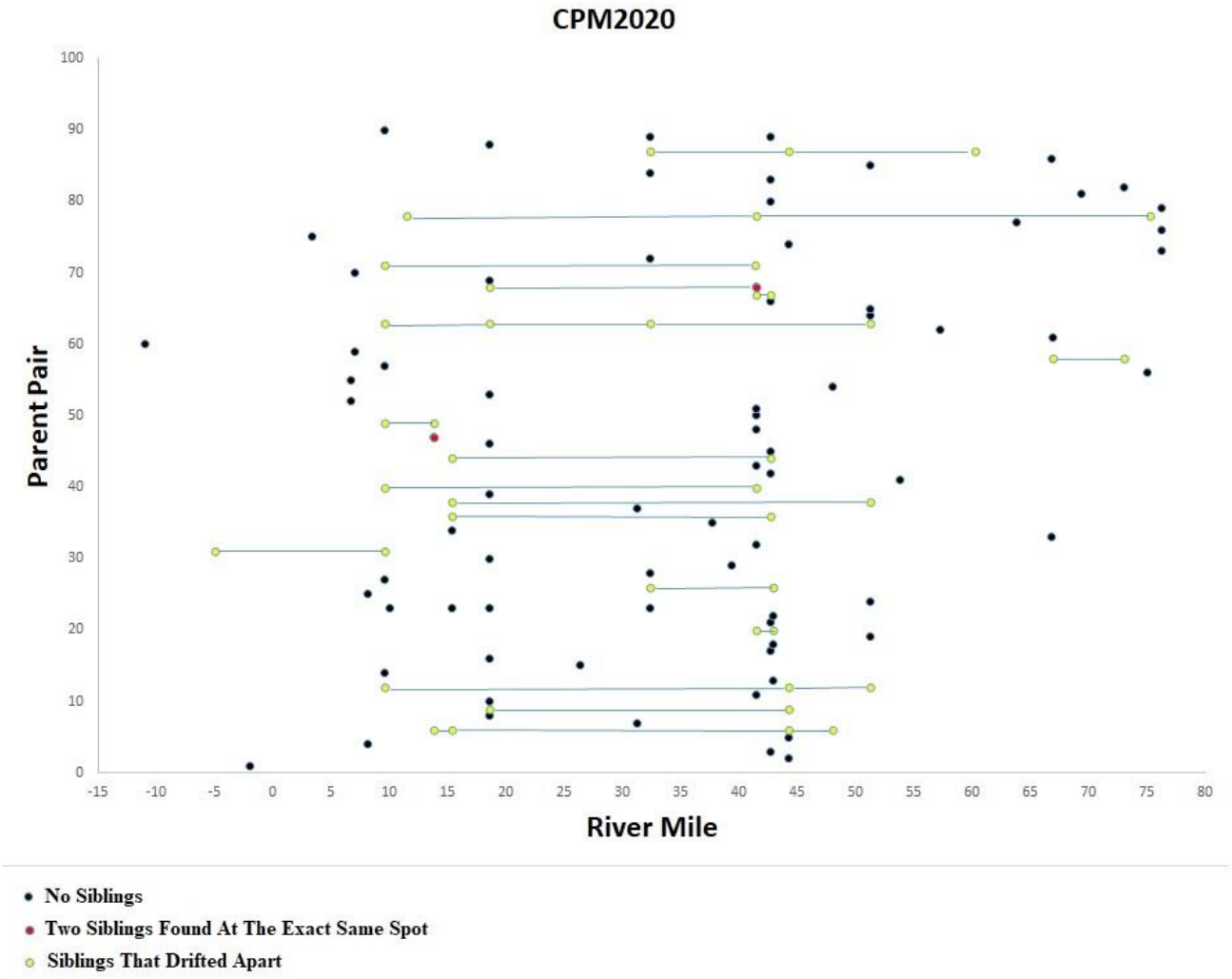

Figure 6.
Estimated number of parental pairs for Colorado Pikeminnow (*Ptychocheilus lucius*) and the spatial distribution of the larval samples calculated with microsatellites. Each circle represents a larval Colorado Pikeminnow with lines connecting each circle indicating full-sibling relatedness. Spatial distribution of samples is shown by plotting larvae according to the river mile at which they were collected (X axis).

