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
May 2022

Final Report For: United States Bureau of Reclamation Agreement No. R17PG00086

Title: Genetic characterization of refuge, reestablished, and natural populations of the Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) in Arizona

Principal Investigators: Steve Mussmann, Alexis Harrison, and Wade Wilson; Southwestern Native Aquatic Resources & Recovery Center, Dexter, NM

Co-Invest./Cooperators: Doug Duncan[1], William Stewart[2], Kent Mosher[2]

1 U.S. Fish and Wildlife Service
2 U.S. Bureau of Reclamation

Executive Summary:

Gila Topminnow (*Poeciliopsis occidentalis occidentalis*), once considered one of the most abundant fishes of the lower Colorado River basin, has suffered severe population declines over the past century. The consequences of these declines and the imprint of historical biogeography on its genome were evaluated previously, yet those efforts did not address concerns about the genetic status or effective size of the many re-established and refuge populations. This study expands upon previous efforts by adding 21 novel microsatellite loci to existing marker panels. Sampling efforts covered a greater number of localities, including captive and wild stocks (43 localities, 2,065 samples), to conduct a more complete genetic evaluation of U.S. populations. Combined, these actions allow for characterization of genetic diversity within and among populations at greater resolution than previously possible. Population structure analysis verified the recorded ancestry of most sampling localities, with just five localities differing from management records. These included Deer Valley High School, the Phoenix Zoo, Secret (Gatewood) Spring, and two captive populations held at the Arizona State University (ASU) Animal Facility. The most surprising of these deviations showed introgression of Yaqui Topminnow (*P. o. sonoriensis*) into the captive ASU Monkey Spring population. Gila Topminnow lineages were also evaluated to identify population structuring within each lineage. This revealed relative homogeneity within several lineages (Bylas Springs, Sharp Spring, and both subdivisions of Sonoita Creek), but a high degree of sub-lineage population structure in Monkey and Cottonwood Springs lineage, as well as the Santa Cruz River. Headwater Livebearer (*P. monacha*) mitochondrial DNA haplotypes were also discovered in three wild populations in the Santa Cruz River, two of which appeared near the international border with Mexico (Potrero Creek and the Santa Cruz River near Nogales, AZ) while the third occurred in the Santa Cruz River at Tucson, AZ. The Headwater Livebearer haplotype individuals were prevalent in this river system, accounting for 76% of all fish sampled from these three sites. Results of genetic analyses were consistent with the hypothesis that these fish represent the hemiclonal species *P. monacha-occidentalis*.

Loss of neutral genetic diversity was apparent in all lineages as evidenced by fixation of alleles at several microsatellite loci in different lineages. This was extreme in some lineages, particularly Monkey and Cottonwood Springs, in which nine loci were consistently monomorphic at all of its localities, with some localities having up to 20 (of 28) loci fixed for a single allele. Fixation of alleles is indicative of founder events and genetic bottlenecks; however, few tests for bottleneck events were statistically significant. This may be a consequence of tests excluding monomorphic loci, which leads to a reduction in power for accurately detecting bottleneck events. Effective population size estimates experienced similar issues due to these monomorphic loci. Finite point estimates were calculated for most populations, but some were infinite, and wide confidence intervals surrounded most point estimates. Eleven localities had effective population sizes lower than 50, indicating they face immediate threats from genetic drift and inbreeding. Furthermore, the lower bound of the 95% confidence interval for effective size was below 50 for 30 sites. The overall results indicate that careful genetic management of the species will be necessary going forward to increase gene flow among localities representing the same genetic lineages, and establish new populations using methods that minimize impacts on genetic diversity of existing populations.

Background:

The Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) is a small (25-45mm), viviparous fish that was once one of the most abundant fishes of the lower Colorado River basin (Hubbs and Miller 1941; Minckley and Marsh 2009). As the northernmost representative of its genus, Gila Topminnow prefers relatively warm, shallow waters, but also tolerates an exceptionally wide range of environmental conditions. It has been recorded living at temperatures ranging from near freezing to 37°C, pH values from 6.6-8.9, salinities from fresh to seawater, and in waters with a wide range of dissolved oxygen content (Heath 1962; Schoenherr 1974; Meffe et al. 1983). It is even able to tolerate sites that briefly run dry by burrowing into the mud for short periods (Meffe et al. 1983).

Despite its broad environmental tolerances and historical abundance, Gila Topminnow has suffered severe declines over the past century, primarily as a result of habitat degradation and the introduction of nonnative species (Minckley 1999). The introduction of the Western Mosquitofish (*Gambusia affinis*) has been particularly problematic, as this species will both compete with the Gila Topminnow for space and directly predate on smaller individuals. When Gila Topminnow was first listed under the Endangered Species Preservation Act (USDI 1967), only two extant natural populations were known to exist in the United States (Minckley 1969). However, at this time natural populations existed in 13 sites in Arizona, the majority of which had not yet been discovered (Simons et al. 1989). Several populations have since been established at captive sites and at refuge sites in the wild. Six populations were also maintained as refuges at the Arizona State University (ASU) Animal Facility; however, two were recently lost due to natural mortality (Bylas Springs and Sharp Spring) and a third was destroyed due to hybridization with another species (Monkey Spring). Three remaining Gila Topminnow populations and a single Yaqui Topminnow (*P. o. sonoriensis*) population were removed from ASU in April 2022. Parker Canyon was brought to the Aquatic Research and Conservation Center (Arizona Game and Fish Department), while the remaining three (Cienega Creek, Red Rock Canyon, and Yaqui Topminnow) were stocked into ponds (Kent Mosher, USBR, personal communication).

Several previous studies examining the population genetics of Gila Topminnow have focused on identifying the degree of differentiation among the 10 extant natural populations initially identified in Arizona using microsatellite loci (Parker et al. 1998, 1999; Hedrick et al. 2001), an MHC locus (Hedrick and Parker 1998; Hedrick et al. 2001), and sequence variation in three mitochondrial genes (Hedrick et al. 2006). These studies reached somewhat different conclusions: microsatellite and MHC loci identified two evolutionarily significant units (ESUs), separating Monkey and Cottonwood Springs from all other sites, the latter of which were further subdivided into four management units (MUs). In contrast,

mitochondrial sequences showed no differentiation among populations (Hedrick and Hurt 2012). A relatively small number of microsatellite loci (5-7) were available for previous studies, and each locus had few alleles. This limited the resolution of population genetic analyses, including estimates of genetic diversity within and among populations, detection of heterozygosity differing significantly from expected values, and reconstruction of the historical relationships among populations.

We expanded upon previous efforts to quantify Gila Topminnow diversity in two distinct ways. First, 21 novel microsatellite loci were developed and added to seven existing markers to create a new panel of 28 neutral loci for quantifying genetic diversity within and among populations at greater resolution than previously attainable. Second, sampling efforts were expanded to include refuges, reestablished populations, and extant natural populations, thus providing a more complete and contemporary analysis of Gila Topminnow populations within the United States. Data were analyzed for this report to 1) validate the genetic lineage of each sampling locality relative to management records, 2) evaluate population structure, 3) identify populations that have reduced genetic diversity due to genetic bottlenecks and founder events, and 4) identify hybrid populations. Data collected for this study were also compared to data collected by Hedrick et al. (2001) to determine whether populations have lost genetic diversity via drift or selection over the past 20 years. Analytical outcomes will ultimately form the basis for a genetic management protocol for captive stock and augmentation programs in the management of the species. This study also supports two recovery objectives in the draft revised recovery plan for Gila Topminnow (USFWS 1998), Task 4: Develop and implement genetic protocol for managing populations, and Task 5: Study life-history, genetics, ecology, and habitat of Gila Topminnow and interactions with nonnative aquatic species.

Materials and Methods:

Sample and Data Collection

The listing of Gila Topminnow under Endangered Species Act of 1973 includes both Gila Topminnow (*P. o. occidentalis*) and Yaqui Topminnow (*P. o. sonoriensis*). These two sister taxa are now commonly regarded as separate species in the scientific community (*P. occidentalis and P. sonoriensis*: Miller et al. 2005, Conway et al. 2019; Mateos et al. 2019a; 2019b). However, the subspecific designations of these taxa are retained in this report, meaning 'Gila Topminnow' refers solely to *P. o. occidentalis*, and 'Yaqui Topminnow' will refer only to *P. o sonoriensis*.

Samples were collected to represent all captive populations of Gila Topminnow and Yaqui Topminnow at ASU. Samples were also collected from extant wild, re-established, and refuge populations throughout the current range of Gila Topminnow in the United States (Figure 1). Fifty samples were targeted for collection during each of 45 sampling events at 43 localities. This goal was met for nearly all instances; however, some collection efforts fell short, and other samples were ultimately excluded from analysis due to excessive missing genotype data or species misidentification, resulting in localities with <50 samples (Table 1). Ultimately, 2,065 *Poeciliopsis* samples were genotyped from 43 localities at which collections occurred between 2017 and 2021 ($\bar{x}$ = 48.02 samples per locality). Whole fish were collected and stored in 95% ethanol. Genomic DNA was extracted from a tissue clip using either the DNeasy® Blood and Tissue Kit (Qiagen, Valencia, CA, USA) or Nucleospin® Tissue kit (Macherey-Nagel, Allentown, PA, USA) following standard protocols. All tissue samples were archived in 95% ethanol at -80°C.

Microsatellite development

Three samples of genomic DNA were sent to the University of Georgia's Savannah River Ecology Laboratory (SREL; Aiken, South Carolina) to identify microsatellite loci via Illumina sequencing. PAL_FINDER v0.02.04 was used to identify 6,725 microsatellite loci for which primers could be developed (Castoe et al. 2012). This candidate list was filtered to remove all imperfect repeats and loci

with repeat motifs < 4 base pairs (bp) in length. This resulted in 58 candidate loci for initial screening. These loci were amplified via polymerase chain reaction (PCR) and subjected to 1.5% agarose gel electrophoresis. Thirty loci that successfully amplified were selected for further testing. The forward primer for each locus was labeled with one of four fluorescent dyes (Dye set G5: Applied Biosystems®) and further evaluated for polymorphism. Twenty-one loci were ultimately selected for genotyping (Table 2).

Genotyping

The 21 novel microsatellite loci were combined with seven loci used in previous Gila Topminnow genetic evaluations (*Pooc-4-44*, *Pooc-G49*, *Pooc-G53*, *Pooc-OO56*, *Pooc-C15*, *Pooc-G10*, and *Pooc-LL53*: Parker et al. 1998). This yielded nine multiplex panels totaling 28 loci. Amplification via PCR was conducted in 10 µL reactions containing 0.6x Qiagen Multiplex Master Mix, up to 0.125 µM of each primer, and 1 µL of template DNA. Conditions for amplification consisted of an initial denaturation step at 95°C for 15 minutes followed by 35 cycles of denaturation at 95°C (45 s), annealing at 56°C (60 s), and extension at 72°C (60 s) with a final extension at 72°C for 30 minutes. Each forward primer was labeled with one of four fluorescent dyes (Dye set G5: Applied Biosystems®). Capillary electrophoresis was carried out on an ABI 3500XL Genetic Analyzer and all fragments were sized using LIZ-500 internal size standard (Applied Biosystems®). Loci were genotyped using GeneMapper® Software 5 (Applied Biosystems® Foster City, CA, USA). Scoring of microsatellite alleles was performed independently by two researchers. Ten percent of samples were re-amplified and scored by an independent party to verify data integrity.

A 581 bp region of cytochrome c oxidase subunit I (COI) was also sequenced for 767 samples. Samples were chosen for sequencing from populations in the Santa Cruz River system (at Nogales, AZ and Tucson, AZ) and Potrero Creek due to observation of unique microsatellite alleles private to these sites. Localities showing full or partial Santa Cruz River ancestry (Phoenix Zoo and captive Parker Canyon), those potentially containing non-*Poeciliopsis* species (Rio Salado Audubon Center and Deer Valley High School), and those containing hybrids (captive Monkey Spring) were also sequenced. Finally, Yaqui Topminnow and Gila Topminnow samples extracted at the same time as the above populations were sequenced as reference samples to assess COI diversity within Gila Topminnow. Polymerase chain reaction was conducted using the FishF2 and FishR2 COI barcoding primers (Ward et al. 2005) in 20 µL reaction volumes containing 8 µL 2x Qiagen® Multiplex Master Mix; 1 µL each of forward and reverse primers at 10 µM concentration, 1 µL DNA template, and 9 µL nuclease-free water. Amplification of COI was performed using the same thermal cycler settings as listed above for microsatellite amplification. Reactions were purified by combining Exonuclease I (Exo) and Shrimp Alkaline Phosphatase (SAP) using manufacturer protocols (New England BioLabs; Ipswich, MA, USA). Sequencing reactions were performed bidirectionally using BigDye v3.1 Terminator chemistry (Applied Biosystems®) according to manufacturer protocols. Sequence products were concentrated via ethanol precipitation, dried, and eluted in 10 µL HiDi-Formamide solution prior to capillary electrophoresis (ABI 3500XL Genetic Analyzer). Forward and reverse sequences for each sample were aligned and edited to verify base calls using the software Sequencher® v.5.1 (Gene Codes Corporation).

mtDNA analysis

Each mtDNA sequence was submitted to NCBI BLAST (Altschul et al. 1990) to verify species identifications and obtain reference sequences to include in a haplotype network. Reference COI sequences were downloaded for Desert Pupfish (*Cyprinodon macularius*: GenBank accession number MW300330.1), Western Mosquitofish (*Gambusia affinis*: AP004422.1), Gila Topminnow (*P. o. occidentalis*: HQ556953.1 and HQ556954.1), Yaqui Topminnow (*P. o. sonoriensis*: MK860197.1), and Headwater Livebearer (*P. monacha*: KX229692.1). Multiple sequence alignment of 773 sequences was

performed using default settings in MAFFT v.7.487 (Katoh and Standley 2013). The resulting alignment was visually inspected to ensure homologous bases were properly aligned. A haplotype network was constructed using the statistical parsimony method (TCS: Templeton et al. 1992) in Popart v1.7 (Leigh and Bryant 2015). Results of this analysis were used to remove misidentified species (Desert Pupfish and Western Mosquitofish) from the dataset and partition Santa Cruz River sites into sample groups comprised of either Gila Topminnow or Headwater Livebearer haplotypes.

Genetic Equilibrium and Diversity

Microsatellite genotypes were screened for deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) using Genepop v4.2 (Rousset 2008). Statistical significance was assessed using a Bonferroni-adjusted alpha value (LD $\alpha = 0.0001$; HWE $\alpha = 0.001$) to compensate for the increased risk of Type I error associated with multiple comparisons (Rice 1989). Data were evaluated for both heterozygote deficiency and excess.

Observed (H_O) and unbiased expected heterozygosity (H_E), in addition to the fixation index (F), was calculated for each sampling locality in GenAlEx v6.5 (Peakall and Smouse 2012). Allelic richness (A_R) was calculated via rarefaction in HP-Rare v1.1 (Kalinowski 2004, 2005) using the lowest number of observed alleles at a site for which at least 30 samples were obtained (Santa Cruz River at Tucson; N = 62). Private alleles were recorded at each sampling locality, and again for putatively 'pure' representatives of the Gila Topminnow lineages identified through mtDNA haplotype and population structure analysis (see below).

Population Structure

The program Structure v2.3.4 was applied in a hierarchical manner to distinguish both global and localized patterns of genetic diversity. The program was initially used to verify the recorded lineage of each sampling locality based upon delineations of previously designated conservation units (Hedrick et al. 2001). A cluster value (K) of 9 was applied to the entire dataset in Structure using the admixture model and assuming correlated allele frequencies (Pritchard et al. 2000; Falush et al. 2003). Eight clusters were assumed to exist in the dataset as representatives of Yaqui Topminnow, Monkey and Cottonwood Springs, Bylas Springs, Sonoita Creek (Fresno / Coalmine Canyon), Sonoita Creek (Red Rock Canyon), Sharp Spring, Cienega Creek, and the Santa Cruz River. The mitochondrial DNA analysis also revealed individuals with Headwater Livebearer mtDNA haplotypes among the Santa Cruz River localities, which were excluded from initial Structure analysis. Twenty-four Structure replicates were performed, each consisting of 250,000 Markov chain Monte Carlo (MCMC) generations of burn-in followed by 1,000,000 generations of data collection. Clumpak (Kopelman et al. 2015) was used to identify the best explanation of population structure among all replicate Structure runs.

Two methods were then utilized to quantify and evaluate sub-lineage population structure. First, an analysis of molecular variance (AMOVA: Excoffier et al. 1992) was calculated in Arlequin v3.5.2.2 (Excoffier and Lischer 2010) with sample localities clustered according to the K = 9 Structure output. Mixed populations were categorized based upon their highest overall cluster assignment value. Pairwise F_{ST} estimates of divergence were calculated for all sample locality pairs using 16,000 permutations to test for significance.

Secondly, the dataset was subdivided into three parts based upon genetic cluster assignments determined by the K = 9 Structure output. The first (Group 1) corresponded primarily to the Monkey and Cottonwood Springs lineage, with representatives of Yaqui Topminnow, Bylas Springs, and Cienega Creek lineage included to provide reference data for possible admixed localities (e.g., Scottsdale Community College and ASU captive Monkey Spring). The second (Group 2) was composed of Bylas Springs, Sonoita Creek (Red Rock Canyon), and Sharp Spring lineages. The third (Group 3) was comprised of Cienega Creek, Santa Cruz River, and Sonoita Creek (Fresno / Coalmine Canyon) lineages. Individuals

from the Santa Cruz River that exhibited Headwater Livebearer mtDNA haplotypes were also added to Group 3. For each of these three datasets, Structure was used to evaluate K = 1 to 20 with 24 replicates at each K. All other parameters matched those used above for the complete dataset. Structure Harvester v0.6.94 (Earl and vonHoldt 2012) facilitated determination of optimal K values through comparison of –log likelihood values and Evanno's ΔK (Evanno et al. 2005). Clumpak was again utilized to assess multimodality among replicate runs, as well as execute Clumpp and Distruct to summarize and visualize Structure output (Rosenberg 2004; Jakobsson and Rosenberg 2007; Kopelman et al. 2015).

