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

Preliminary Draft Report For: United States Bureau of Reclamation

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

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

Co-Invest./Cooperators: Ross Timmons[1], Doug Duncan[2], Tony Robinson[1], William Stewart[3], Kent Mosher[3]

1 Arizona Game and Fish Department
2 U.S. Fish and Wildlife Service
3 U.S. Bureau of Reclamation

Executive Summary:

Gila Topminnow (*Poeciliopsis occidentalis*), once considered one of the most abundant fishes of the lower Colorado River basin, has suffered severe declines over the past century. The genetic consequences of these declines and the imprint of historical biogeography on its genome were evaluated previously, yet these efforts did not address concerns about effective population size or the genetic status of the many re-established and refuge populations. This study expands upon previous efforts by adding 22 novel microsatellite loci to existing marker panels. Sampling efforts covered a greater number of localities, including captive and wild stocks (40 localities, 1,952 samples), to conduct a more complete genetic evaluation of U.S. populations. Combined, these actions to 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 deviation shows introgression of Yaqui Topminnow (*P. o. sonorensis*) 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.

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 eight loci were consistently monomorphic at all of its localities, with some localities having up to 20 (of 29) 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, however some were infinite and wide confidence intervals surrounded many 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 23 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. Finally, a population within the Santa Cruz River near the international border with Mexico contains alleles not found in any other population, and may be evidence of hybridization, or presence of another species within this drainage. These samples should be evaluated alongside a variety of samples from neighboring drainage basins in Mexico to test hypotheses concerning their lineage.

Background:

The Gila Topminnow (*Poeciliopsis 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, *P. occidentalis* prefers relatively warm, shallow waters, but is also able to tolerate 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 historic abundance, *P. occidentalis* has suffered severe declines over the past century, primarily as a result of habitat degradation and the introduction of nonnative species (USFWS 1998). 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 *P. occidentalis* was first listed under the Endangered Species Preservation Act (USDI 1967), only 10 extant natural populations were known to exist in the United States. With the discovery of four new populations and the probable extirpation of a population in the North Fork of Ash Creek, there are now 13 known extant wild populations of *P. occidentalis* in the United States, all occurring in Arizona (AZGFD, pers. comm.). In addition, six populations established as refugia are currently maintained at the Arizona State University (ASU) Animal Facility. An additional 12 populations have been established at refuge sites in the wild, and another 10 populations have been reestablished at captive and wild sites.

Several previous studies examining the population genetics of *P. occidentalis* 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 *P. occidentalis* diversity in two distinct ways. First, 22 novel microsatellite loci were developed and added to seven existing markers to create a new panel of 29 neutral loci for quantifying genetic diversity within and among populations at greater

resolution than previously attainable. Second, sampling efforts were expanded to include refugia, reestablished populations, and extant natural populations, thus providing a more complete analysis of *P. occidentalis* populations within the United States. Data were analyzed for this preliminary 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 *P. occidentalis* (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

Samples were collected to represent all captive populations of *P. occidentalis* and *P. o. sonorensis* (Yaqui Topminnow) at ASU. Extant wild, re-established, and refugium populations throughout its modern range in the United States were also collected (Figure 1). Fifty samples were targeted for collection during each of 41 sampling events at 40 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, resulting in localities with <50 samples (Table 1). This yielded 1,952 samples from 40 localities collected from 2017 to 2019 ($\overline{x}$ = 48.8 samples per locality). Whole fish were collected and stored in 95% ethanol. Genomic DNA was extracted from a tissue clip using the DNeasy® Blood and Tissue Kits (Qiagen, Valencia, CA, 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 bases. 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-two loci were ultimately selected for genotyping (Table 2).

Genotyping

The 22 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 29 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.

Genetic Equilibrium and Diversity

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 α =0.000123; HWE α = 0.0017) 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 6.5 (Peakall and Smouse 2012). Allelic richness (A_R) was calculated via rarefaction in HP-Rare (Kalinowski 2004, 2005) using the second lowest number of observed alleles at a site (Larry Creek; N=74). Private alleles were recorded at each sampling locality, and again for putatively 'pure' representatives of the eight lineages identified in population structure analyses (see below).

Population Structure

The program Structure 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 8 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). The eight clusters assumed to exist in the dataset were 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. Twelve replicates were performed, each consisting of 250,000 generations of burn-in followed by 1,000,000 generations of data collection. Clumpak (Kopelman et al. 2015) was used to identify multimodality among 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 (Excoffier and Lischer 2010) with groups clustered according to the K=8 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 calculated by the K=8 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 Bylas Springs). 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. For each of these three datasets, Structure was used to evaluate K=1 to 20 with twelve 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. Clumpak was again utilized to assess multimodality among replicate runs, as well as execute Clumpp and Distruct to summarize and visualize Structure output (Kopelman et al. 2015).