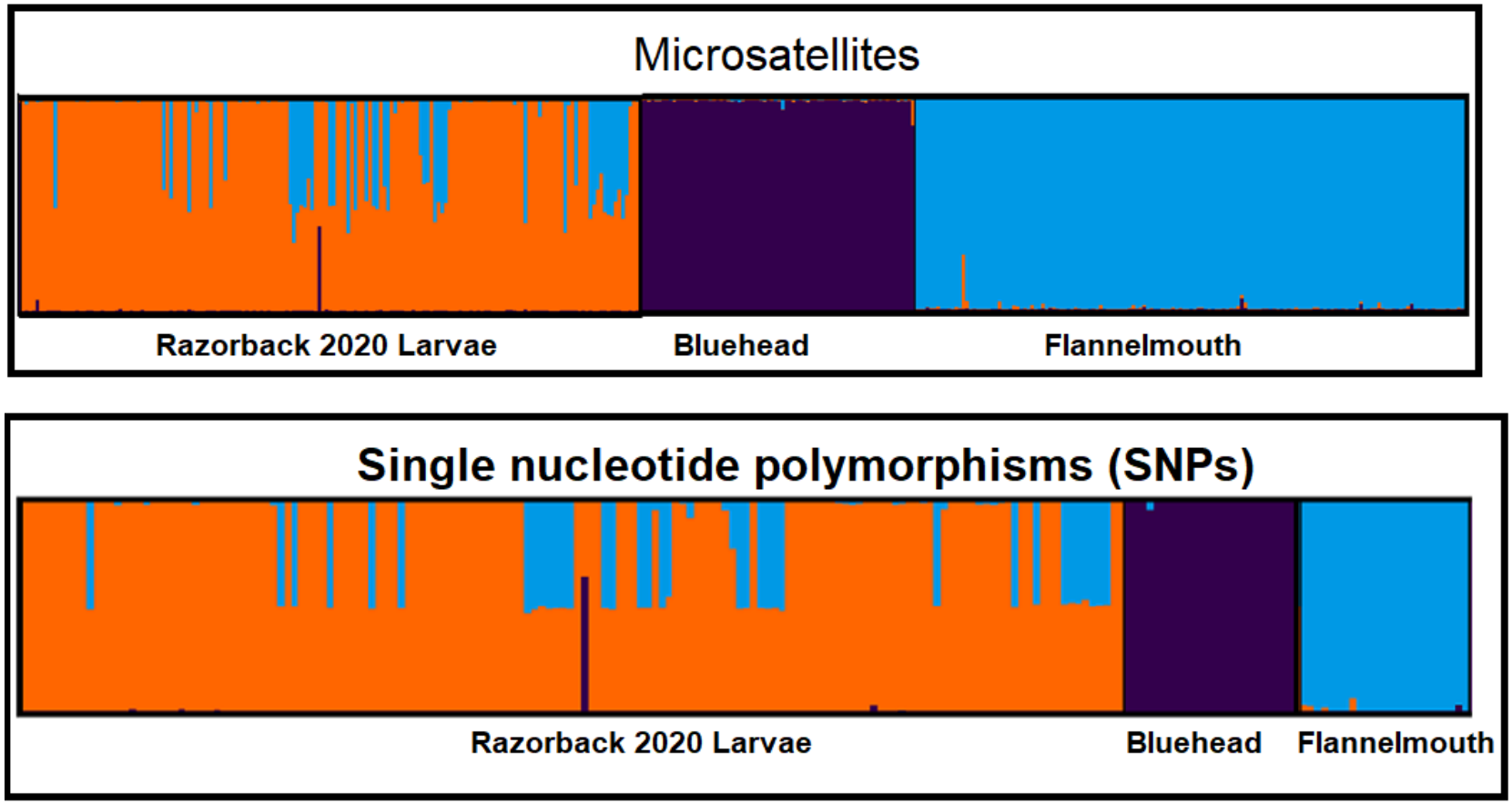

Figure 7.
The results of an Admixture hybrid analysis conducted in AdmixPipe and visualized in CLUMPAK for single nucleotide polymorphisms in Razorback Sucker (*Xyrauchen texanus*) larvae collected from the San Juan River in 2020. Each individual is represented by a single vertical bar, with the proportion of color in each bar representing the estimated proportion of ancestry attributed to each of three of three genetic clusters. Orange represents Razorback Sucker, purple represents Bluehead Sucker (*Catostomus discobolus*), and blue represents Flannelmouth Sucker (*C. latipinnis*).