Effective Population Size and Genetic Bottlenecks

NeEstimator v2.1 (Do et al. 2014) was utilized to calculate effective population size (N_E) and 95% jackknife confidence intervals for each sample locality. Rare alleles were excluded from analysis (P_{Crit}) following recommendations of Waples and Do (2010). In summary, for sample localities with collection size of $N \geq 25$, $P_{Crit} = 0.02$ was utilized. For localities with $N < 25$, P_{Crit} was calculated as $1/(2N)$.

Multiple methods were applied to test for the presence of a recent genetic bottleneck in each sample group. The method of Cornuet and Luikart (1996) as implemented in Bottleneck v1.2.02 (Piry et al. 1999) was first used to evaluate each locus assuming mutation-drift equilibrium. The expected equilibrium heterozygosity (H_{EQ}) was calculated using 10,000 iterations. It was then determined whether H_E exceeded H_{EQ}, and evaluated for significance using the Wilcoxon signed rank test. This test differs from other tests of heterozygosity employed in this report by testing for an excess of heterozygosity ($H_E > H_{EQ}$) rather than testing for an excess of heterozygous individuals ($H_O > H_E$). Heterozygosity excess can occur in recently bottlenecked populations because the number of alleles in a population will be reduced more quickly than levels of heterozygosity (Cornuet and Luikart 1996; Piry et al. 1999). Results were evaluated for all model options (IAM: infinite alleles model; SMM: stepwise mutation model; TPM: two-phase model) due to persistent uncertainty surrounding microsatellite mutation mechanisms (Oliveira et al. 2006; Amos 2016; Kosman and Jokela 2019). Significance was assessed using a two-tailed P-value that was Bonferroni-adjusted ($\alpha = 0.001$) to compensate for the increased risk of Type I error associated with multiple comparisons (Rice 1989). The presence of a bottleneck was also evaluated by checking for a "mode shift" in allele frequency distribution away from the L-shaped distribution expected under mutation-drift equilibrium (Luikart et al. 1998).

Additional methods to identify genetic bottlenecks were carried out in Arlequin (Excoffier and Lischer 2010). The *M*-Ratio ($M = k/r$: Garza and Williamson 2001) assumes that a reduction in the number of alleles (k) will occur more quickly than a reduction in allele size range (r) when a stochastic event causes a decrease in population size. The *M*-Ratio therefore ranges from 0 to 1, with $M < 0.68$ indicating a bottleneck event when seven or more microsatellite loci are used in analysis (Garza and Williamson 2001).

Unfortunately, several monomorphic loci were detected in our dataset, and the above methods for detecting bottleneck events only use polymorphic loci in their calculations. A modified version of the *M*-Ratio [$M = k/(r+1)$] was used to compensate for this problem by removing the possibility of 'division by zero' errors that would occur if monomorphic loci were included under the original version (Garza and Williamson 2001; Excoffier et al. 2007).

Hybridization

Structure analysis indicated introgression of Yaqui Topminnow alleles into the ASU captive Monkey Spring population. To verify, the Bayesian clustering program NewHybrids v2.0 (Anderson and Thompson 2002) was used to explicitly test whether this population represents Gila Topminnow by Yaqui Topminnow hybrids. This was accomplished by utilizing the 'z' option in NewHybrids to assign captive Yaqui Topminnow (ASU) and Monkey Spring (extant wild population) as 'pure' populations. All other Monkey and Cottonwood Springs lineage individuals were included in the analysis (N=607;

Arizona-Sonora Desert Museum, Cold Spring, Cottonwood Spring, La Barge Canyon, Mud Spring, Tortilla Creek, Tule Creek, Unnamed Drainage #68b, and Walnut Spring #20), as well as six Yaqui Topminnow individuals from San Bernardino National Wildlife Refuge. The program calculated the probability of each sample belonging to one of six pre-defined categories: pure Gila Topminnow, pure Yaqui Topminnow, first generation (F1) Gila by Yaqui Topminnow hybrid, second generation (F2) Gila by Yaqui Topminnow hybrid, an F1 by pure Gila Topminnow backcross, or an F1 by pure Yaqui Topminnow backcross. The program was run for 1,000,000 MCMC generations of burn-in followed by 5,000,000 generations of data collection.

Results:

mtDNA Analysis

Most samples sequenced for COI yielded Gila Topminnow haplotypes (N=619) that identically matched the two reference Gila Topminnow samples downloaded from GenBank (Figure 2). Seven samples from Rio Salado Audubon Center and one from the Santa Cruz River at Tucson were exact matches for the reference Western Mosquitofish sequence. A single sample from Deer Valley High School was an exact match for Desert Pupfish. The Yaqui Topminnow from Arizona State University and San Bernardino National Wildlife Refuge exhibited haplotypes that were identical to one another, but deviated from the GenBank reference sequence at four nucleotides. All captive Monkey Spring individuals had Gila Topminnow haplotypes despite microsatellite data analysis revealing a hybrid ancestry for this population involving Yaqui Topminnow. A large proportion of samples from the Santa Cruz River at Tucson (31/36; 88.6%) and Nogales (41/50; 82%), as well as Potrero Creek (14/28; 50%), exhibited haplotypes identical to the GenBank Headwater Livebearer reference sequence. Gila Topminnow and Headwater Livebearer sample groups were analyzed separately for all Santa Cruz River sites in subsequent analyses.

Genetic Equilibrium and Diversity

Hardy-Weinberg deviations were rare, with most observations occurring among wild Santa Cruz River populations that displayed Headwater Livebearer mtDNA haplotypes. Each of these three populations (Santa Cruz River at Nogales and Tucson; Potrero Creek) exhibited deviation from HWE for at least 21 of 28 genotyped microsatellite loci (75%; Table 3). None of the groupings of individuals with Gila Topminnow mtDNA haplotypes at these localities exhibited deviation from HWE (Table 3). Only three other populations showed evidence of HWE deviation at any locus: Cold Spring (CSE) at *Pocc27*, Arizona-Sonora Desert Museum (ASDM) at *Pocc34,* and Yaqui Topminnow at *Pocc43* (Table 3).

Linkage disequilibrium was also exceptionally rare, with just four observations of deviations from equilibrium expectations. These included significant linkage of *Pocc25* and *Pocc27* for the wild Monkey Spring population, linkage of *Pocc09* and *Pooc-LL5* for the captive Yaqui Topminnow population, and linkage of *Pocc31* with *Pooc-C15* in both the captive Monkey Spring and Parker Canyon populations. Previously, we had noted significant linkage disequilibrium among 145 pairs of loci in the Santa Cruz River population at Nogales (Mussmann et al. 2020b); however, these linkage patterns disappeared in the reanalysis of the data because we considered groups of individuals with Gila Topminnow and Headwater Livebearer haplotypes separately in this updated report.

Measures of genetic diversity were generally low in all populations (Table 3). Expected heterozygosity ranged from a low of 0.072 in the wild Bylas Springs population to a high of 0.601 in the Potrero Creek population for the grouping of individuals with a Headwater Livebearer haplotype. With the exception of Gila Topminnow haplotype individuals from Potrero Creek (H_E = 0.552), the top seven H_E values were observed among collections of Yaqui Topminnow and populations known or suspected to contain individuals of mixed ancestry (captive Monkey Spring: H_E = 0.560; Potrero Creek Gila Topminnow haplotype: H_E = 0.552; Santa Cruz River at Nogales Headwater Livebearer haplotype: H_E = 0.467; Santa

Cruz River at Tucson Headwater Livebearer haplotype: $H_E = 0.415$; captive Yaqui Topminnow: $H_E = 0.411$; Yaqui Topminnow at San Bernardino National Wildlife Refuge: $H_E = 0.381$). A similar trend was noted for allelic richness (A_R), which ranged from 1.39 to 4.5. Five of the sites with the highest H_E values were among those with the highest A_R: Potrero Creek Gila Topminnow haplotype ($A_R = 4.5$), Potrero Creek Headwater Livebearer haplotype ($A_R = 4.39$), captive Monkey Spring ($A_R = 4.21$), captive Yaqui Topminnow ($A_R = 3.65$), and Santa Cruz River at Nogales Headwater Livebearer haplotype ($A_R = 3.57$). The lowest value was observed for Gila Topminnow haplotype individuals from the Santa Cruz River at Tucson ($A_R = 1.39$).

Several microsatellite loci were fixed (i.e., monomorphic) for a single allele in many populations (Figure 3) and many lineages had multiple fixed loci (Table 4). On average, each locality had 11.19 fixed loci out of 28 (39.96%), ranging from 0 (captive Monkey Spring) to 20 (Cottonwood Springs). The lack of fixed loci among captive Monkey Spring fish was assumed to be a consequence of its hybrid ancestry involving Yaqui Topminnow (see below). The Santa Cruz River at Nogales had just one fixed locus among its Headwater Livebearer haplotype individuals, while the Potrero Creek Headwater Livebearer and Gila Topminnow haplotype groupings each had two fixed loci. Yaqui Topminnow exhibited three fixed loci (but see caveats in Table 4). When the Santa Cruz River and Yaqui Topminnow localities were excluded, all other Gila Topminnow localities had nine or more fixed loci. The number of fixed loci observed at any individual locality (Table 3) tended to be greater than the number of fixed loci detected collectively from all localities representing its lineage (Table 4). For example, nine loci were consistently fixed across all Monkey and Cottonwood Springs lineage localities (excluding the captive Monkey Spring fish; Table 4). However, the number of fixed loci varied from 11 to 20 at any individual Monkey and Cottonwood Springs lineage locality (Table 3), indicating that different loci are fixed at different localities representing the same lineage.

Private alleles were detected in each lineage (Table 5), and several individual collection localities (Table 3). When considering individual lineages (Table 5), the Yaqui Topminnow exhibited the most private alleles (61). The next greatest number of private alleles was observed among all individuals with a Headwater Livebearer mtDNA haplotype (21). The Santa Cruz River Gila Topminnow haplotype exhibited 16 private alleles. Monkey and Cottonwood Springs lineage also exhibited a high number of private alleles (13), while all other lineages had four or fewer private alleles. Sonoita Creek (Red Rock Canyon) had the fewest of any lineage (2).

When sampling localities were evaluated individually for private alleles (Table 3), the captive Yaqui Topminnow had the greatest overall number of private alleles (13), while the Potrero Creek populations each had six. The next greatest number was found among captive Parker Canyon samples (5) while all other sample localities exhibited two or fewer private alleles. Most localities (N = 28) had no private alleles.

Population Structure

The program Structure confirmed the ancestry of most sampling localities when the dataset was evaluated for nine genetic clusters (Figure 4). However, there were five localities that did not conform to suspected histories. The captive Monkey Spring population at ASU indicated admixture with Yaqui Topminnow. Secret Spring was thought to represent a pure Bylas Springs lineage, however Structure indicated introgression with the Sonoita Creek (Red Rock Canyon) lineage. Deer Valley High School was also thought to be Bylas Springs lineage, but instead clustered with Cienega Creek lineage. This result was also recovered in our previous report (Mussmann et al. 2020b), and it was suspected to result from a potential sampling error, prompting fresh collection of Deer Valley High School in 2020. However, the 2020 samples evaluated herein also cluster with Cienega Creek. Finally, there was difficulty classifying Parker Canyon and the Phoenix Zoo samples. Captive samples from Parker Canyon (suspected Sharp Spring lineage) clustered with those from the Santa Cruz River rather than with Sharp Spring lineage

sites. This is an unusual pairing due to the large number of unique alleles found exclusively in the Santa Cruz River lineage. Phoenix Zoo exhibited a mixture of Santa Cruz and Sonoita Creek (Fresno / Coal Mine Canyon) lineage, but was also suspected to have some Sharp Spring ancestry.

Despite these assignment difficulties, genetic data and recorded histories of sites were mostly congruent. Most sites representing mixed lineages were accurately identified. For example, fish at the Rio Salado Audubon Center were accurately identified as a mixture of Sharp Spring and Sonoita Creek (Red Rock Canyon). Scottsdale Community College was also accurately reconstructed as a mixture of Bylas Spring, Cienega Creek, and Monkey Spring.

Table 6 shows the pairwise F_{ST} values for all localities calculated via AMOVA (F_{ST} = 0.52, P < 0.001). Genetic differences among lineages were relatively high (F_{CT} = 0.42, P < 0.001) compared to differences among sampling sites within lineages (F_{SC} = 0.17, P < 0.001). The greatest source of genetic variation was found within localities (47.73%) rather than among lineages (42.31%), however these values were similar, and the greater variation within localities is likely driven by the few admixed localities, most of which occurred artificially. The lowest source of genetic variation was found among localities within lineages (9.97%).

The patterns revealed through AMOVA were evident when tests for population substructure were conducted in the program Structure (Figure 5). These tests indicated that some lineages can be further subdivided into discrete entities, while others are homogenous. Evaluation of Evanno's ΔK values were used to identify the best K for each group. This method is biased towards selecting K = 2 as the best explanation of population structure (Janes et al. 2017), and this bias was indeed observed for all three groups of *Poeciliopsis* evaluated for this report. Therefore, the second highest ΔK was selected as the best explanation of population structure within each group. This yielded K values of 9, 3, and 9 for Groups 1, 2, and 3 respectively. In Group 1, Structure indicated population substructure among Monkey and Cottonwood Springs sites, with the 11 representative localities of this lineage being split among seven genetic clusters (Figure 5A). Wild and captive Bylas Springs and Cienega Creek representatives were included in this group for the purpose of evaluating the mixed ancestry of Scottsdale Community College (SCC) fish. The K = 9 Structure plot (Figure 4) had recovered some Cienega Creek ancestry among SCC fish; however, the Structure analysis of Group 1 (Figure 5A) assigned very little Cienega Creek ancestry to SCC.

Structure did not find any additional population substructure in Group 2 (Figure 5B) for Bylas Springs, Sonoita Creek (Red Rock Canyon) and Sharp Spring lineages. The Deer Valley High School populations were included in Group 2 because previous substructure analysis found Deer Valley High School to be a poor fit among Cienega Creek lineage (Mussmann et al. 2020b). In our updated population substructure analysis, the Deer Valley High School samples were assigned unambiguously to Bylas Springs lineage (Figure 5B).

Sonoita Creek (Fresno / Coal Mine Canyon) was homogenous in Group 3 (Figure 5C). However, the Cienega Creek lineage was split into three genetic clusters which rarely corresponded to a sampling locality. Most population structure observed in Group 3 was recovered among sites with Santa Cruz River ancestry. The captive Parker Canyon, Phoenix Zoo, and Gila Topminnow mtDNA haplotype samples from Potrero Creek, Santa Cruz River at Tucson, and Santa Cruz River at Nogales were all assigned to six separate genetic clusters. The genetic cluster assigned to the Santa Cruz River at Tucson was also found among the Cienega Creek populations. A seventh genetic cluster was found within the Santa Cruz River samples with Headwater Livebearer mtDNA haplotypes, however, each individual with this ancestry was also partially assigned ancestry from the Gila Topminnow individuals recovered at the same locality. This result indicates potential mixed ancestry (e.g., *P. monacha-occidentalis*) for all Headwater Livebearer mtDNA haplotype individuals recovered from the Santa Cruz River.

Effective Population Size

Estimates of N_E were highly variable among sites, ranging from 3 to infinity (Table 7; Figure 6). Eleven sampling events yielded $N_E < 50$, while an additional five had an $N_E < 100$. Sixteen sites had an $N_E > 500$, however 14 of these sites had infinite N_E values and very wide confidence intervals. N_E can also be viewed through the lens of the lower bound of the confidence interval, which provides a "floor" for the estimates when adequate numbers of polymorphic markers and samples are used (Waples and Do 2010). The lower bound of the 95% confidence interval (CI) ranges from 0.7 (Santa Cruz River at Tucson Gila Topminnow mtDNA haplotype) to 300 (Sabino Canyon), with 30 sampling events (62.5%) having lower 95% CI $N_E < 50$. An additional 11 had lower 95% CI $N_E < 100$, and only two sites had lower 95% CI $N_E > 200$. The upper bound of the 95% CI ranged from 20.1 to infinity, with one site having upper 95% CI $N_E < 50$ (2019 Lousy Canyon), and an additional site with upper 95% CI $N_E < 100$ (captive Parker Canyon). Overall, 10 sampling events yielded upper 95% CI $N_E < 500$. Thirty-seven of the 38 values above $N_E = 500$ (97.4%) were infinite.

Tests for recent genetic bottlenecks yielded few consistent results (Table 7). Wilcoxon signed rank tests typically yielded different results depending upon which mutation model was assumed for the data. Under the infinite alleles model (IAM), only captive Parker Canyon, captive Monkey Spring, Robbins Butte Wildlife Area, and Headwater Livebearer haplotype samples from Potrero Creek and Santa Cruz River at Tucson tested positive for a genetic bottleneck. In contrast, the two-phase mutation model (TPM) indicated a bottleneck for the captive Monkey Spring population, and the stepwise (SMM) mutation model only indicated a bottleneck for Cienega Creek – Pima County Preserve. The evaluation of the allele frequency distribution indicated a mode shift in 2018 and 2019 Lousy Canyon samples, Robbins Butte Wildlife Area, Gila Topminnow from Potrero Creek and Santa Cruz River at Tucson, and Headwater Livebearer haplotype samples from Santa Cruz River at Tucson.

The Garza-Williamson (G-W) *M* Ratio test (Table 7) indicated several more potential bottlenecks, with 12 sample groups (25%) showing evidence of a recent bottleneck. The modified G-W test, which in contrast to all other bottleneck detection methods employed here uses fixed loci in its calculations, showed that every site has experienced a bottleneck event (Figure 7).

Hybridization

NewHybrids revealed the presence of multigenerational Gila by Yaqui Topminnow hybrids among the ASU captive Monkey Spring fish (Figure 8). Four of 51 (7.8%) fish in this captive population were detected as pure Gila Topminnow. The remainder were multigenerational hybrids, with 33 (64.7%) classified as F2 hybrids and the remaining 14 (27.5%) as Gila Topminnow backcross hybrids. Additionally, one fish at the Arizona-Sonora Desert Museum, and a second within the ASU captive Yaqui Topminnow lineage were classified as F2 hybrids. All assignments were made with high confidence [Posterior probability (Pr) > 0.90] with two exceptions, both of which were F2 hybrids (Pr = 0.896 and Pr = 0.704).