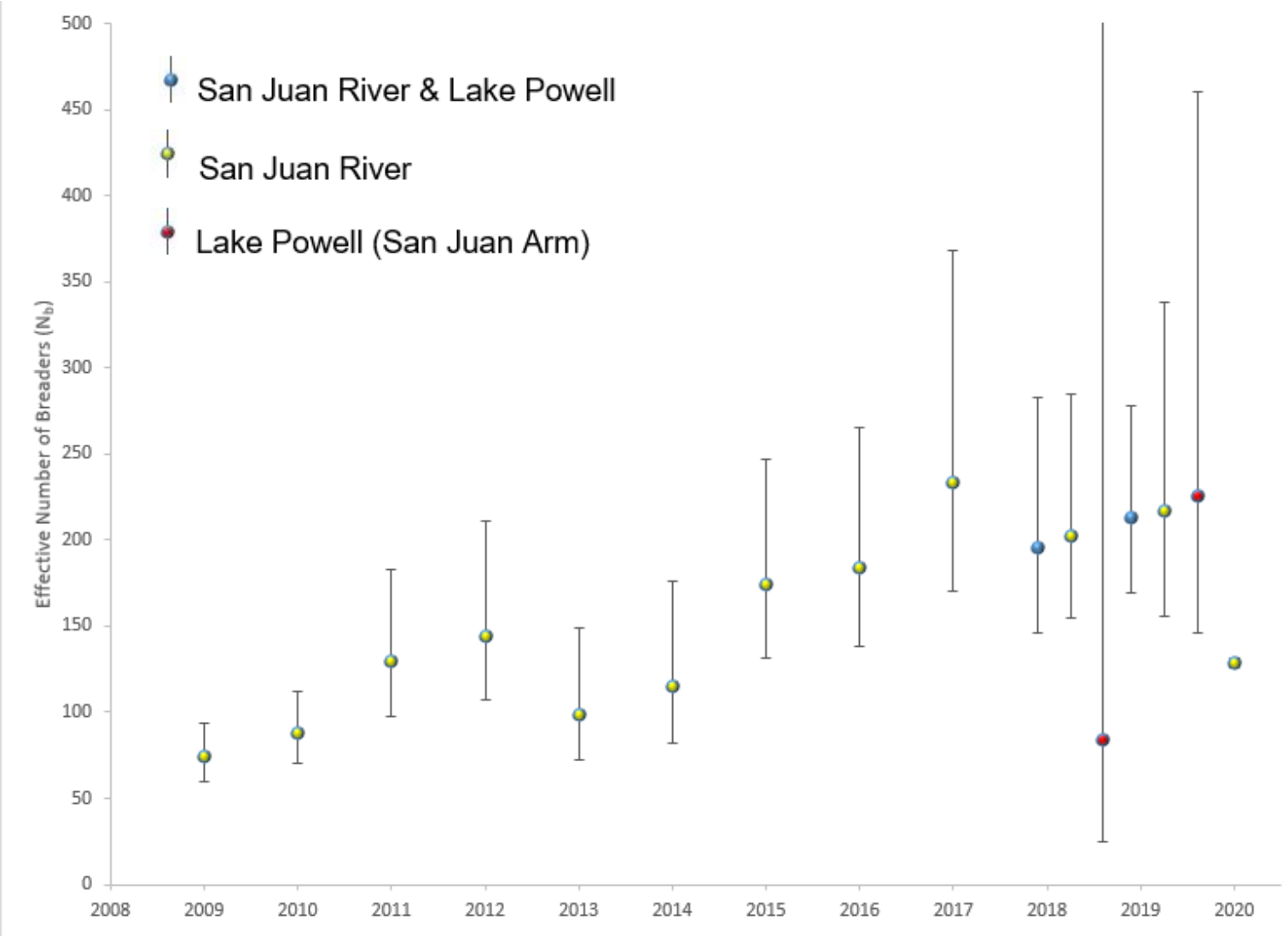

Figure 8.
Effective number of breeder (N_b) estimates for Razorback Sucker (*Xyrauchen texanus*) for 2009-2020 calculated with microsatellites (2009-2019) and SNPs (2020). The 2018 and 2019 samples are broken down by below the waterfall (Lake Powell), above the waterfall (San Juan River) and both combined.