Discussion

Genetic analysis of Gila Topminnow has revealed some concerning patterns that will need to be confronted through population management plans guided by genetic data. These issues are discussed below, beginning with the surprising lack of diversity observed in microsatellite data which indicate that diversity has been lost in both historical and contemporary time. Next, the within-lineage patterns of genetic diversity are discussed using examples from the Bylas and Monkey / Cottonwood Springs lineages to illustrate the processes that are impacting individual localities within lineages. Third, the effective population size and bottleneck results are evaluated, as are implications for short- and long-term impacts on Gila Topminnow populations. Finally, known and suspected hybrid populations are discussed with a special emphasis on the Santa Cruz River before presenting some preliminary management implications for the species.

A Holistic View

Overall patterns of population structure indicate that lineages have experienced isolation from one another over time. This is unsurprising given the threats that native fishes have faced in the lower Colorado River basin throughout the 20th and 21st centuries (Minckley and Deacon 1968) and the documented decline of once-common species throughout the Gila River basin (Olden and Poff 2005). The AMOVA results revealed that the greatest source of genetic variation was found within localities (47.73%), however variation among lineages remained high (42.31%) and variation among localities was low. Such numbers are close to the partitioning of within-species diversity observed among isolated endorheic basins (Mussmann et al. 2020a) rather than a species whose natural habitat was once connected via stream network (Meffe and Vrijenhoek 1988).

A loss of genetic diversity (i.e., the loss of alleles and sequence diversity) is an expected outcome of sudden demographic changes in natural populations (Nei et al. 1975; Tajima 1989). Overall patterns in allele frequencies are consistent with a historic bottleneck event that impacted the entire species at an indeterminate time point in the past, followed by additional bottleneck events that impacted individual lineages. This is backed up by the microsatellite loci that are fixed for the same allele in all or most populations, and the lack of diversity in mitochondrial DNA data observed in other studies (Hedrick et al. 2006), indicating a wide-ranging impact on the species. Subsequent bottlenecks are evident in individual lineages, with fixation occurring for different alleles in different lineages (Sonoita Creek Fresno / Coal Mine Canyon lineage; loci *Pocc02* and *Pocc28*: Table 4). Probable loss of diversity is also evident in contemporary time. For example, Sharp Spring lineage is now fixed for allele 159 at locus *Pooc-G49*. Previously this lineage had one additional allele detected at this locus as recently as the year 2000 (Hedrick et al. 2001).

Microsatellites typically exhibit high levels of polymorphism (Oliveira et al. 2006), making cases of multi-locus monomorphism in microsatellites very rare. However, this phenomenon has been documented in other animal species, and ascribed to either long-term low effective population size or sudden demographic changes (Aguilar et al. 2004). One such example verified that microsatellite alleles had been fixed for at least a century (Habel et al. 2008). Superficially, this seems detrimental to the long-term genetic health of a population. However, it is important to note that microsatellite loci are adaptively neutral (Oliveira et al. 2006), meaning they do not directly evaluate evolutionary or adaptive potential (Holderegger et al. 2006), but can elucidate population dynamics that may have also impacted adaptive loci (Westemeier et al. 1998). Unfortunately, we do not know the length of time for which Gila Topminnow has persisted with such low levels of diversity, but similarly low levels of heterozygosity were observed in many of these populations in the 1990s (H_O = 0.071-0.309: Hedrick et al. 2012), indicating that genetic diversity loss may have occurred prior to its initial endangered species status listing in 1967. Although we have not assessed loci that may contribute to the adaptive potential of the species, the documented decline of neutral diversity in contemporary time remains a cause for concern (Spielman et al. 2004).

Within-Lineage Genetic Diversity

Patterns observed for fixed loci, private alleles and genetic diversity estimates make it apparent that genetic drift has impacted all lineages, in addition to collection sites within lineages. This is noted despite a general lack of statistically significant genetic bottleneck tests (Modified G-W test excluded). Sampling localities representing different lineages have been impacted to varying degrees as well. The Monkey and Cottonwood Springs lineage represents one of the best documented instances of this phenomenon. A total of nine loci are fixed for the same allele across all ten sample sites that are putatively 'pure' representatives of this lineage based upon Structure (Table 4; Figure 4) and NewHybrids results (Figure 8). However, the number of fixed loci at each sampling locality ranges from

11 to 20 (Table 3), meaning that each of these sites is fixed for a minimum of two loci that are variable in other representatives of this lineage.

This variable fixation of alleles at different localities that are descended from the same lineage is most likely is a consequence of founder events and a lack of within-lineage gene flow. For example, locations such as Tortilla Creek (TTC) and Unnamed Drainage #68b (D68B) were populated by unintentional release of Gila Topminnow from a holding tank (Mesquite Tank #2) located upstream of these sites. Mesquite Tank #2 was stocked with 1,000 fish in 1982. Unnamed Drainage, where Gila Topminnow were first observed in 1985, was populated by fish from an unintentional release from Mesquite Tank #2. It is assumed that the Gila Topminnow population in lower Tortilla Creek, where topminnow were first detected in 2005, resulted from natural dispersal of fish from Unnamed Drainage. This is consistent with genetic data, as the population substructure analysis indicated these two sites form a distinct genetic cluster within the Monkey and Cottonwood Springs lineage (TTC and D68B in Figure 5A). The population in Unnamed Drainage was last observed in 2017, and is considered extirpated.

Measures of genetic diversity for Tortilla Creek and Unnamed Drainage are consistent with the expectations of founder events and restricted gene flow (i.e., the stocking of their source population occurred in 1982, meaning no gene flow has occurred with the other eight localities representing their lineage). Here, Unnamed Drainage is monomorphic at 15/28 microsatellite loci, and Tortilla Creek is fixed for 16/28 loci. Neither locality exhibits any private alleles. Both have exceptionally low measures of genetic diversity (D68B H_E = 0.190; TTC H_E = 0.186; D68B A_R = 1.76; TTC A_R = 1.67). Unnamed Drainage has a surprisingly high N_E estimate (N_E = 660), but with a very wide confidence interval (95% CI = 35.3 - ∞). Tortilla Creek has a low N_E with a similarly wide confidence interval (N_E = 56; 95% CI = 18.3 - ∞). These metrics again show a decline from Unnamed Drainage to Tortilla Creek.

These patterns observed in the Monkey and Cottonwood Springs linage are overall consistent with multiple founder events, in that Unnamed Drainage contains a fraction of the diversity that exists collectively within the Monkey and Cottonwood Springs lineage, while Tortilla Creek contains a further reduction of diversity relative to Unnamed Drainage. However, a statistically significant bottleneck event could not be detected for these sites using any method except the modified G-W test. Similar trends are also apparent in all other Gila Topminnow lineages, where each lineage except for the Santa Cruz River contains a high base number of fixed loci (≥ 9: Table 4) with greater numbers of fixed loci occurring at individual sampling localities, and nearly every site exceeding its lineage's base number of fixed loci (Table 3).

Bylas Springs lineage also exhibits populations which have undergone extreme allele frequency changes. One such site (Deer Valley High School) was classified as Cienega Creek lineage in Structure analyses in both this report and our previous report (Mussmann et al. 2020b), despite records showing that it was founded from just 55 Bylas Springs lineage fish from the easternmost spring of the Bylas Springs Complex (known as 'S3') in 1999. The S3 population itself had likely undergone multiple founder events prior to the founding of the Deer Valley High School population because its history can be traced to Middle Spring of the Bylas Springs Complex, from which 300 fish were sourced in 1986 and moved to Roper Lake State Park. Just twenty fish were moved from Roper Lake to the Arizona State University in 1994 before an unknown number of fish were transplanted from ASU to S3 in 1996. The changes in allele frequencies resulting from this series of events has led to a Deer Valley High School population that is now coincidentally similar to Cienega Creek lineage sites (pairwise F_{ST} = 0.23-0.29; Table 6) but divergent from other Bylas Springs lineage sites (pairwise F_{ST} = 0.35-0.55; Table 6). However, the Deer Valley High School collections cluster with Bylas Springs in the absence of Cienega Creek lineage fish (Figure 5B). In our previous analysis of population substructure, the Deer Valley High School collection was distinct from all Cienega Creek sites (Figure 3C in Mussmann et al. 2020b). Consequently, we suspect Deer Valley High School has no Cienega Creek ancestry, and any similarities to Cienega Creek are

a consequence of the multiple founder events that exacerbated the effects of genetic drift in its population history between 1986 and present.

Effective Population Size

Generational effective population size estimates (N_E) are commonly used to inform population viability. They estimate the number of individuals in a population that effectively contribute to the next generation, and help quantify a population's potential for adaptation to changing environmental conditions (Jensen and Bachtrog 2011). Most threatened species experience the negative effects of genetic factors that are amplified by small population size (Spielman et al. 2004), however distinct cutoff points for when organisms will experience these effects are challenging to quantify. The long recognized 50/500 rule (Franklin 1980) is frequently used to provide context for N_E estimates by providing a guideline to assess short- and long-term genetic viability of a population. In other words, a minimum population of 50 is required to minimize the negative impacts of inbreeding, and a minimum of 500 individuals to combat loss of alleles due to genetic drift. However, this rule does not exist without controversy. Some researchers have criticized the applicability of the 50/500 rule to complex systems, and suggested the need for revised rules with greater minimum N_E thresholds (Jamieson and Allendorf 2012, 2013; Frankham et al. 2013). Regardless, this rule remains as a minimum baseline by which N_E estimates can be compared.

Few Gila Topminnow populations definitively exceed the 50/500 rule benchmarks for N_E estimates, meaning most are susceptible to long- or short-term effects of genetic drift and inbreeding. The analysis found several populations to have "infinite" population sizes (Table 7), however, these seemingly large population estimates cannot be taken at face value. The linkage disequilibrium method used for calculating N_E employs a correction to compensate for expected sampling error (Waples and Do 2008). Therefore, these values could be indicative of a large population size, or alternatively the correction employed in this method could by chance be greater than the estimator for which it is trying to compensate (Waples and Do 2010). This method can also experience difficulty distinguishing between large and truly infinite population sizes (Marandel et al. 2019). However, the lower bound of the 95% confidence interval will be a finite number when adequate data are available, and can be used for setting a lower bound on the population size (Waples and Do 2010). For Gila Topminnow, the lower bound of each site's 95% confidence interval is < 500, indicating none of the populations, even those with infinite N_E point estimates, can currently be excluded from these dangers.

The uncertainty and wide confidence intervals for some populations are likely driven by monomorphic and otherwise low-variability loci which are prevalent in many sites. Investigations of interactions between number of loci, locus variability, and sample size show that a decrease in any of these variables can impact precision of N_E estimates (Waples and Do 2010). Sample size was approximately constant across populations (N = 44 - 52 in 38/48 sample groups); however, the number of alleles present at each locus was highly variable. The linkage disequilibrium method disregards any monomorphic loci and ignores low frequency alleles (P_{Crit}) based upon user-specified input (Do et al. 2014). Therefore, nearly all sampling localities evaluated in this report were functionally assessed in these tests at fewer than the 28 total microsatellite loci used for genotyping. Thirty-three sample groups had ≥ 10 fixed microsatellite loci, meaning most estimates were based upon ≤ 18 loci, with some estimates being derived from as few as eight polymorphic loci (Figure 3; Table 3), thus yielding wide confidence intervals for N_E estimates in many populations.

Hybridization

Two populations displayed either unique allele frequencies or mtDNA haplotypes that warranted investigation to determine if they had mixed with non-Gila Topminnow populations. These were the ASU captive Monkey Spring population, and the wild populations sampled from Santa Cruz

River and Potrero Creek. NewHybrids was used to verify that all but four Monkey Spring fish resulted from a hybridization event that occurred with Yaqui Topminnow at least two generations ago, and involved hybrid offspring mating with pure Gila Topminnow individuals to produce individuals with varying degrees of hybrid ancestry. This trend was not observed in other Monkey and Cottonwood Springs lineage fish, with the exception of two samples in other populations which were identified as F2 hybrids (Figure 8). Somewhat surprisingly, no Yaqui Topminnow mtDNA was detected within the captive Monkey Springs population (Figure 2). However, mtDNA is prone to selective sweeps, and due to its strictly maternal inheritance pattern in most animals, mtDNA haplotypes introduced via hybridization could be purged from a population in as little as two generations (Wilson et al. 1985; Bazin et al. 2006).

Poeciliopsis monacha-occidentalis in the Santa Cruz River

A much more complex hybrid system was identified in the Santa Cruz River system. Wild Santa Cruz River populations exhibited a prevalence of Headwater Livebearer mtDNA haplotypes that outnumbered Gila Topminnow haplotypes at all localities except Potrero Creek. Hybridization in *Poeciliopsis* can lead to unique clonal lineages and variable ploidy (Mateos and Vrijenhoek 2005; Conway et al. 2019), with *P. monacha-occidentalis* being one of the known hemiclonal lineages that reproduces by hybridogenesis (Quattro et al. 1992). This is a form of unisexual reproduction in which the maternal genome is transmitted to offspring without recombination, but the paternal genome is discarded via pre-meiotic cell divisions (Lavanchy and Schwander 2019). Consequently, offspring of Gila Topminnow and *P. monacha-occidentalis* should have the appearance of being F1 hybrids (i.e., 50% Gila Topminnow and 50% Headwater Livebearer ancestry) regardless of the number of generations that have occurred since the initial hybridization event. Furthermore, all documented *P. monacha-occidentalis* are female, meaning all should exhibit Headwater Livebearer mtDNA haplotypes (Quattro et al. 1992). Lastly, because *P. monacha-occidentalis* would have a genetic fingerprint similar to F1 hybrids, we would anticipate deviations from Hardy-Weinberg equilibrium in these populations (Shockley 1973).

The above predictions regarding the hemiclonal *P. monacha-occidentalis* were observed among *Poeciliopsis* of the Santa Cruz River localities. Before conducting all microsatellite analyses, we separated Gila Topminnow haplotype individuals in these sites from Headwater Livebearer haplotype individuals. In Structure analyses, all Gila Topminnow haplotype individuals were assigned Gila Topminnow ancestry while Headwater Livebearer haplotype individuals were assigned partial Gila Topminnow ancestry and partial ancestry that was not observed in any other Gila Topminnow lineage (Figures 4; 5). Furthermore, Headwater Livebearer haplotype individuals showed evidence of deviation from Hardy-Weinberg equilibrium at nearly all loci, whereas no such deviations were observed among the Gila Topminnow haplotype individuals (Table 3). Additionally, no linkage disequilibrium was observed in these haplotype-based groupings, indicating they were more appropriate divisions of genetic diversity than lumping of samples purely by sample site. This is in contrast to our previous report, in which 145 of 406 locus pairs (35.7%) exhibited linkage disequilibrium for the Santa Cruz River sample locality at Nogales (Mussmann et al. 2020b). These prior deviations were likely caused by the Wahlund effect (Law et al. 2003). Therefore, we conclude from the results above that the Headwater Livebearer haplotype individuals observed at Potrero Creek, Santa Cruz River at Nogales, and Santa Cruz River at Tucson represent the hemiclonal species *P. monacha-occidentalis*.

Despite this conclusion, uncertainty surrounds the Gila Topminnow haplotype individuals observed in the Santa Cruz River. These individuals exhibited no detectable evidence of Headwater Livebearer introgression, and we therefore considered these individuals to be “Santa Cruz River” lineage Gila Topminnow. However, our analyses in this regard are limited by a lack of pure Headwater Livebearer reference samples. Collectively, these Santa Cruz River lineage fish exhibit only two monomorphic microsatellite loci, while all other Gila Topminnow lineages have at least nine monomorphic loci (Table 4). They also exhibit more private alleles (N=16) than all other Gila Topminnow

lineages (Table 5). Thus, we cannot exclude the possibility that some of these alleles come from other *Poeciliopsis* species. To appropriately test hypotheses surrounding the genetic ancestry of Santa Cruz River Gila Topminnow, it would be necessary to have population-level sampling of any *Poeciliopsis* species from neighboring drainages of the Santa Cruz in Mexico (e.g., Rio Concepción; Rio Sonora) and pure reference samples of Headwater Livebearer.

Conclusions and Management Implications

Through our assessment of captive and wild genetic diversity patterns of Gila Topminnow, multiple points are apparent that should be considered for future management of this species. The first is that care must be taken to minimize artificial hybridization among different *Poeciliopsis* species. For example, our analyses revealed that the captive Monkey and Cottonwood Springs lineage at ASU was no longer representative of its lineage. Methods for hybrid analysis revealed that 92% of the samples from this population were of mixed ancestry with Yaqui Topminnow. As a consequence, this captive population was destroyed (Marsh 2021).

Secondly, analyses indicate that genetic drift has played a major role in driving within-lineage divergence of populations, as identified through fixation of alleles at different loci within subpopulations of each lineage. This suggests that the current Gila Topminnow stocking practices, as well as transfer of fish among different subpopulations of the same lineage, should be evaluated and revised to maintain genetic equilibrium within the different Gila Topminnow lineages. Primarily, the effects of removing large numbers of individuals from one site to found a new site need to be evaluated for both the source and destination sites. For example, it may be more appropriate to found a new population of Monkey Spring lineage Gila Topminnow by combining 100 fish from each of five existing sites rather than removing 500 fish from a single site. This would minimize the impact on any one site, and allow for transfer of genetic diversity from multiple localities representing the same lineage. Additionally, gene flow among representative sites of the same lineage needs to be re-established to mitigate effects of genetic drift that have impacted many Gila Topminnow sites. If few replicates exist for a lineage, this could necessitate usage of Gila Topminnow from sites at which undesirable non-native species exist (e.g., Western Mosquitofish). Although managers currently try to avoid usage of these sites, current protocols are in place which require at least three experts to agree upon the identification of each translocated fish from a population known to contain Western Mosquitofish (Doug Duncan, USFWS, personal communication).

Finally, the implications of *P. monacha-occidentalis* presence in the Santa Cruz River need to be considered. Ideally, it would be helpful to know the timing of introduction of this species into the river (i.e., was it recent or have they persisted undetected in the Santa Cruz River system since historical times?), but the difficulty level for answering such questions often ranges from “challenging” to “impossible” and is highly dependent upon availability of suitable museum specimens. Current information suggests this is a recent introduction into the river, although most relevant work has relied upon morphological separation of *P. monacha-occidentalis* from Gila Topminnow via study of the dentition. The teeth of over 1,000 *Poeciliopsis* samples collected prior to 1970 were previously evaluated, and no evidence of *P. monacha-occidentalis* was found in the United States (Moore et al. 1970). Genetic evaluations of Gila Topminnow conducted in the Santa Cruz River during the 1990s also yielded no evidence of *P. monacha-occidentalis* (Phil Hedrick, ASU, personal communication). Therefore, this appears to be a recent, novel introduction. However, the most important issues to address will be their contemporary distribution and prevalence within the Santa Cruz River system, as well as their impact on sympatric Gila Topminnow populations in the Santa Cruz River given their reproduction method via hybridogenesis.

The above issues will be considered in greater depth as part of a forthcoming, comprehensive genetic management and translocation plan. This will be developed to propose methods that will

mitigate the impacts of genetic drift and founder effects by weighing the benefits and risks of different strategies. Overall, this will provide a guide to ongoing restoration efforts of this endangered species, specifically considering extant genetic diversity in the context of logistical issues, life history traits, biosecurity, and fish health concerns.

References:

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proceedings of the National Academy of Sciences of the United States of America* 101:3490-3494.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215(3):403-410.

Amos W (2016) Heterozygosity increases microsatellite mutation rate. *Biology Letters* 12:20150929.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160:1217-1229.

Barton NH, Otto SP (2005) Evolution of recombination due to random drift. *Genetics* 169:2353-2370.

Bazin E, Glemin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. *Science* 312(5773):570-572.

Castoe TA, Poole AW, de Koning APJ, Jones KL, Tomback DF, Oyler-McCance SJ, Fike JA, Lance SL, Streicher JW, Smith EN, Pollock DD (2012) Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS One* 7:e30953.

Conway KW, Mateos M, Vrijenhoek RC (2019) A new species of the livebearing fish genus *Poeciliopsis* from northern Mexico (Cyprinodontiformes: Poeciliidae). *ZooKeys* 883:91-118.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014.

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (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.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359–361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14(8):2611-2620.

Excoffier L, Laval G, Schneider S (2007) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47–50.

Excoffier L, Lischer HE (2010) Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10:564–567.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.

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

Frankham R, Brook BW, Bradshaw CJA, Traill LW, Spielman D (2013) 50/500 rule and minimum viable populations: response to Jamieson and Allendorf. *Trends in Ecology & Evolution* 28:187–188.

Franklin LR (1980) Evolutionary change in small populations. In: Soulé ME, Wilcox BA (eds) *Conservation Biology: an Evolutionary-Ecological Perspective*. Sinauer Associates, Sunderland, MA.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* 10:305–318.

Habel JC, Zachos FE, Finger A, Meyer M, Louy D, Assmann T, Schmitt T (2008) Unprecedented long-term genetic monomorphism in an endangered relict butterfly species. *Conservation Genetics* 10:1659-1665.

Heath WG (1962) Maximum Temperature Tolerance as a Function of Constant Temperature Acclimation in the Gila Topminnow (*Poeciliopsis occidentalis*). Ph.D. Dissertation, University of Arizona.

Hedrick PW, Hurt CR (2012) Conservation genetics and evolution in an endangered species. *Evolutionary Applications* 5:806–819.

Hedrick PW, Lee RN, Hurt CR (2006) The endangered Sonoran topminnow: Examination of species and ESUs using three mtDNA genes. *Conservation Genetics* 7:483-492.

Hedrick PW, Parker KM (1998) MHC variation in the endangered Gila Topminnow. *Evolution* 52:194–199.

Hedrick PW, Parker KM, Lee RN (2001) Using microsatellite and MHC variation to identify species, ESUs, and MUs in the endangered Sonoran topminnow. *Molecular Ecology* 10:1399–1412.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology* 21:797–807.

Hubbs CL, Miller RR (1941) Studies of the fishes of the order Cyprinodontes. VI. Materials for a revision of the American genera and species. *Miscellaneous Publication of the Museum of Zoology, University of Michigan* 19:1–87.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23(14):1801-1806.

Jamieson IG, Allendorf FW (2012) How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution* 27:578–584.

Jamieson IG, Allendorf FW (2013) A school of red herring: reply to Frankham et al. *Trends in Ecology & Evolution* 28:188–189.

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL (2017) The K = 2 conundrum. *Molecular Ecology* 26(14):3594-3602.

Jensen JD, Bachtrog D (2011) Characterizing the influence of effective population size on the rate of adaptation. *Genome Biology and Evolution* 3:687–701.

Kalinowski ST (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics* 5:539–543.

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology* Notes 5:187–189.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30(4):772-780.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15:1179–1191.

Kosman E, Jokela J (2019) Dissimilarity of individual microsatellite profiles under different mutation models: Empirical approach. *Ecology and Evolution* 9:4038–4054.

Lavanchy G, Schwander T (2019) Primer: hybridogenesis. *Current Biology* 29(1):R9-R11.

Law B, Buckleton JS, Triggs CM, Weir BS (2003) Effects of population structure and admixture on exact tests for association between loci. *Genetics* 164:381–387.

Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6:1110-1116.

Luikart G, Allendorf F, Cornuet J-M, Sherwin W (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* 89:238–247.

Marandel F, Lorance P, Berthelé O, Trenkel VM, Waples RS, Lamy J-B (2019) Estimating effective population size of large marine populations, is it feasible? *Fish and Fisheries* 20:189–198.

Marsh PC (2021) Histories of captive stocks of Gila Topminnow *Poeciliopsis occidentalis* and Yaqui (Sonoran) topminnow *P. sonoriensis* at Arizona State University. Final performance report for the period June 1, 2016 to April 30, 2021 for agreement no. R16AP00035. U.S. Bureau of Reclamation, Phoenix, AZ.

Mateos M, Vrijenhoek RC (2005) Independent origins of allotriploidy in the fish genus *Poeciliopsis*. *Journal of Heredity* 96(1):32-39.

Mateos M, Domínguez-Domínguez O, Varela-Romero A (2019a) A multilocus phylogeny of the fish genus *Poeciliopsis*: solving taxonomic uncertainties and preliminary evidence of reticulation. *Ecology and Evolution* 9(4):1845-1857.

Mateos M, Kang D, Klopp C, Parrinello H, García-Olazábal M, Schumer M, Jue NK, Guiguen Y, Schartl M (2019b) Draft genome assembly and annotation of the Gila Topminnow Poeciliopsis occidentalis. *Frontiers in Ecology and Evolution* 7:404.

Meffe GK, Hendrickson DA, Minckley WL, Rinne JN (1983) Factors resulting in decline of the endangered Sonoran topminnow *Poeciliopsis occidentalis* (Atheriniformes: Poeciliidae) in the United States. *Biological Conservation* 25:135–159.

Meffe GK, Vrijenhoek RC (1988) Conservation genetics in the management of desert fishes. *Conservation Biology* 2:157–169.

Minckley WL (1969) Native Arizona fishes, part 1 – livebearers. *Wildlife Views (Arizona Game and Fish Department)* 16:6-8.

Minckley WL (1999) Ecological review and management recommendations for recovery of the endangered Gila Topminnow. *Great Basin Naturalist* 59(3):230-244.

Minckley WL, Deacon JE (1968) Southwestern fishes and the enigma of "endangered species." *Science* 159:1424–1432.

Minckley WL, Marsh PC (2009) *Inland Fishes of the Greater Southwest: Chronicle of a Vanishing Biota*. University of Arizona Press, Tucson, AZ.

Moore WS, Miller RR, Schultz RJ (1970) Distribution, adaptation, and probable origin of an all-female form of Poeciliopsis (Pisces: Poeciliidae) in northwestern Mexico. *Evolution* 24(4):789-795.

Mussmann SM, Douglas MR, Oakey DD, Douglas ME (2020a) Defining relictual biodiversity: conservation units in Speckled Dace (Leuciscidae: *Rhinichthys osculus*) of the greater Death Valley ecosystem. *Ecology and Evolution* 10(19):10798-10817.

Mussmann SM, Harrison AS, Brown DP, Wilson WD (2020b) Genetic characterization of refuge, reestablished, and natural populations of the Gila Topminnow (*Poeciliopsis occidentalis*) in Arizona. Preliminary Draft Report submitted to the U.S. Bureau of Reclamation, Phoenix, AZ.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution* 29:1–10.

Olden J, Poff N (2005) Long-term trends of native and non-native fish faunas in the American Southwest. *Animal Biodiversity and Conservation* 28:75–89.

Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC (2006) Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology* 29:294–307.

Parker KM, Hughes K, Kim TJ, Hedrick PW (1998) Isolation and characterization of microsatellite loci from the Gila Topminnow (*Poeciliopsis o. occidentalis*) and their utility in guppies (*Poecilia reticulata*). *Molecular Ecology* 7:361–363.

Parker KM, Sheffer RJ, Hedrick PW (1999) Molecular variation and evolutionarily significant units in the endangered Gila Topminnow. *Conservation Biology* 13:108–116.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539.

Piry S, Luikart G, Cornuet J-M (1999) Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* 90:502–503.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Quattro JM, Avise JC, Vrijenhoek RC (1992) An ancient clonal lineage in the fish genus *Poeciliopsis* (Atheriniformes: Poeciliidae). Proceedings of the National Academy of Sciences of the United States of America 89(1):348-352.

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* 4:137-138.

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8:103–106.

Schoenherr AA (1974) Life History of the Topminnow, *Poeciliopsis occidentalis* (Baird and Girard) in Arizona, and an Analysis of Its Interaction with the Mosquitofish *Gambusia affinis* (Baird and Girard). Ph.D. Dissertation, Arizona State University.

Shockley W (1973) Deviations from Hardy-Weinberg frequencies caused by assortative mating in hybrid populations. *Proceedings of the National Academy of Sciences of the United States of America* 70(3):732-736.

Simons LH, Hendrickson DA, Papoulias D (1989) Recovery of the Gila Topminnow: a success story? *Conservation Biology* 3(1):11-15.

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America* 101:15261-15264.

Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics* 123:597–601.

Templeton AR, Crandall KA, Sing CF (1992) A cladistics analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132(2):619-633.

USDI (U.S. Department of the Interior) (1967) Native fish and wildlife. *Federal Register* 35:4001.

USFWS (U.S. Fish and Wildlife Service) (1998) Gila Topminnow, *Poeciliopsis occidentalis occidentalis*, Revised Recovery Plan. U.S. Fish and Wildlife Service, Albuquerque, NM.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications* 3:244–262.

Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources* 8:753–756.

Ward RD, Zemlak TS, Innes BH, Last PR, Herbert PDN (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360:1847-1857.

Westemeier RL, Brawn JD, Simpson SA, Esker TL, Jansen RW, Walk JW, Kershner EL, Bouzat JL, Paige KN (1998) Tracking the long-term decline and recovery of an isolated population. *Science* 282:1695-1698.

Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* 26:375-400.

Table 1. Sampling localities (Site) for Gila Topminnow (*Poeciliopsis occidentalis occidentalis*), Yaqui Topminnow (*P. o. sonoriensis*), and potential *P. monacha-occidentalis*. The number of samples (N), captive ® or wild (W) status, two-digit collection year, site locality code, and suspected lineage are provided for each locality.

Site	N	Status	Year	Code	Suspected lineage
Bylas Spring Complex	49	C	18	BYL	Bylas Springs
Bylas Spring Complex	51	W	19	BYL	Bylas Springs
Deer Valley High School (#113)	50	W	18	DVHS	Bylas Springs
Deer Valley High School (#113)	44	W	21	DVHS	Bylas Springs
Lower San Pedro River Preserve Pond	50	W	18	SPR	Bylas Springs
Secret Spring (Gatewood Spring)	49	W	19	SSP	Bylas Springs
Swamp Springs Canyon	50	W	19	SWC	Bylas Springs
Cienega Creek - Las Cienegas	50	W	18	CCLC	Cienega Creek
Cienega Creek - Pima County Preserve	50	W	18	CCPC	Cienega Creek
Cienega Creek	50	C	18	CNG	Cienega Creek
Road Canyon Tank (Las Cienegas NCA)	50	W	19	RCT	Cienega Creek
Sabino Canyon	50	W	18	SBC	Cienega Creek
Rio Salado Audubon Center	44	W	18	RSAC	Mixed
Scottsdale Community College	50	W	19	SCC	Mixed
Arizona-Sonora Desert Museum	50	W	19	ASDM	Monkey & Cottonwood Springs
Cold Spring (#85)	50	W	18	CSE	Monkey & Cottonwood Springs
Cottonwood Spring	50	W	18	CWS	Monkey & Cottonwood Springs
Unnamed Drainage (#68b)	50	W	17	D68B	Monkey & Cottonwood Springs
La Barge Canyon	50	W	17	LBC	Monkey & Cottonwood Springs
Mud Spring (#18)	50	W	19	MDS	Monkey & Cottonwood Springs
Monkey Spring	51	C	18	MNK	Monkey & Cottonwood Springs
Monkey Spring	50	W	19	MNK	Monkey & Cottonwood Springs
Tortilla Creek	50	W	17	TTC	Monkey & Cottonwood Springs
Tule Creek (#75)	50	W	19	TUC	Monkey & Cottonwood Springs
Walnut Spring (#20)	50	W	18	WS20	Monkey & Cottonwood Springs
Potrero Creek	14	W	21	POTC_O	Santa Cruz River
Santa Cruz River - North of Nogales	9	W	17	SCN_O	Santa Cruz River
Santa Cruz River - Tucson	4	W	20	SCT_O	Santa Cruz River
Phoenix Zoo	49	W	18	PZ	Santa Cruz River (Sharp Spring/Sonoita Creek)
AD Wash (#242)	51	W	18	ADW	Sharp Spring
Buckhorn Spring	50	W	19	BSP	Sharp Spring
Lime Creek (#301)	50	W	19	LMC	Sharp Spring
Parker Canyon	50	C	18	PRK	Sharp Spring
Robbins Butte Wildlife Area - Swimming Pool Tank	50	W	19	RBW	Sharp Spring
Sharp Spring	50	C	18	SHP	Sharp Spring
Coal Mine Canyon	50	W	19	CMC	Sonoita Creek (Fresno / Coal Mine Canyon)
Fresno Canyon	50	W	19	FCN	Sonoita Creek (Fresno / Coal Mine Canyon)
Lousy Canyon (#306)	10	W	18	LCN	Sonoita Creek (Fresno / Coal Mine Canyon)
Lousy Canyon (#306)	9	W	19	LCN	Sonoita Creek (Fresno / Coal Mine Canyon)
Larry Creek	37	W	19	LYC	Sonoita Creek (Fresno / Coal Mine Canyon)
Red Rock Canyon	52	C	18	RRK	Sonoita Creek (Red Rock Canyon)
Timbucktwo Tank	50	W	19	TMB	Sonoita Creek (Red Rock Canyon)
Walnut Spring (#392)	50	W	18	WS392	Sonoita Creek (Red Rock Canyon)
Potrero Creek	14	W	21	POTC_M	*P. monacha-occidentalis*
Santa Cruz River - North of Nogales	41	W	17	SCN_M	*P. monacha-occidentalis*
Santa Cruz River - Tucson	31	W	20	SCT_M	*P. monacha-occidentalis*
San Bernardino National Wildlife Refuge	6	W	20	SBNWR	Yaqui Topminnow
Tule and North Springs	50	C	18	TNS	Yaqui Topminnow

Table 2. Twenty-one novel microsatellite loci developed for Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) that cross-amplify in Yaqui Topminnow (*P. o. sonoriensis*) and *P. monacha-occidentalis*. The forward (F) and rever®(R) sequences are provided for each primer pair, along with the allelic size range measured in base pairs (BP), total number of alleles detected (N_A), and observed heterozygosity (H_O).