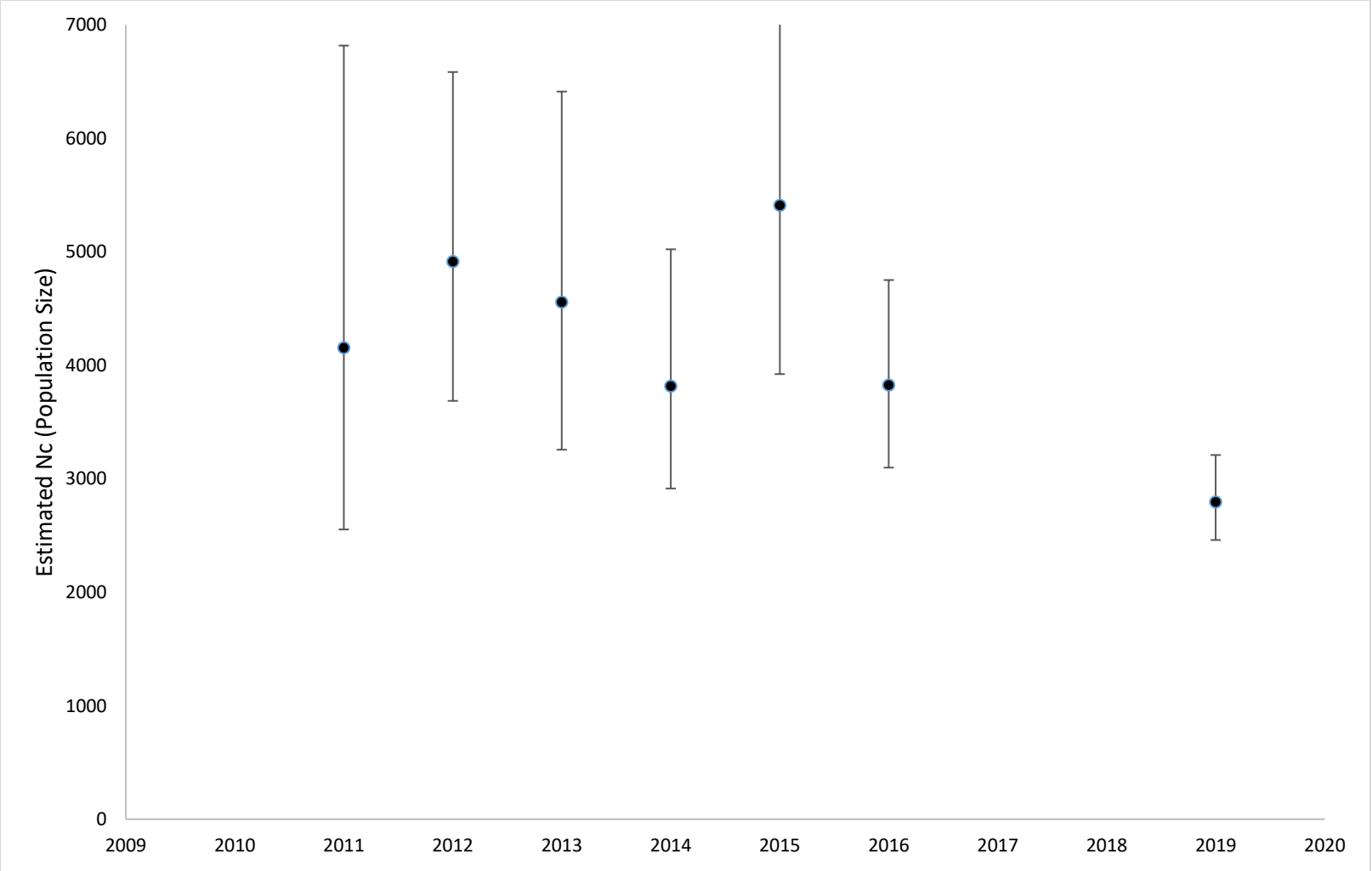

Figure 9.
Estimated adult census sizes (N_c) per year for Razorback Sucker (*Xyrauchen texanus*). Circles represent mean adult census sizes and bars represent 95% confidence intervals.