Loci		Primer Sequence	Size (BP)	N_A	H_O
Pocc02	F	CTAACCGAAGTCAGTGGCAAGC	187-207	5	0.206
	R	GGCTGAACAATGACAGAGGAGG			
Pocc03	F	GGAAGCAAGTCTAGAATTGACGC	177-196	6	0.170
	R	AGGAACAGTGTGCTCTTGTTAAGG			
Pocc07	F	AACATGAAGGCCTCACTTGC	165-177	3	0.052
	R	CCACAGGAGTAAGATGAAGAGAGC			
Pocc09	F	TATGTCGTTGTAGGTGCTCTGG	174-232	8	0.236
	R	ACCGGATGATCATTACAACAGG			
Pocc15	F	CAGGAACAGGAGCACTCATTGG	123-169	7	0.311
	R	TTCCAGGACAGAAGCGCTAACC			
Pocc16	F	CACAGGCACACGCATTAGAAGG	125-190	12	0.357
	R	AAGTAGGATGGCCTGGCAACG			
Pocc18	F	ATGCTCACAGCGCATCTGG	275-329	13	0.476
	R	TGCGCACCTGTTATTATTCCGC			
Pocc21	F	ACCAAGGCTCAGAATACACACC	115-131	4	0.061
	R	CAGAATCTGCGCCAGTCTGC			
Pocc25	F	GGAGGTGGCAGCATCATTACG	324-428	19	0.479
	R	CGCCACATCGTTACATAATAGACG			
Pocc26	F	GCTGCTTGCTTAAGAGTGCG	180-252	17	0.569
	R	ACCTTGCTATAACCTTGTTGCGC			
Pocc27	F	TCACCTTCAGTGTGAGTTCTCC	124-193	17	0.422
	R	CGGCTCCATCCTACTCCTATCC			
Pocc28	F	AGTCATCACATCACTGCTGGC	207-255	11	0.151
	R	ACTTGAGAATGAGCTATGCATGC			
Pocc29	F	ATTGACGACGATGGAACAAGCC	157-232	10	0.203
	R	AGACCGGTGACCTGTTCATGG			
Pocc31	F	CACACACAAGCCTCAACTTCTACC	122-173	9	0.173
	R	ACGTAAGAGGAGGAACACAGCC			
Pocc34	F	GAGTCACCGCTTCTCCACATCG	233-415	21	0.521
	R	GTCGTACAACAGGAGCCAGAGG			
Pocc38	F	TTGAGTGTATGTATGTGTCCATCC	182-309	22	0.492
	R	ATCAAGAGCATCAGAAGACCGG			
Pocc40	F	ATGATTACTGACATTCTGACCAGG	176-208	6	0.053
	R	ATAACGGCAATTCAAGAGCTGC			
Pocc43	F	AAGATGCTCAGCAATCACCACG	189-243	4	0.031
	R	TCTTCCTCTCCTCGCTGAGC			
Pocc44	F	GCAGTTATCACAGTTGCTTGTGC	263-403	24	0.554
	R	GCTGATAGACACGAGCAACTCC			
Pocc45	F	TGTGGAAGTAGAACCAACAACAGG	158-256	14	0.072
	R	GCAGTGAAGGTTCAACTCCAGC			
Pocc47	F	AGTTGCTTGTGGTTCTGACAGC	195-274	15	0.367
	R	GGAGACTGACTCATGACTGTCCG			

Table 3. Genetic diversity estimates for all Gila Topminnow (*Poeciliopsis occidentalis occidentalis*), Yaqui Topminnow (*P. o. sonoriensis*), and potential *P. monacha-occidentalis* localities in this report. Population names are represented by a combination of status, two-digit collection year, and locality code from Table 1. The number of monomorphic microsatellite loci (out of 28), observed (H_O) and expected (H_E) heterozygosity, fixation index (F), allelic richness (A_R), private alleles, Hardy-Weinberg equilibrium deviations (HWE) and instances of Linkage Disequilibrium (LD; out of 378) are also provided for each sampled population.

Population	Suspected Lineage	Monomorphic	H_O	H_E	F	A_R	Private	HWE	LD
C18BYL	Bylas Springs	16	0.125	0.128	0.005	1.64	0	0	0
W18DVHS	Bylas Springs	15	0.140	0.141	-0.015	1.88	0	0	0
W18SPR	Bylas Springs	16	0.093	0.100	0.029	1.77	0	0	0
W19BYL	Bylas Springs	17	0.076	0.072	-0.026	1.58	0	0	0
W19SSP	Bylas Springs	10	0.298	0.291	-0.030	2.63	0	0	0
W19SWC	Bylas Springs	18	0.081	0.083	0.044	1.55	0	0	0
W21DVHS	Bylas Springs	13	0.147	0.143	-0.037	1.97	1	0	0
C18CNG	Cienega Creek	13	0.228	0.226	-0.026	2.58	1	0	0
W18CCLC	Cienega Creek	14	0.248	0.245	-0.019	2.51	0	0	0
W18CCPC	Cienega Creek	9	0.230	0.242	0.153	2.92	2	0	0
W18SBC	Cienega Creek	13	0.255	0.248	-0.033	2.35	0	0	0
W19RCT	Cienega Creek	12	0.240	0.247	0.011	2.51	0	0	0
W18RSAC	Mixed	9	0.328	0.327	0.000	2.8	1	0	0
W19SCC	Mixed	9	0.282	0.290	0.004	2.55	0	0	0
C18MNK	Monkey & Cottonwood Springs	0	0.556	0.560	0.003	4.21	2	0	1
W17D68B	Monkey & Cottonwood Springs	15	0.194	0.190	-0.035	1.76	0	0	0
W17LBC	Monkey & Cottonwood Springs	11	0.297	0.291	-0.039	2.61	0	0	0
W17TTC	Monkey & Cottonwood Springs	16	0.187	0.186	-0.041	1.67	0	0	0
W18CSE	Monkey & Cottonwood Springs	13	0.298	0.285	-0.044	2.65	2	1	0
W18CWS	Monkey & Cottonwood Springs	20	0.141	0.134	-0.068	1.47	0	0	0
W18WS20	Monkey & Cottonwood Springs	11	0.321	0.310	-0.055	3.35	1	0	0
W19ASDM	Monkey & Cottonwood Springs	11	0.223	0.238	0.199	2.58	0	1	0
W19MDS	Monkey & Cottonwood Springs	13	0.226	0.214	-0.058	2.17	0	0	0
W19MNK	Monkey & Cottonwood Springs	11	0.301	0.302	-0.021	3.02	0	0	1
W19TUC	Monkey & Cottonwood Springs	11	0.290	0.294	0.024	2.72	0	0	0
W17SCN_M	*P. monacha-occidentalis*	1	0.571	0.467	-0.038	3.57	1	23	0
W20SCT_M	*P. monacha-occidentalis*	4	0.536	0.415	-0.186	2.57	0	22	0
W21POTC_M	*P. monacha-occidentalis*	2	0.564	0.601	0.025	4.39	6	21	0
W17SCN_O	Santa Cruz River	7	0.317	0.355	0.047	2.57	2	0	0
W20SCT_O	Santa Cruz River	19	0.152	0.143	-0.209	1.39	0	0	0
W21POTC_O	Santa Cruz River	2	0.561	0.552	-0.048	4.5	6	0	0
W18PZ	Santa Cruz River (Sharp Spring/Sonoita Creek)	8	0.358	0.360	-0.007	3.36	2	0	0
C18PRK	Sharp Spring	8	0.398	0.385	-0.051	3.18	5	0	1
C18SHP	Sharp Spring	9	0.296	0.291	-0.019	2.48	0	0	0
W18ADW	Sharp Spring	11	0.255	0.266	0.037	2.33	1	0	0
W19BSP	Sharp Spring	11	0.267	0.264	-0.025	2.24	0	0	0
W19LMC	Sharp Spring	12	0.214	0.217	0.009	2.06	0	0	0
W19RBW	Sharp Spring	11	0.269	0.272	0.008	1.98	0	0	0
W18LCN	Sonoita Creek (Fresno / Coal Mine Canyon)	14	0.221	0.228	-0.036	1.71	1	0	0
W19CMC	Sonoita Creek (Fresno / Coal Mine Canyon)	9	0.257	0.256	-0.015	2.24	1	0	0
W19FCN	Sonoita Creek (Fresno / Coal Mine Canyon)	11	0.253	0.263	0.016	2.34	1	0	0
W19LCN	Sonoita Creek (Fresno / Coal Mine Canyon)	14	0.155	0.184	0.054	1.64	0	0	0
W19LYC	Sonoita Creek (Fresno / Coal Mine Canyon)	12	0.234	0.234	0.022	1.92	0	0	0
C18RRK	Sonoita Creek (Red Rock Canyon)	17	0.178	0.168	-0.070	1.76	0	0	0
W18WS392	Sonoita Creek (Red Rock Canyon)	15	0.167	0.169	-0.010	1.86	1	0	0
W19TMB	Sonoita Creek (Red Rock Canyon)	15	0.189	0.189	-0.015	1.98	0	0	0
sC18TNS	Yaqui Topminnow	2	0.421	0.411	-0.035	3.65	13	1	1
sW20SBNWR	Yaqui Topminnow	7	0.381	0.392	-0.051	2.93	2	0	0

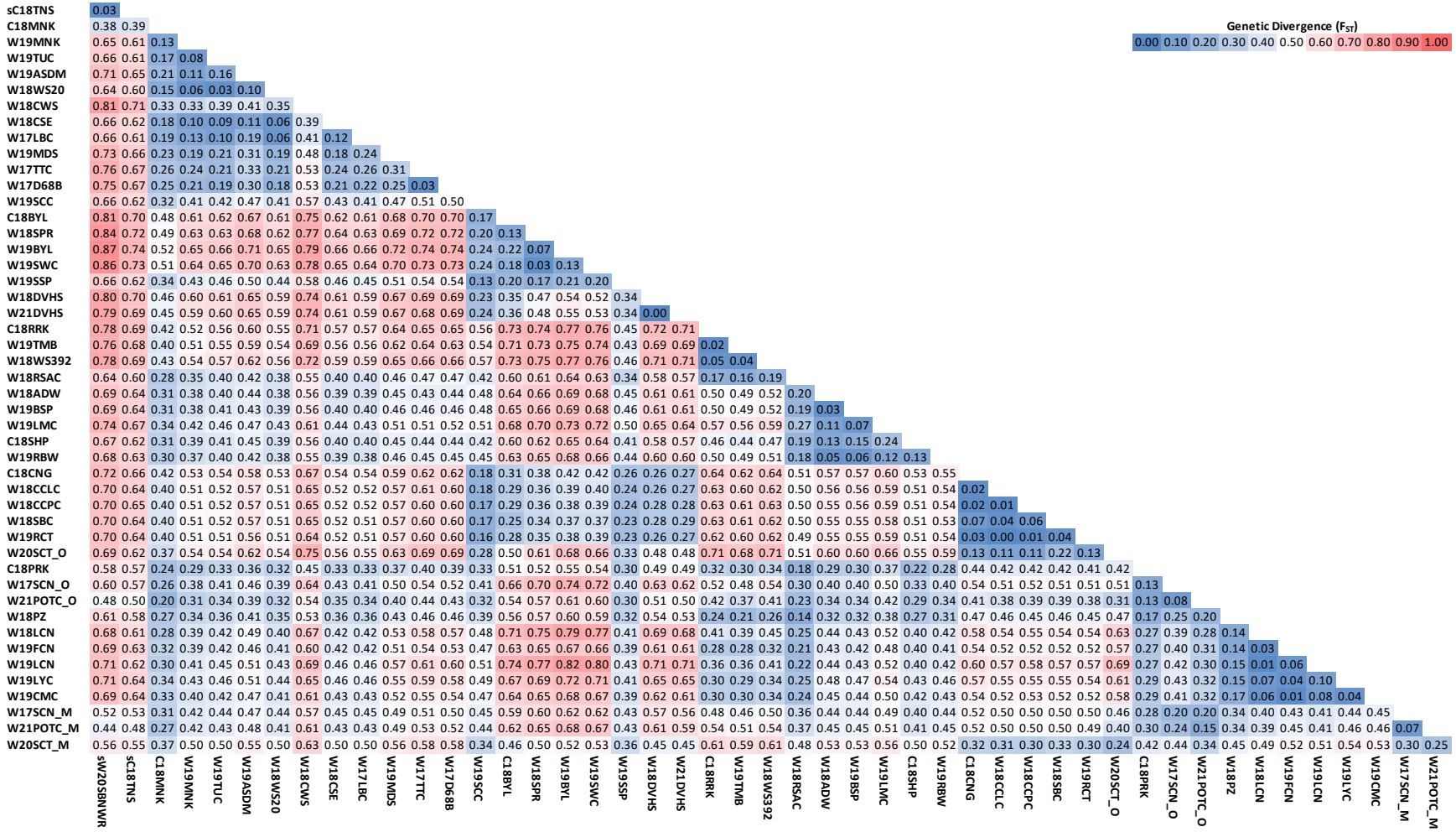
Table 4. Microsatellite loci that are fixed for specific alleles in each of the seven Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) genetic lineages, Yaqui Topminnow (*P. o. sonoriensis*), and putative *P. monacha-occidentalis*. The number associated with each locus for each lineage represents the allele (measured in base pairs) that is fixed in that lineage. Fixed alleles in the Monkey and Cottonwood Springs lineage ignore the captive Arizona State University Monkey Spring population, which represents a hybrid population of Gila and Yaqui Topminnow. Two loci in Yaqui Topminnow (*) were questionable with regards to fixation: *Pocc07* was monomorphic except for one individual that may be a hybrid, and *Pooc-OO56* only amplified in three individuals.

Locus	Bylas Springs	Cienega Creek	Monkey and Cottonwood Springs	Santa Cruz River	Sharp Spring	Sonoita Creek (Red Rock Canyon)	Sonoita Creek (Fresno / Coal Mine Canyon)	*P. monacha-occidentalis*	Yaqui Topminnow
Pocc02	203	-	-	-	-	-	199	-	-
Pocc03	181	-	-	-	-	-	-	-	-
Pocc07	169	169	169	169	169	169	169	-	177*
Pocc09	-	-	-	-	-	177	177	-	-
Pocc15	-	-	-	-	-	-	-	-	-
Pocc16	-	-	-	-	-	-	174	-	-
Pocc18	-	-	-	-	-	-	-	-	-
Pocc21	127	127	127	-	127	127	127	-	-
Pocc25	-	-	-	-	-	-	-	-	-
Pocc26	-	-	-	-	-	-	-	-	-
Pocc27	-	-	-	-	-	-	-	-	-
Pocc28	207	207	207	-	-	-	246	-	-
Pocc29	-	-	-	-	-	-	-	-	-
Pocc31	157	-	-	-	-	-	157	-	-
Pocc34	-	-	-	-	-	-	-	-	-
Pocc38	-	-	-	-	-	-	-	-	-
Pocc40	192	192	-	-	-	-	192	-	-
Pocc43	189	189	189	189	189	189	189	189	-
Pocc44	-	-	-	-	-	-	-	-	-
Pocc45	162	-	162	-	162	162	162	-	-
Pocc47	-	-	-	-	-	-	-	-	-
Pooc-4-44	107	107	107	-	107	107	107	-	-
Pooc-C15	-	-	-	-	-	-	-	-	-
Pooc-G10	192	192	192	-	192	192	192	-	180
Pooc-G49	-	159	159	-	159	159	159	-	-
Pooc-G53	101	101	101	-	101	101	101	-	-
Pooc-LL53	-	-	-	-	-	-	-	-	-
Pooc-OO56	-	-	-	-	149	-	149	-	149*
Total Monomorphic	**12**	**9**	**9**	**2**	**9**	**9**	**15**	**1**	**1**

Table 5. The number of private alleles detected in each of seven Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) lineages, Yaqui Topminnow (*P. o. sonoriensis*), and putative *P. monacha-occidentalis*. The number of alleles unique to a lineage is presented for each microsatellite locus. Collections of known mixed ancestries (e.g., captive Monkey Spring lineage at Arizona State University) were not included in calculations.

Locus	Bylas Springs	Cienega Creek	Monkey and Cottonwood Springs	Santa Cruz River	Sharp Spring	Sonoita Creek (Red Rock Canyon)	Sonoita Creek (Fresno / Coal Mine Canyon)	*P. monacha-occidentalis*	Yaqui Topminnow
Pocc02	-	-	-	-	-	-	-	1	1
Pocc03	-	-	-	-	-	-	-	1	1
Pocc07	-	-	-	-	-	-	-	1	1
Pocc09	-	-	-	-	-	-	-	1	3
Pocc15	-	-	-	-	-	-	-	1	1
Pocc16	-	-	1	-	-	-	-	1	1
Pocc18	-	-	-	-	-	-	-	-	2
Pocc21	-	-	-	1	-	-	-	-	1
Pocc25	-	-	-	1	-	-	-	1	4
Pocc26	-	1	1	1	-	-	-	1	4
Pocc27	-	-	1	1	1	-	-	-	-
Pocc28	-	-	-	1	-	-	2	1	2
Pocc29	-	-	-	1	-	-	-	-	1
Pocc31	-	-	1	1	-	-	-	1	-
Pocc34	-	-	1	1	-	-	-	2	5
Pocc38	1	-	-	-	-	-	-	-	5
Pocc40	-	-	1	1	-	-	-	-	1
Pocc43	-	-	-	-	-	-	-	-	3
Pocc44	1	-	2	1	1	2	-	1	2
Pocc45	-	1	-	-	-	-	-	3	5
Pocc47	-	-	1	1	-	-	-	-	3
Pooc-4-44	-	-	-	1	-	-	-	-	2
Pooc-C15	1	-	2	1	1	-	1	2	6
Pooc-G10	-	-	-	-	-	-	-	1	1
Pooc-G49	-	-	-	-	-	-	-	-	-
Pooc-G53	-	-	-	-	-	-	-	1	2
Pooc-LL53	-	1	2	2	-	-	-	1	4
Pooc-OO56	-	1	-	1	-	-	1	-	-
Total Private	**3**	**4**	**13**	**16**	**3**	**2**	**4**	**21**	**61**

Table 6. Pairwise estimates of genetic divergence (F_{ST}) calculated from 28 microsatellite loci. Colors range from low divergence (Blue; F_{ST} = 0.0) to high (Red; F_{ST} = 1.0). Locality names are constructed from a combination of status, two-digit collection year, and locality code from Table 1. Locality names beginning with a lower case 's' represent Yaqui Topminnow.



	sW20SRNWR	sC18TNS	C18MNK	W19MNK	W19TUC	W19ASDM	W18WS20	W18CWS	W18CSE	W17LBC	W19MDS	W17TTC	W17D68B	W19SCC	C18BYL	W18SPR	W19BYL	W19SWC	W19SSP	W18DVHS	W21DVHS	C18RRK	W19TMB	W18WS392	W18RSAC	W18ADW	W19BSP	W19LMC	C18SHP	W19RBW	C18CNG	W18CCLC	W18CCPC	W18SBC	W19RCT	W20SCT_O	C18PRK	W17SCN_O	W21POTC_O	W18PZ	W18LCN	W19FCN	W19LCN	W19LYC	W19CMC	W17SCN_M	W21POTC_M
sC18TNS	0.03																																															
C18MNK	0.38	0.39																																														
W19MNK	0.65	0.61	0.13																																													
W19TUC	0.66	0.61	0.17	0.08																																											
W19ASDM	0.71	0.65	0.21	0.11	0.16																																										
W18WS20	0.64	0.60	0.15	0.06	0.03	0.10																																										
W18CWS	0.81	0.71	0.33	0.33	0.39	0.41	0.35																																									
W18CSE	0.66	0.62	0.18	0.10	0.09	0.11	0.06	0.39																																								
W17LBC	0.66	0.61	0.19	0.13	0.10	0.19	0.06	0.41	0.12																																							
W19MDS	0.73	0.66	0.23	0.19	0.21	0.31	0.19	0.48	0.18	0.24																																						
W17TTC	0.76	0.67	0.26	0.24	0.21	0.33	0.21	0.53	0.24	0.26	0.31																																					
W17D68B	0.75	0.67	0.25	0.21	0.19	0.30	0.18	0.53	0.21	0.22	0.25	0.03																																				
W19SCC	0.66	0.62	0.32	0.41	0.42	0.47	0.41	0.57	0.43	0.41	0.47	0.51	0.50																																			
C18BYL	0.81	0.70	0.48	0.61	0.62	0.67	0.61	0.75	0.62	0.61	0.68	0.70	0.70	0.17																																		
W18SPR	0.84	0.72	0.49	0.63	0.63	0.68	0.62	0.77	0.64	0.63	0.69	0.72	0.72	0.20	0.13																																	
W19BYL	0.87	0.74	0.52	0.65	0.66	0.71	0.65	0.79	0.66	0.66	0.72	0.74	0.74	0.24	0.22	0.07																																
W19SWC	0.86	0.73	0.51	0.64	0.65	0.70	0.63	0.78	0.65	0.64	0.70	0.73	0.73	0.24	0.18	0.03	0.13																															
W19SSP	0.66	0.62	0.34	0.43	0.46	0.50	0.44	0.58	0.46	0.45	0.51	0.54	0.54	0.13	0.20	0.17	0.21	0.20																															
W18DVHS	0.80	0.70	0.46	0.60	0.61	0.65	0.59	0.74	0.61	0.59	0.67	0.69	0.69	0.23	0.35	0.47	0.54	0.52	0.34																														
W21DVHS	0.79	0.69	0.45	0.59	0.60	0.65	0.59	0.74	0.61	0.59	0.67	0.68	0.69	0.24	0.36	0.48	0.55	0.53	0.34	0.00																													
C18RRK	0.78	0.69	0.42	0.52	0.56	0.60	0.55	0.71	0.57	0.57	0.64	0.65	0.65	0.56	0.73	0.74	0.77	0.76	0.45	0.72	0.71																												
W19TMB	0.76	0.68	0.40	0.51	0.55	0.59	0.54	0.69	0.56	0.56	0.62	0.64	0.63	0.54	0.71	0.73	0.75	0.74	0.43	0.69	0.69	0.02																										
W18WS392	0.78	0.69	0.43	0.54	0.57	0.62	0.56	0.72	0.59	0.59	0.65	0.66	0.66	0.57	0.73	0.75	0.77	0.76	0.46	0.71	0.71	0.05	0.04																									
W18RSAC	0.64	0.60	0.28	0.35	0.40	0.42	0.38	0.55	0.40	0.40	0.46	0.47	0.47	0.42	0.60	0.61	0.64	0.63	0.34	0.58	0.57	0.17	0.16	0.19																								
W18ADW	0.69	0.64	0.31	0.38	0.40	0.44	0.38	0.56	0.39	0.39	0.45	0.43	0.44	0.48	0.64	0.66	0.69	0.68	0.45	0.61	0.61	0.50	0.49	0.52	0.20																						
W19BSP	0.69	0.64	0.31	0.38	0.41	0.43	0.39	0.56	0.40	0.40	0.46	0.46	0.46	0.48	0.65	0.66	0.69	0.68	0.46	0.61	0.61	0.50	0.49	0.52	0.19	0.03																					
W19LMC	0.74	0.67	0.34	0.42	0.46	0.47	0.43	0.61	0.44	0.43	0.51	0.51	0.52	0.51	0.68	0.70	0.73	0.72	0.50	0.65	0.64	0.57	0.56	0.59	0.27	0.11	0.07																			
C18SHP	0.67	0.62	0.31	0.39	0.41	0.45	0.39	0.56	0.40	0.40	0.45	0.44	0.44	0.42	0.60	0.62	0.65	0.64	0.41	0.58	0.57	0.46	0.44	0.47	0.19	0.13	0.15	0.24																		
W19RBW	0.68	0.63	0.30	0.37	0.40	0.42	0.38	0.55	0.39	0.38	0.46	0.45	0.45	0.45	0.63	0.65	0.68	0.66	0.44	0.60	0.60	0.50	0.49	0.51	0.18	0.05	0.06	0.12	0.13																	
C18CNG	0.72	0.66	0.42	0.53	0.54	0.58	0.53	0.67	0.54	0.54	0.59	0.62	0.62	0.18	0.31	0.38	0.42	0.42	0.26	0.26	0.27	0.64	0.62	0.64	0.51	0.57	0.57	0.60	0.53	0.55																
W18CCLC	0.70	0.64	0.40	0.51	0.52	0.57	0.51	0.65	0.52	0.52	0.57	0.61	0.60	0.18	0.29	0.36	0.39	0.40	0.24	0.26	0.27	0.63	0.60	0.62	0.50	0.56	0.56	0.59	0.51	0.54	0.02														
W18CCPC	0.70	0.65	0.40	0.51	0.52	0.57	0.51	0.65	0.52	0.52	0.57	0.60	0.60	0.17	0.29	0.36	0.38	0.39	0.24	0.28	0.28	0.63	0.61	0.63	0.50	0.55	0.56	0.59	0.51	0.54	0.02	0.01													
W18SBC	0.70	0.64	0.40	0.51	0.52	0.57	0.51	0.65	0.52	0.51	0.57	0.60	0.60	0.17	0.25	0.34	0.37	0.37	0.23	0.28	0.29	0.63	0.61	0.62	0.50	0.55	0.55	0.58	0.51	0.53	0.07	0.04	0.06												
W19RCT	0.70	0.64	0.40	0.51	0.51	0.56	0.51	0.64	0.52	0.51	0.57	0.60	0.60	0.16	0.28	0.35	0.38	0.39	0.23	0.26	0.27	0.62	0.60	0.62	0.49	0.55	0.55	0.59	0.51	0.54	0.03	0.00	0.01	0.04											
W20SCT_O	0.69	0.62	0.37	0.54	0.54	0.62	0.54	0.75	0.56	0.55	0.63	0.69	0.69	0.28	0.50	0.61	0.68	0.66	0.33	0.48	0.48	0.71	0.68	0.71	0.51	0.60	0.60	0.66	0.55	0.59	0.13	0.11	0.11	0.22	0.13										
C18PRK	0.58	0.57	0.24	0.29	0.33	0.36	0.32	0.45	0.33	0.33	0.37	0.40	0.39	0.33	0.51	0.52	0.55	0.54	0.30	0.49	0.49	0.32	0.30	0.34	0.18	0.29	0.30	0.37	0.22	0.28	0.44	0.42	0.42	0.42	0.41	0.42									
W17SCN_O	0.60	0.57	0.26	0.38	0.41	0.46	0.39	0.64	0.43	0.41	0.50	0.54	0.52	0.41	0.66	0.70	0.74	0.72	0.40	0.63	0.62	0.52	0.48	0.54	0.30	0.40	0.40	0.50	0.33	0.40	0.54	0.51	0.52	0.51	0.51	0.51	0.13								
W21POTC_O	0.48	0.50	0.20	0.31	0.34	0.39	0.32	0.54	0.35	0.34	0.40	0.44	0.43	0.32	0.54	0.57	0.61	0.60	0.30	0.51	0.50	0.42	0.37	0.41	0.23	0.34	0.34	0.42	0.29	0.34	0.41	0.38	0.39	0.39	0.38	0.31	0.13	0.08							
W18PZ	0.61	0.58	0.27	0.34	0.36	0.41	0.35	0.53	0.36	0.36	0.43	0.46	0.46	0.39	0.56	0.57	0.60	0.59	0.32	0.54	0.53	0.24	0.21	0.26	0.14	0.32	0.32	0.38	0.27	0.31	0.47	0.46	0.45	0.46	0.45	0.47	0.17	0.25	0.20						
W18LCN	0.68	0.61	0.28	0.39	0.42	0.49	0.40	0.67	0.42	0.42	0.53	0.58	0.57	0.48	0.71	0.75	0.79	0.77	0.41	0.69	0.68	0.41	0.39	0.45	0.25	0.44	0.43	0.52	0.40	0.42	0.58	0.54	0.55	0.54	0.54	0.63	0.27	0.39	0.28	0.14				
W19FCN	0.69	0.63	0.32	0.39	0.42	0.46	0.41	0.60	0.42	0.42	0.51	0.54	0.53	0.47	0.63	0.65	0.67	0.66	0.39	0.61	0.61	0.28	0.28	0.32	0.21	0.43	0.42	0.48	0.40	0.41	0.54	0.52	0.52	0.52	0.52	0.57	0.27	0.40	0.31	0.14	0.03			
W19LCN	0.71	0.62	0.30	0.41	0.45	0.51	0.43	0.69	0.46	0.46	0.57	0.61	0.60	0.51	0.74	0.77	0.82	0.80	0.43	0.71	0.71	0.36	0.36	0.41	0.22	0.44	0.43	0.52	0.40	0.42	0.60	0.57	0.58	0.57	0.57	0.69	0.27	0.42	0.30	0.15	0.01	0.06		
W19LYC	0.71	0.64	0.34	0.43	0.46	0.51	0.44	0.65	0.46	0.46	0.55	0.59	0.58	0.49	0.67	0.69	0.72	0.71	0.41	0.65	0.65	0.30	0.29	0.34	0.25	0.48	0.47	0.54	0.43	0.46	0.57	0.55	0.55	0.55	0.54	0.61	0.29	0.43	0.32	0.15	0.07	0.04	0.10		
W19CMC	0.69	0.64	0.33	0.40	0.42	0.47	0.41	0.61	0.43	0.43	0.52	0.55	0.54	0.47	0.64	0.65	0.68	0.67	0.39	0.62	0.61	0.30	0.30	0.34	0.24	0.45	0.44	0.50	0.42	0.43	0.54	0.52	0.53	0.52	0.52	0.58	0.29	0.41	0.32	0.17	0.06	0.01	0.08	0.04	
W17SCN_M	0.52	0.53	0.31	0.42	0.44	0.47	0.44	0.57	0.45	0.45	0.49	0.51	0.50	0.45	0.59	0.60	0.62	0.62	0.43	0.57	0.56	0.48	0.46	0.50	0.36	0.44	0.44	0.49	0.40	0.44	0.52	0.50	0.50	0.50	0.50	0.46	0.28	0.20	0.20	0.34	0.40	0.43	0.41	0.44	0.45
W21POTC_M	0.44	0.48	0.27	0.42	0.43	0.48	0.41	0.61	0.43	0.43	0.49	0.53	0.52	0.44	0.62	0.65	0.68	0.67	0.43	0.61	0.59	0.54	0.51	0.54	0.37	0.45	0.45	0.51	0.41	0.45	0.52	0.50	0.50	0.50	0.49	0.40	0.30	0.24	0.15	0.34	0.39	0.45	0.41	0.46	0.46	0.07
W20SCT_M	0.56	0.55	0.37	0.50	0.50	0.55	0.50	0.63	0.50	0.50	0.56	0.58	0.58	0.34	0.46	0.50	0.52	0.53	0.36	0.45	0.45	0.61	0.59	0.61	0.48	0.53	0.53	0.56	0.50	0.52	0.32	0.31	0.30	0.33	0.30	0.24	0.42	0.44	0.34	0.45	0.49	0.52	0.51	0.54	0.53	0.30	0.25

Table 7. Results of effective population size (N_E) and genetic bottleneck tests per population. The 95% jackknife confidence interval (95% CI) is provided for N_E estimates in addition to the P_{Crit} value used in calculations. Large or infinite NE values cannot be taken at face value, and likely reflect statistical anomalies that resulted from the large number of monomorphic microsatellite loci observed in most populations. P-values are displayed for genetic bottleneck tests, which were conducted under three models (IAM: infinite alleles model; SMM: stepwise mutation model; TPM: two-phase model). Significance was assessed using a Bonferroni-adjusted two-tailed P-value ($\alpha = 0.001$). Assessment of a shift in the mode of allele frequencies from a typical L-shaped distribution was conducted (Shifted = indication of a genetic bottleneck). The Garza-Williamson *M* Ratio (G-W) and its modified version (G-W Mod) were also calculated. $M < 0.68$ indicates significant evidence of a bottleneck. Population names are constructed from the status, two-digit collection year, and locality code from Table 1. Locality names beginning with a lower case 's' represent Yaqui Topminnow. Significant values are bolded.

Population	P_{Crit}	N_E	95% CI	IAM	TPM	SMM	Mode-Shift	G-W	G-W Mod
C18BYL	0.020	446	22.1 - ∞	0.733	1.000	0.791		0.706	**0.178**
C18CNG	0.020	83	41.9 - 360.8	0.978	0.107	0.003		0.785	**0.244**
C18MNK	0.020	53	32.2 - 108.8	**0.000**	**0.000**	0.070		0.694	**0.423**
C18PRK	0.020	34	21.3 - 62.8	**0.000**	0.003	0.898		0.723	**0.304**
C18RRK	0.020	38	16.3 - 177.4	0.016	0.240	0.898		0.696	**0.189**
C18SHP	0.020	144	55.7 - ∞	0.018	0.225	0.860		0.697	**0.245**
W17D68B	0.020	660	35.3 - ∞	0.005	0.057	0.542		0.816	**0.185**
W17LBC	0.020	157	54.1 - ∞	0.013	0.109	0.854		0.843	**0.242**
W17SCN_O	0.056	∞	20.7 - ∞	0.393	0.919	0.191		**0.628**	**0.269**
W17TTC	0.020	56	18.3 - ∞	0.001	0.003	0.013		0.789	**0.177**
W18ADW	0.020	41	17.7 - 198.3	0.011	0.190	0.644		0.704	**0.229**
W18CCLC	0.020	180	58.2 - ∞	0.391	0.670	0.058		0.813	**0.236**
W18CCPC	0.020	13	1.6 - 718.9	0.123	0.005	**0.000**		0.730	**0.285**
W18CSE	0.020	265	74.6 - ∞	0.026	0.252	0.252		0.822	**0.244**
W18CWS	0.020	183	17.8 - ∞	0.008	0.008	0.020		**0.617**	**0.160**
W18DVHS	0.020	∞	48.8 - ∞	0.893	0.305	0.057		0.755	**0.205**
W18LCN	0.050	3	0.9 - ∞	0.002	0.020	0.153	**Shifted**	0.723	**0.192**
W18PZ	0.020	279	93.4 - ∞	0.030	0.571	0.105		0.781	**0.327**
W18RSAC	0.020	∞	114.8 - ∞	0.008	0.055	0.984		**0.675**	**0.260**
W18SBC	0.020	∞	300 - ∞	0.095	0.978	0.252		0.792	**0.225**
W18SPR	0.020	147	22.7 - ∞	0.077	0.006	0.002		0.793	**0.182**
W18WS20	0.020	291	84.1 - ∞	0.080	0.712	0.027		0.835	**0.292**
W18WS392	0.020	57	21.4 - ∞	0.497	0.635	0.216		0.737	**0.197**
W19ASDM	0.020	17	3.1 - 145.1	0.890	0.132	0.005		**0.654**	**0.254**
W19BSP	0.020	49	25.4 - 146.7	0.005	0.174	0.818		**0.645**	**0.218**
W19BYL	0.020	∞	10.7 - ∞	0.102	0.012	0.009		0.768	**0.174**
W19CMC	0.020	202	44.2 - ∞	0.210	0.922	0.241		0.809	**0.245**
W19FCN	0.020	245	46.8 - ∞	0.089	0.963	0.207		0.769	**0.250**
W19LCN	0.056	3	1.3 - 20.1	0.542	0.952	0.542	**Shifted**	0.708	**0.185**
W19LMC	0.020	∞	124.4 - ∞	0.093	0.782	0.298		**0.667**	**0.216**
W19LYC	0.020	45	16.6 - ∞	0.021	0.144	0.323		0.732	**0.201**
W19MDS	0.020	110	38.7 - ∞	0.121	1.000	0.135		0.801	**0.218**
W19MNK	0.020	215	65.9 - ∞	0.057	0.404	0.378		0.851	**0.268**
W19RBW	0.020	903	69.1 - ∞	**0.000**	0.001	0.023	**Shifted**	**0.623**	**0.201**
W19RCT	0.020	∞	104.8 - ∞	0.464	0.632	0.074		0.822	**0.243**
W19SCC	0.020	159	61 - ∞	0.006	0.210	0.568		**0.629**	**0.244**
W19SSP	0.020	244	76.1 - ∞	0.142	0.734	0.119		**0.668**	**0.258**
W19SWC	0.020	∞	33.4 - ∞	0.492	0.160	0.105		0.788	**0.165**
W19TMB	0.020	∞	205.2 - ∞	0.168	0.893	0.244		0.736	**0.201**
W19TUC	0.020	46	26.7 - 100.8	0.011	0.284	0.190		0.847	**0.256**
W20SCT_O	0.125	∞	0.7 - ∞	0.652	0.734	0.734	**Shifted**	0.807	**0.161**
W21DVHS	0.020	141	24.5 - ∞	0.277	0.048	0.010		0.743	**0.205**
W21POTC_O	0.036	∞	76.5 - ∞	0.157	0.940	0.111	**Shifted**	0.714	**0.448**
W17SCN_M	0.020	∞	∞ - ∞	0.012	0.336	0.455		**0.648**	**0.396**
W20SCT_M	0.020	∞	4.9 - ∞	**0.000**	0.003	0.114	**Shifted**	**0.644**	**0.296**
W21POTC_M	0.036	∞	∞ - ∞	**0.000**	0.036	0.940		**0.667**	**0.453**
sC18TNS	0.020	56	35 - 106.2	0.269	0.452	0.004		0.752	**0.413**
sW20SBNWR	0.083	17	2.8 - ∞	0.048	0.277	0.927		0.730	**0.261**