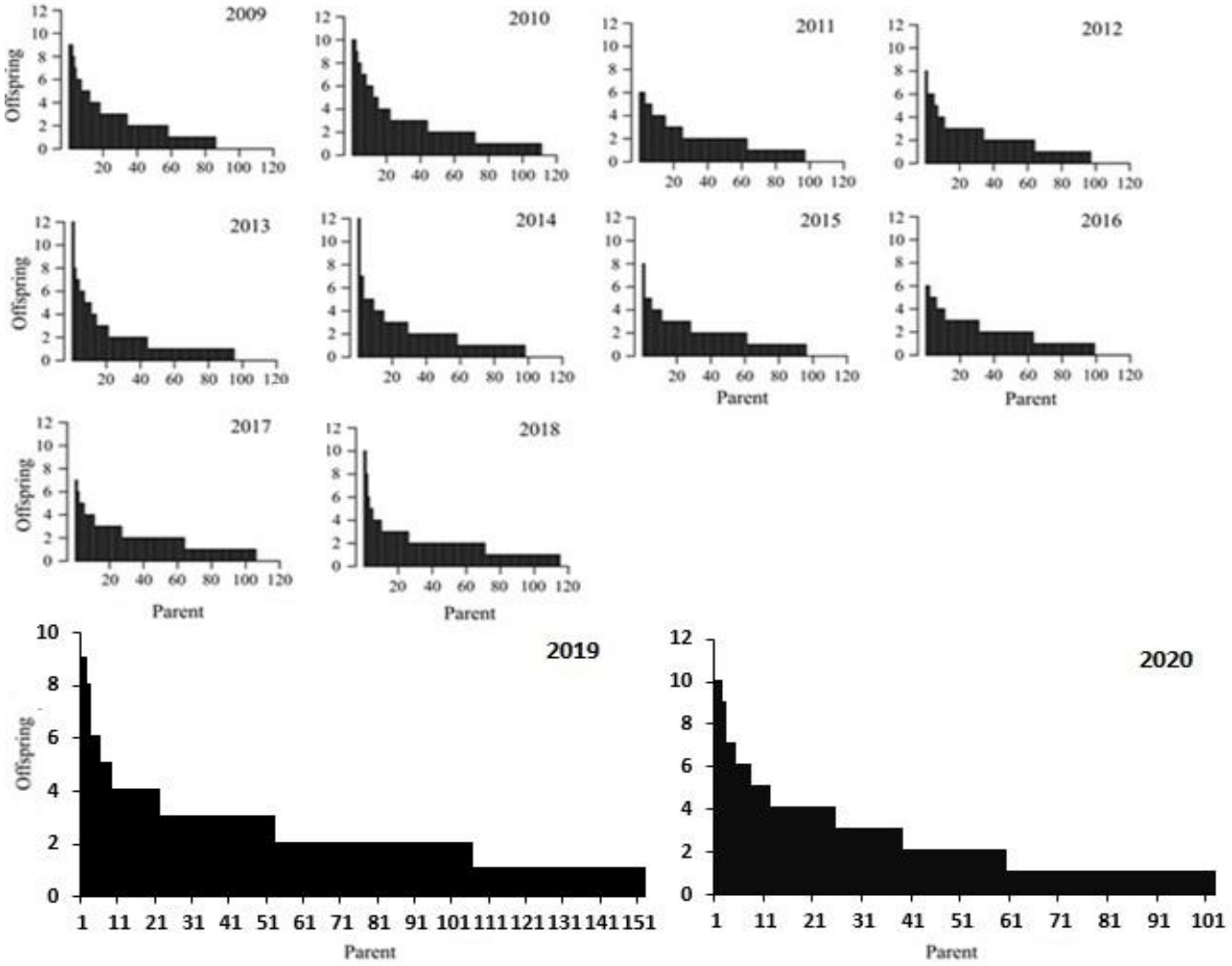

Figure 10.
Frequency of offspring contributed by individual Razorback Sucker (*Xyrauchen texanus*) parents from the San Juan River spanning 2009-2020 calculated with microsatellites (2009-2019) and SNPs (2020).