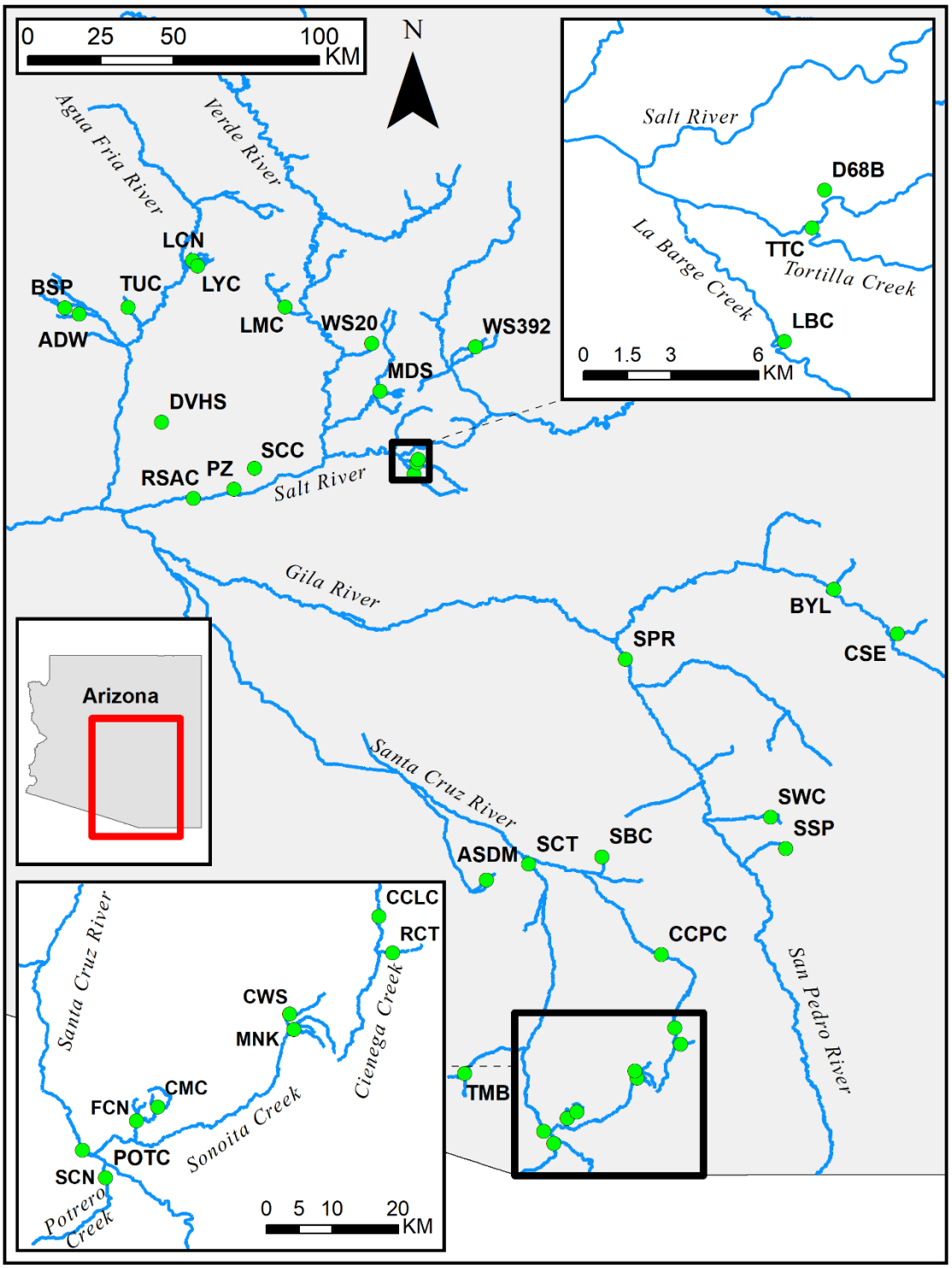


Figure 1. Gila Topminnow (*Poeciliopsis occidentalis*) sampling localities located in Arizona (center left inset map). The Upper right inset shows detail within the La Barge and Tortilla Creek drainages. The lower left inset shows detail within the Santa Cruz River, Sonoita Creek, and Cienega Creek drainages. Major rivers are labeled in all maps. Locality codes correspond to those in Table 1.

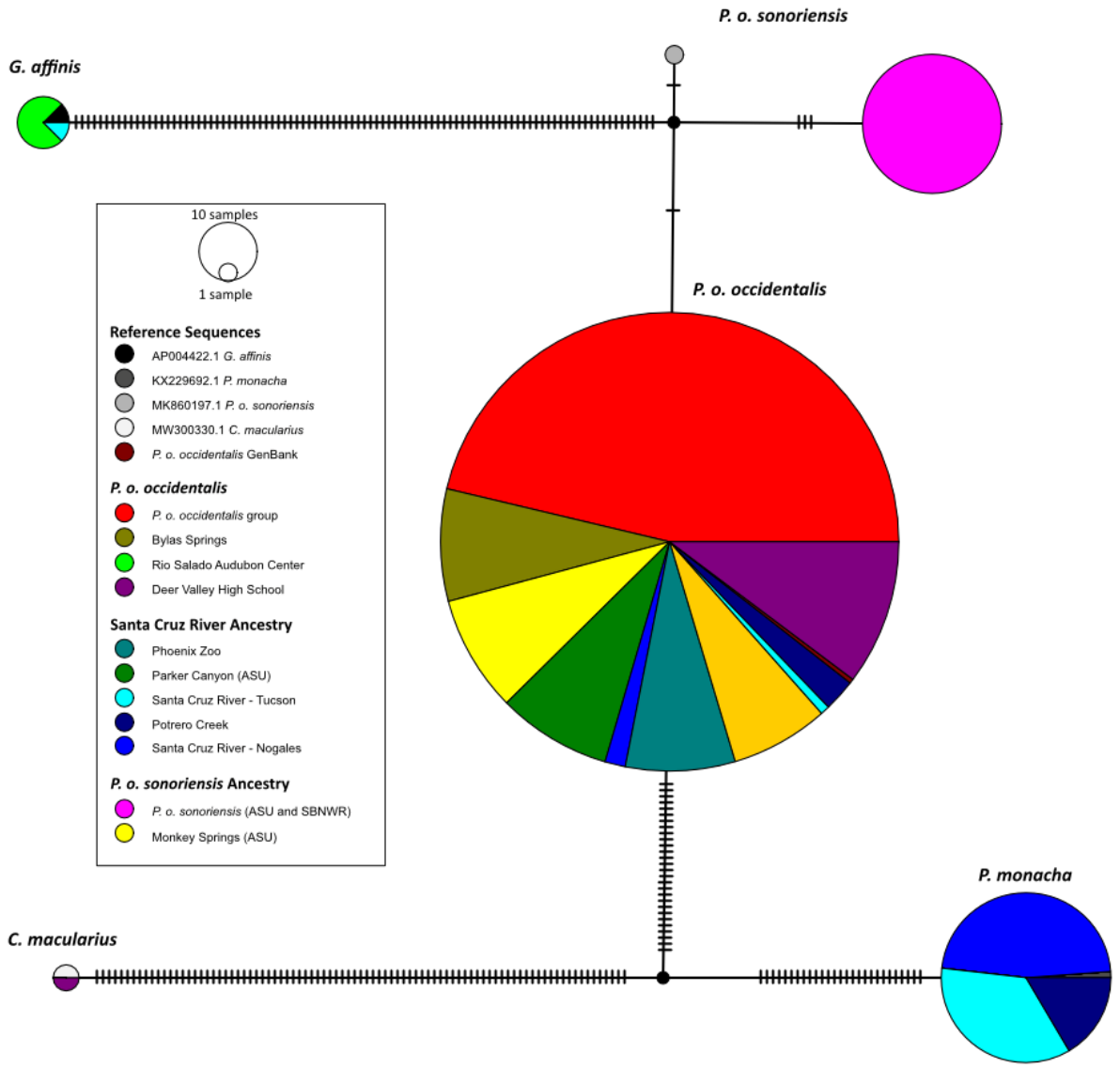


Figure 2. Haplotype network for 767 samples selected for sequencing of 581 base pairs (bp) of the mitochondrial gene cytochrome c oxidase subunit I (COI). Each circle represents a haplotype with its size being proportional to the number of times a haplotype was observed. Colored sections of each circle are proportional to the number of times each haplotype was observed in a population. Six reference sequences downloaded from NCBI GenBank representing *Cyprinodon macularius* (N=1), *Gambusia affinis* (N=1), *Poeciliopsis monacha* (N=1), *P. occidentalis occidentalis* (N=2; GenBank accessions HQ556953.1 and HQ556954.1), and *P. o. sonoriensis* (N=1) are included. Sample locality names are provided in the legend, with the exception of "*P. o. occidentalis* group" which is comprised of individuals from Cienega Creek (Arizona State University), Sharp Spring (Arizona State University), Unnamed Drainage #68b, Tortilla Creek, AD Wash #242, Cienega Creek – Las Cienegas, Cienega Creek – Pima County Preserve, Cold Spring #85, and Sabino Canyon.

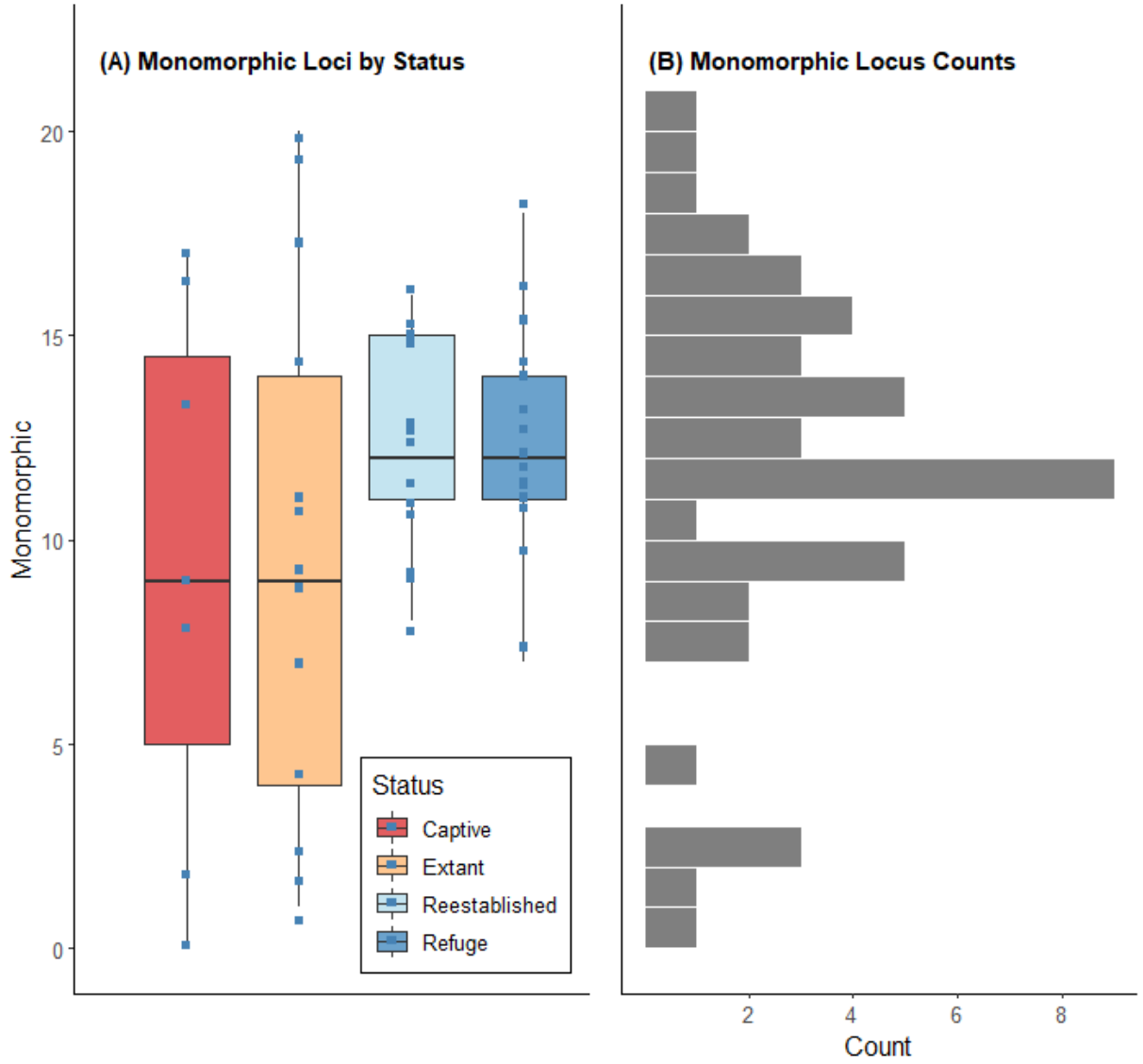


Figure 3. Monomorphic microsatellite loci were observed in nearly every population evaluated. Plot (A) shows the distribution of the number of monomorphic loci per population, with localities grouped according to their status (i.e., whether they represent captive populations, extant populations that have never been augmented, populations that have been reestablished, and refuge populations). Plot (B) shows a histogram of the global distribution of monomorphic loci. Plots (A) and (B) share a common Y-axis. Refuge and Reestablished sites appear to have a greater median number of monomorphic loci; however, captive and extant groupings are biased by containing hybrid populations, putative *Poeciliopsis monacha-occidentalis*, and *P. occidentalis sonoriensis* populations which tended to have fewer monomorphic loci per population. Furthermore, any difference among these groupings was not statistically significant (one-way ANOVA; degrees of freedom = 3, F = 1.456; P = 0.24).

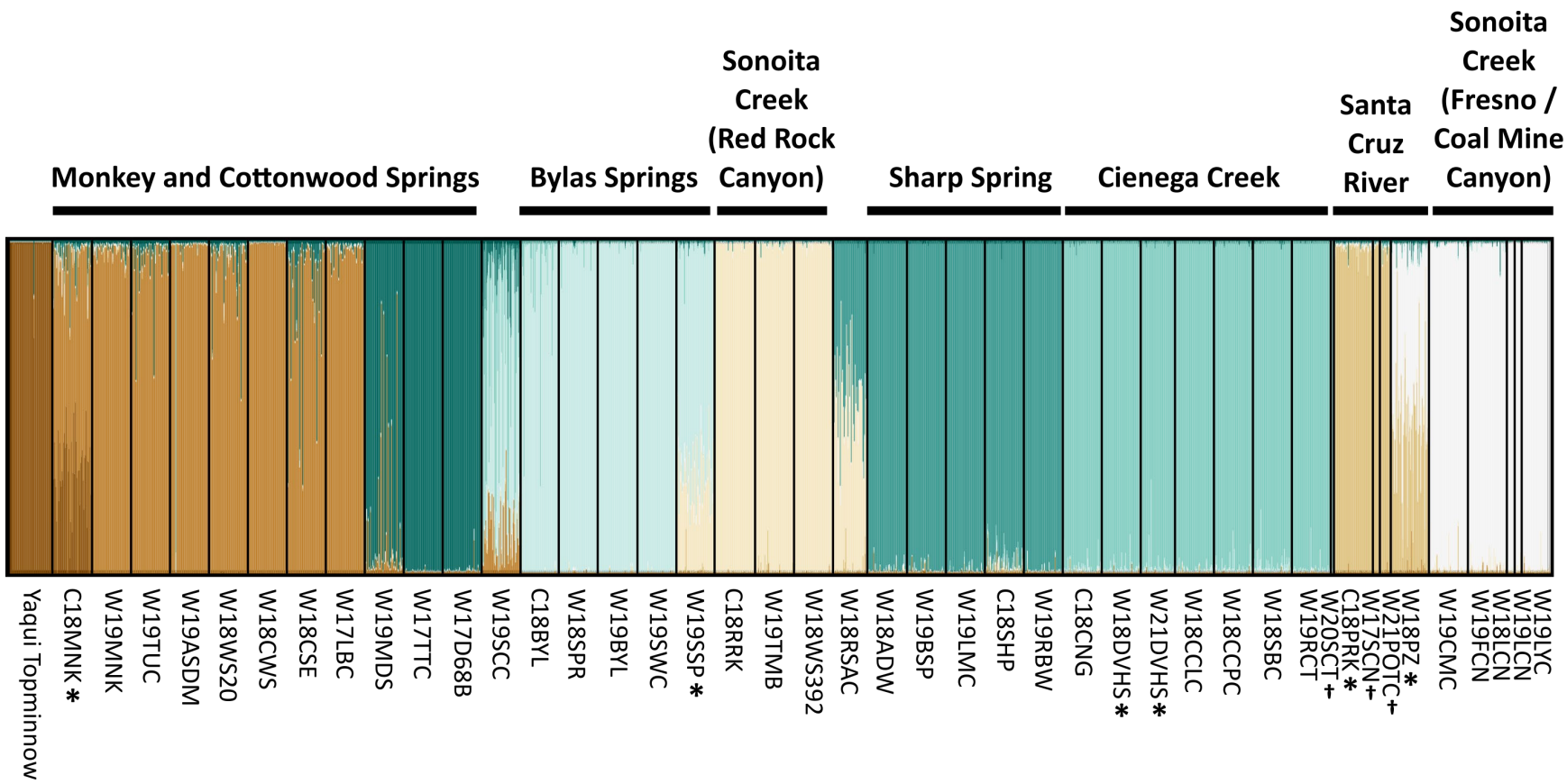


Figure 4. Population structure results showing the assignment of each sampling locality to each of nine genetic clusters representing Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) and Yaqui Topminnow (*P. o. sonoriensis*) lineages defined in earlier works. Monkey and Cottonwood Springs lineage is divided among two clusters. Each genetic cluster is represented by a color, and each individual sample is represented by a vertical bar. The proportions of color within each vertical bar represent the proportion of ancestry for that individual as calculated from population allele frequencies. Horizontal black bars at the top of the figure show the sample localities assigned to each genetic lineage. Sample locality names are represented by a combination of status, two-digit collection year, and locality code from Table 1. Those denoted with an asterisk (*) were either assigned to a cluster that conflicted with management records, or showed evidence of admixture with another lineage. Localities labeled with a dagger (†) are Santa Cruz River sites that harbor individuals with a Headwater Livebearer (*P. monacha*) mtDNA haplotype. Individuals with a Headwater Livebearer haplotype were suspected to be *P. monacha-occidentalis*, and were excluded from analyses conducted to produce this figure.