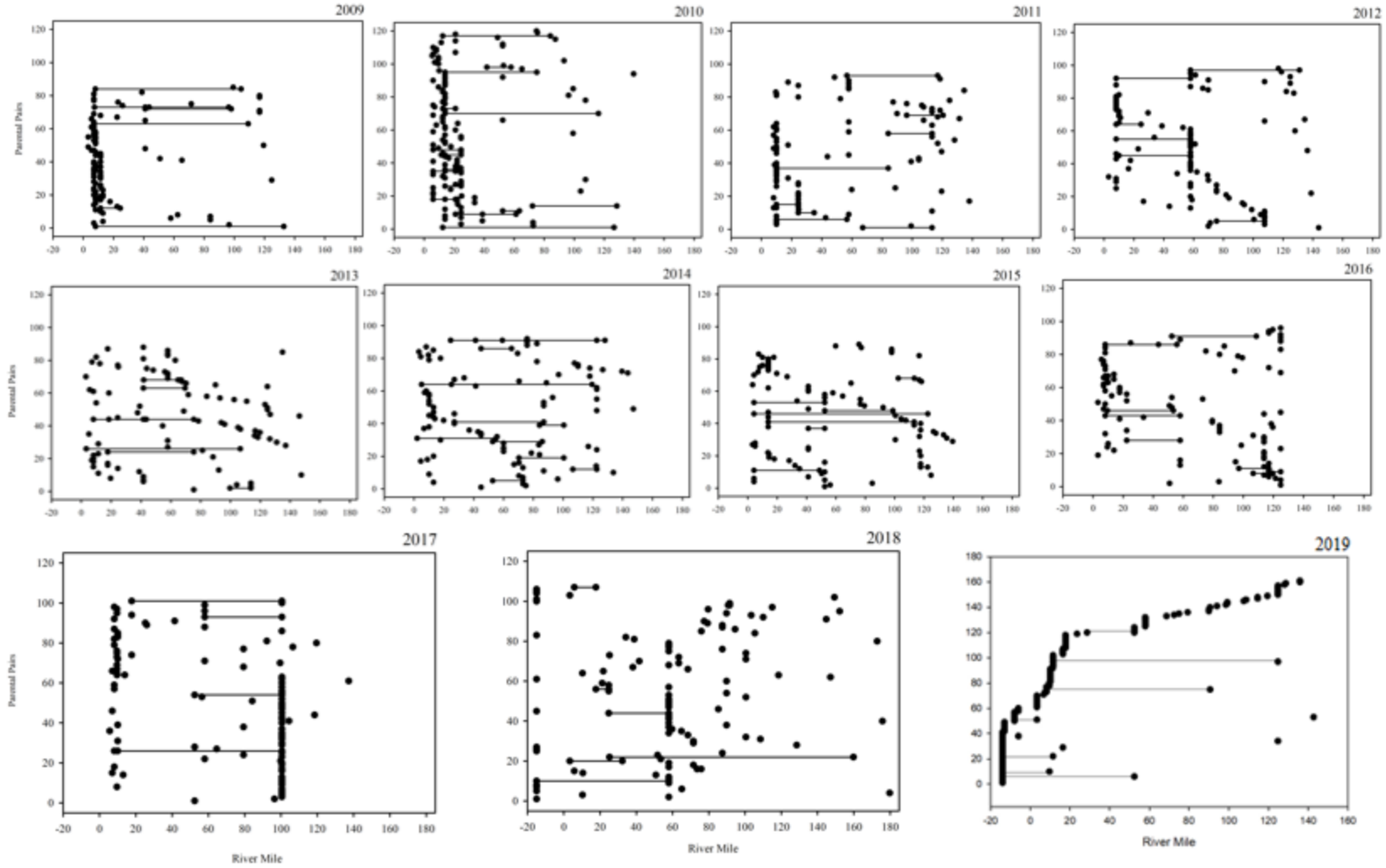

Figure 11.
Estimated number of parental pairs for Razorback Sucker (*Xyrauchen texanus*) and the spatial distribution of the larval samples calculated with microsatellites (2009-2019) and SNPs (2020). Each circle represents a larval Razorback Sucker with lines connecting circles indicating full-sibling relatedness. Spatial distribution of samples is shown by plotting larvae according to the river mile at which they were collected (X axis).