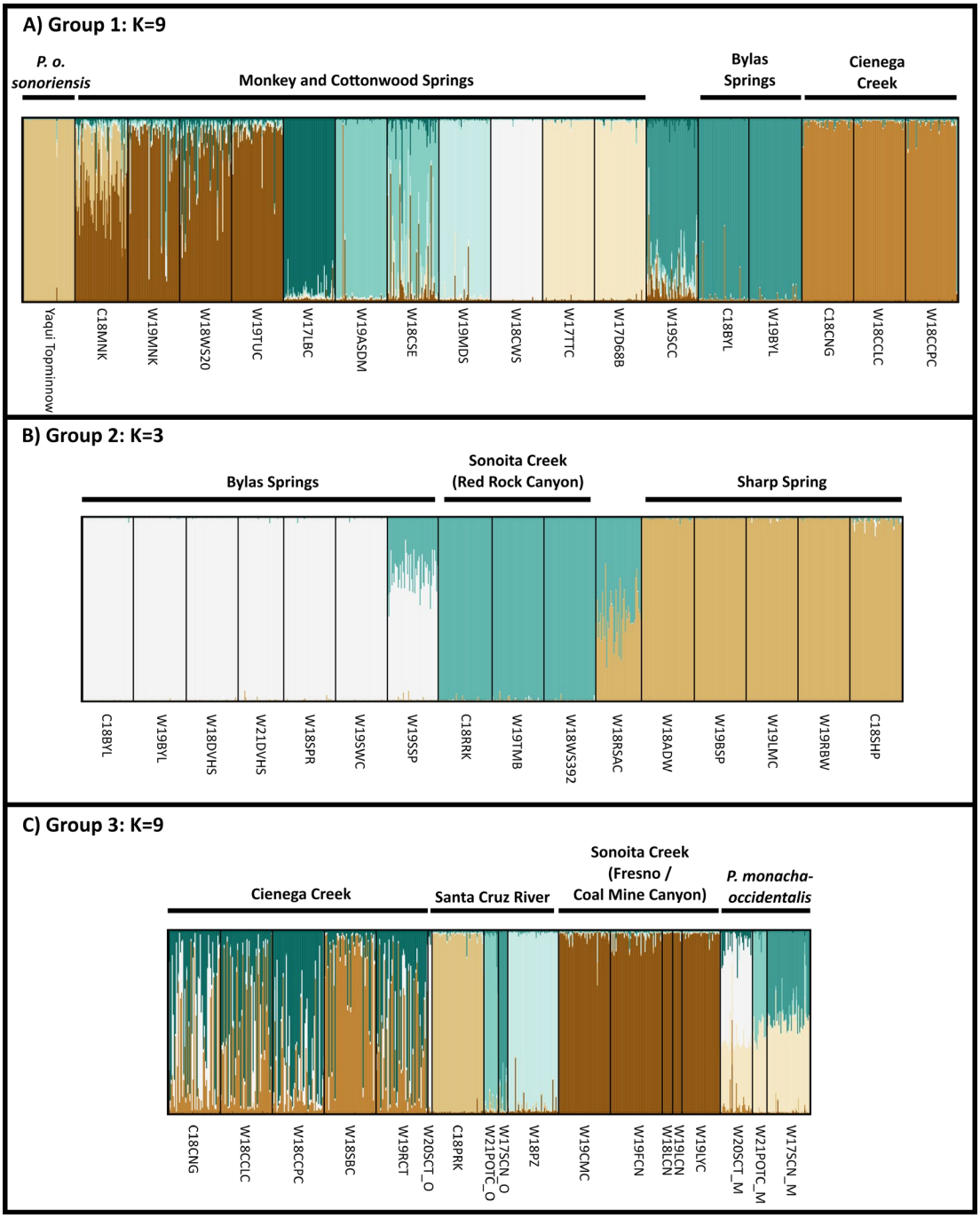


Figure 5. Subpopulation structure for the genetic lineages shown in Figure 4. Each Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) lineage was evaluated for within-lineage population structuring. Each genetic cluster is represented by a color, and each individual sample is represented by a vertical bar. The proportions of color within each vertical bar represent the proportion of ancestry for that individual. Colors representing populations are exclusive to each group, and are not shared across the three panels in the figure. Horizontal black bars at the top of the figure show the name for each genetic lineage. Sample locality names are represented by a combination of status, two-digit collection year, and locality code from Table 1. Some sample localities appear in multiple groups to account for mixing among lineages in captive sites and refuges, meaning a total of 19 genetic clusters were detected across all lineages previously outlined in Figure 4.

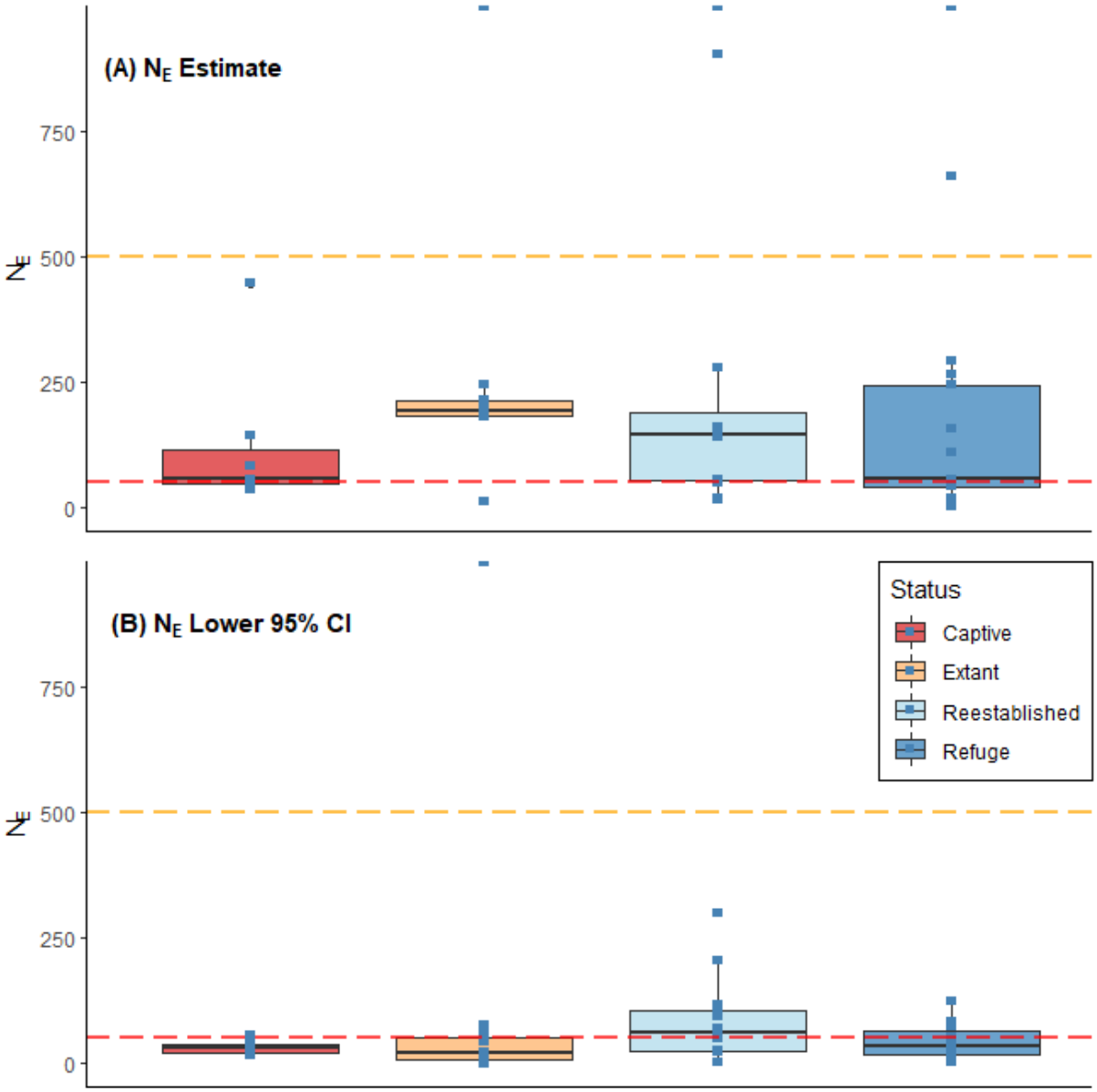


Figure 6. Effective population size (N_E) was calculated for each sampling locality. Localities were classified as captive populations, extant populations that have never been augmented, populations that have been reestablished, and refuge populations. Box plots show the distribution of N_E for each group of populations, with boxplots in panel (A) showing the point estimate for N_E, and panel (B) showing the lower bound of the 95% confidence interval (CI) for each estimate. The upper bound of the CI was not plotted because it was infinite for 37 of 48 sampling events (77%). The dashed red line on each plot is placed at N_E=50, which is commonly considered the minimum population size necessary to minimize the negative effects of inbreeding. The dashed orange line is placed at N_E=500, which is considered the minimum population size necessary to mitigate long-term effects of genetic drift. Comparisons of N_E point estimates among the four groupings were not statistically significant (one-way ANOVA; degrees of freedom = 3, F = 0.346, P = 0.792).

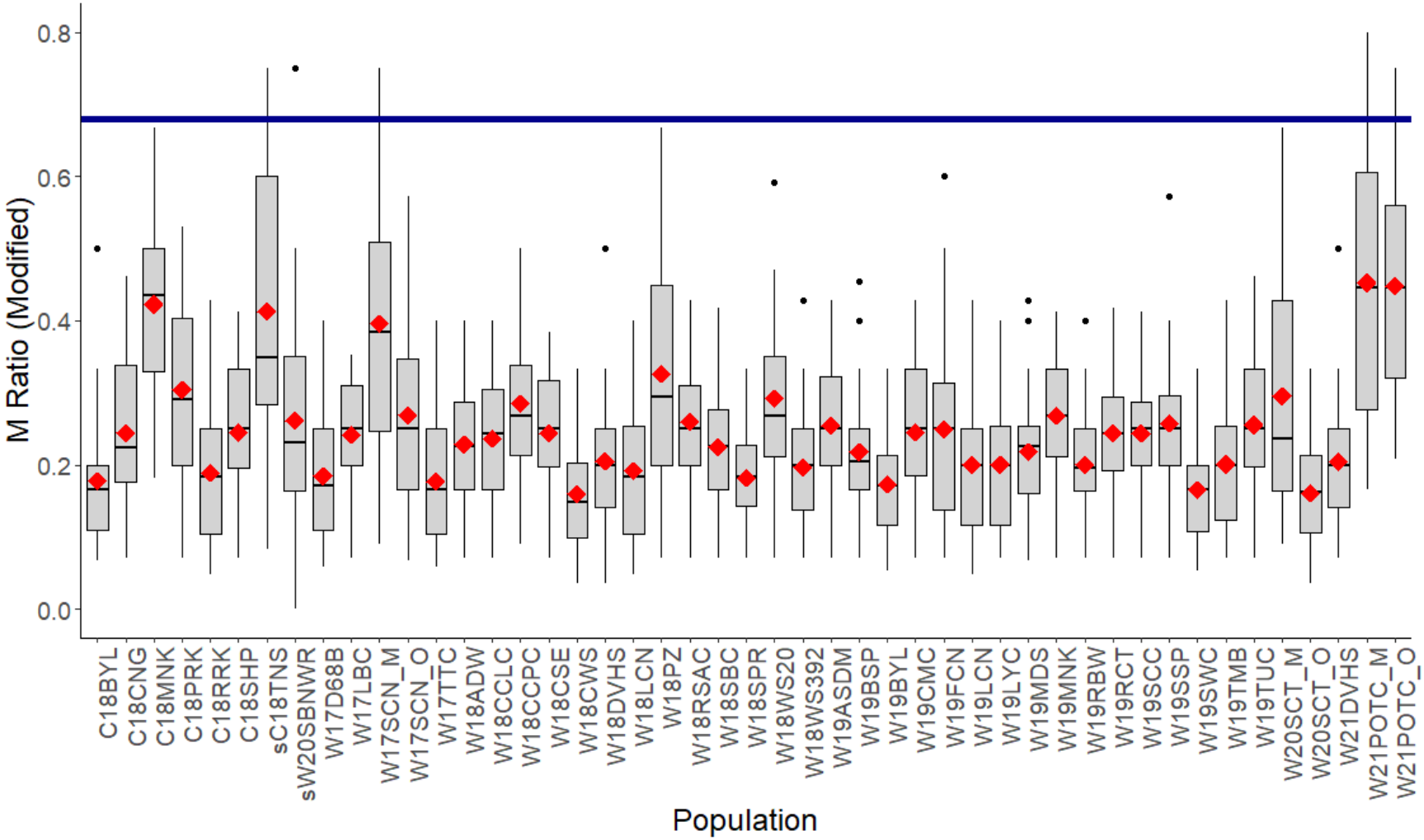


Figure 7. The modified Garza-Williamson *M* Ratio was calculated per locus per population to assess each population for evidence of a genetic bottleneck. The distribution of the *M* Ratio for each population is plotted here, with red diamonds representing the mean value of all loci each population. All mean values fall below the critical value (0.68) which is plotted as a solid horizontal blue line. Population codes on the X axis correspond to those in Table 1. Locality codes that start with a lower case 's' indicate Yaqui Topminnow populations.

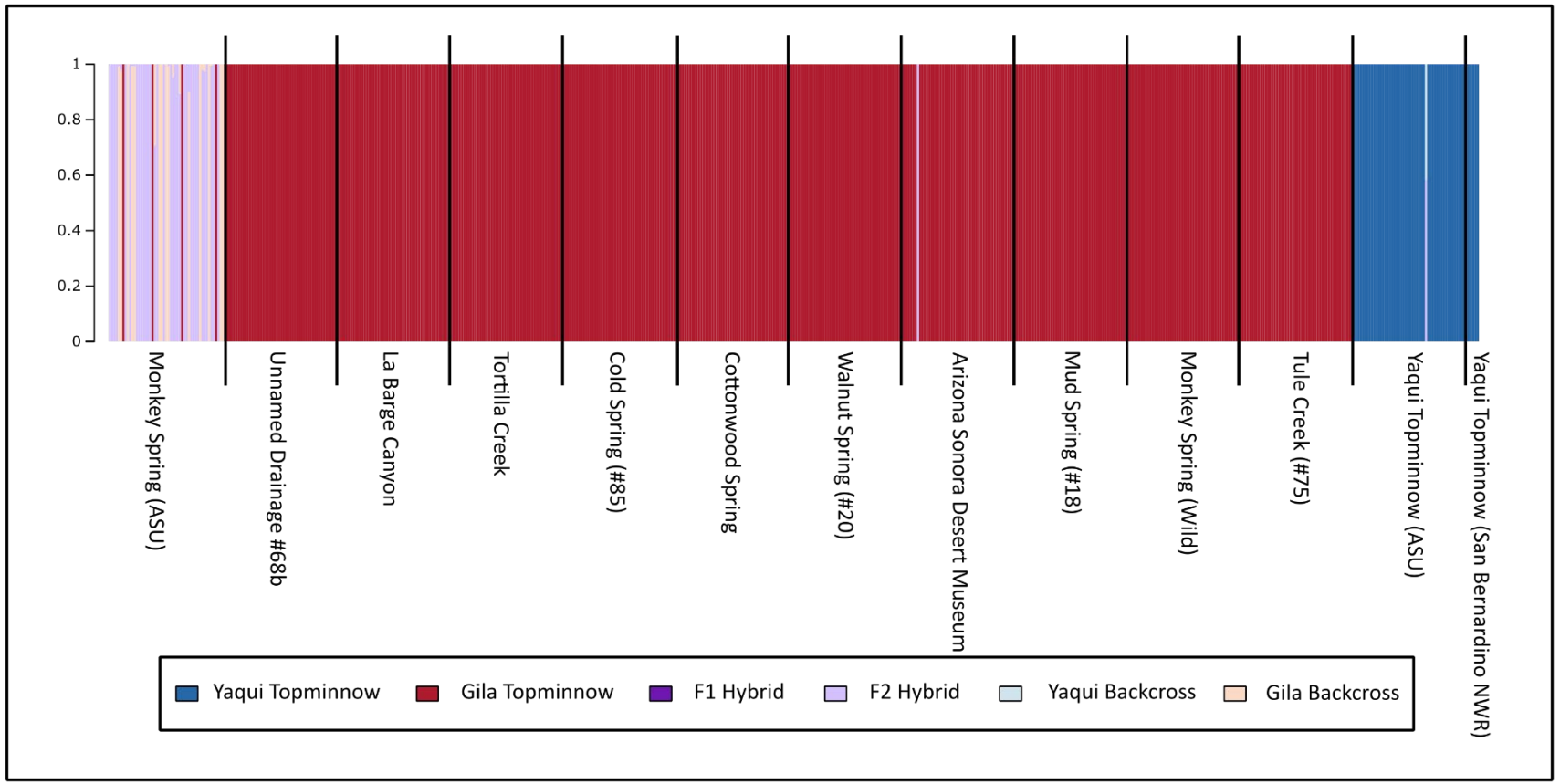


Figure 8. Hybrid analysis of the Monkey and Cottonwood Springs lineage verified that the captive Monkey Spring population previously maintained at Arizona State University (ASU) represented a hybrid group of Gila Topminnow (*Poeciliopsis occidentalis occidentalis*) and Yaqui Topminnow (*P. o. sonoriensis*). Each vertical bar in the figure represents an individual sample (N=607). The proportions of colors in each bar represent the probability of assignment to each of six hybrid categories as shown in the figure legend. The ASU Monkey Spring fish were comprised of second-generation hybrids (F2 hybrid; N = 33), Gila Topminnow backcrosses [i.e., offspring of a first-generation (F1) hybrid and a Gila Topminnow; N = 14], and Gila Topminnow (N = 4). Assignments were made with high confidence [Posterior probability (Pr) > 0.90] with two exceptions, both of which were F2 hybrids (Pr = 0.896 and Pr = 0.704).