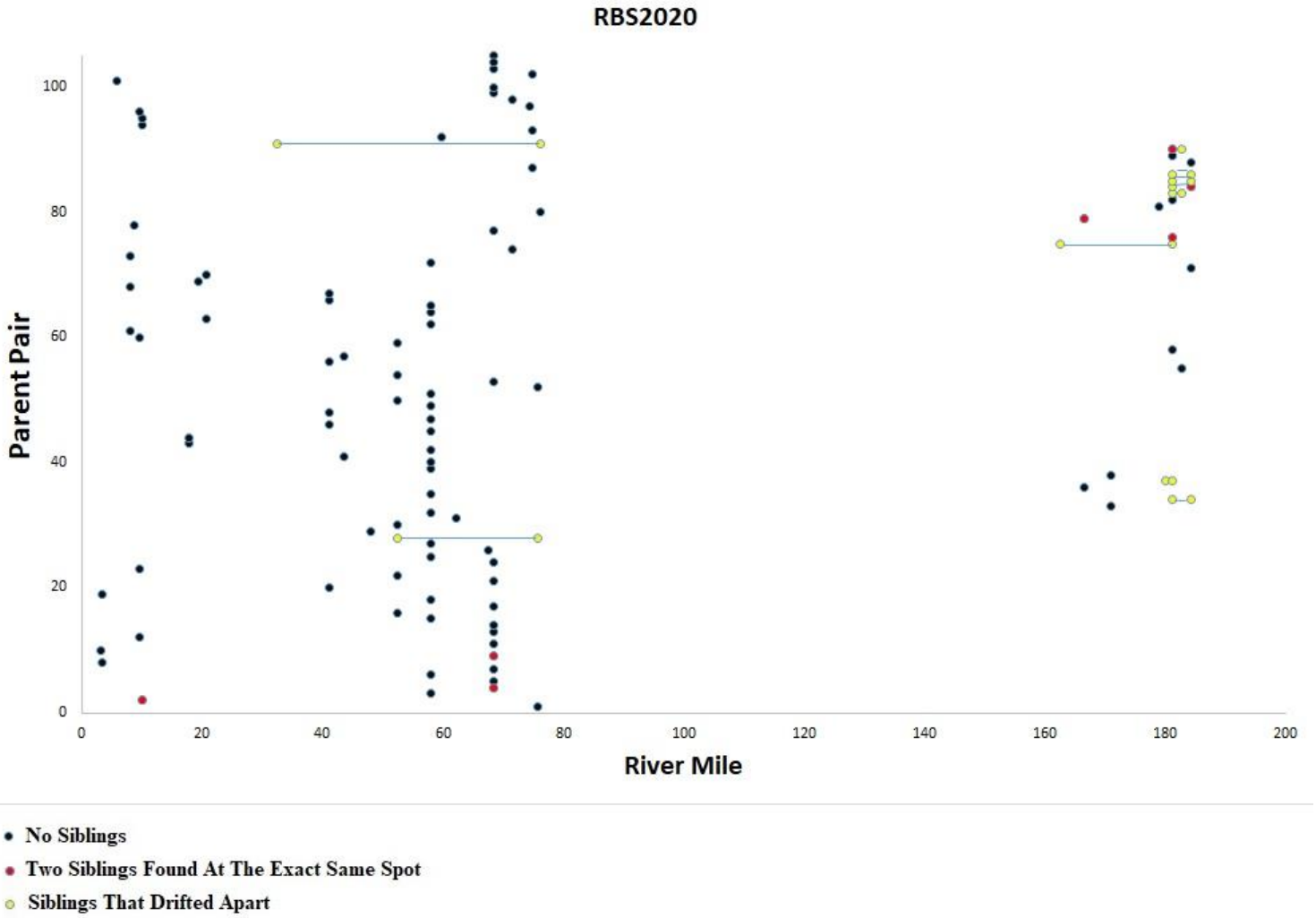

Figure 12.
Estimated number of parental pairs for Razorback Sucker (*Xyrauchen texanus*) and the spatial distribution of the larval samples calculated with SNPs. Each circle represents a larval Razorback Sucker with lines connecting circles indicating full-sibling relatedness. Spatial distribution of samples is shown by plotting larvae according to the river mile at which they were collected (X axis).