Effective Population Size and Genetic Bottlenecks

NeEstimator 2.1 (Do et al. 2014) was utilized to calculate effective population size (N_E) and 95% jackknife confidence intervals for each sample locality. Rare alleles ($P_{CRIT} < 0.02$) were excluded from analysis following recommendations of Waples and Do (2010).

Multiple methods were applied to test for the presence of a recent genetic bottleneck at each sampling locality. 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 ($\alpha = 0.0012$) 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 evaluate the presence of a genetic bottleneck 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 potential introgression of Yaqui Topminnow alleles into the ASU captive Monkey Spring population. To verify, the Bayesian clustering program NewHybrids v1.1 beta 3 (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 wild Monkey Spring (extant population) as 'pure' populations. All other Monkey and Cottonwood Springs lineage individuals were included in the analysis (N=601; Arizona-Sonora Desert Museum, Cold Spring, Cottonwood Spring, La Barge Canyon, Mud Spring, Tortilla Creek, Tule Creek, Unnamed Drainage #68b, and Walnut Spring #20). 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 generations of burn-in followed by 5,000,000 generations of data collection.

Results:

Genetic Equilibrium and Diversity

Hardy-Weinberg deviations were mostly evident in the Santa Cruz population located north of Nogales (SCN). Here, eight loci showed signs of heterozygote deficiency while 14 showed evidence of heterozygote excess. Only two other populations showed evidence of deficiency at any locus: Cienega Creek – Pima County Preserve (CCPC) at *Pocc36* and Arizona-Sonora Desert Museum (ASDM) at *Pocc34*. Similarly, just two other populations showed evidence of heterozygote excess: Cold Spring (CSE) at *Pocc27* and Yaqui Topminnow at *Pocc43*.

The Santa Cruz River-Nogales population also indicated significant linkage disequilibrium among 145 pairs of loci. However, these patterns were not repeated for other sampled localities. Just three other pairs of loci exhibited linkage disequilibrium: *Pocc09* by *Pooc-LL53* for Yaqui Topminnow, *Pocc16* by *Pocc36* for the ASU captive Parker Canyon (PRK) population, and *Pocc31* by *Pooc-C15* for both the ASU captive Monkey Spring (MNK) and Parker Canyon populations. Overall, patterns of linkage disequilibrium and deviation from Hardy Weinberg expectations rarely occurred for the same loci at multiple sampling localities, with *Pocc31* by *Pooc-C15* providing the sole exception to this trend. All 29 loci were therefore retained for downstream analysis, since consistent linkage disequilibrium problems appear localized to Santa Cruz River-Nogales indicating these patterns may be a symptom of its population composition and demographic history.

Measures of genetic diversity were generally low in all populations (Table 3). Expected heterozygosity ranged from a low of 0.070 in the wild Bylas Springs population to a high of 0.568 in the captive Monkey Spring population. Heterozygosity tended to be highest in populations that were suspected or demonstrated (see below) to be comprised of mixed lineages (e.g., captive Monkey Spring: 0.568; Santa Cruz River-Nogales: 0.488; Phoenix Zoo: 0.365; Rio Salado Audubon Center: 0.344). A similar trend was noted for allelic richness (A_R), which ranged from 1.5 to 4.37. These same four populations were again among the highest values for this metric (captive Monkey Spring = 4.37; Santa Cruz River-Nogales = 3.89; Phoenix Zoo = 3.5; Rio Salado Audubon Center = 2.9). The lowest value was observed in Cottonwood Spring (A_R = 1.5).

Several microsatellite loci were fixed for a single allele in many populations. The number of fixed loci per sampling locality is found in Table 3, while the fixed loci per lineage are reported in Table 4. On average, each locality had 11.87 fixed loci out of 29 (40.9%), and this value ranged from 0 fixed loci (captive Monkey Spring up to 20 (Cottonwood Springs). The captive Money Spring population was the only with no fixed loci, likely a consequence of being of hybrid ancestry involving Yaqui Topminnow (see below). Santa Cruz River-Nogales had just one fixed locus, while Yaqui Topminnow exhibited three (but see caveats in Table 4). All other localities had eight or more fixed loci. The number of fixed loci per locality (Table 3) tended to be greater than the number detected in each locality's respective lineage (Table 4), and these numbers sometimes varied substantially. The most extreme example is Cottonwood Springs, for which its 20 monomorphic loci were more than double the number of loci fixed across all Monkey and Cottonwood Springs sites (9).

Private alleles were detected in each lineage (Table 5), and several individual collection localities (Table 3). Yaqui Topminnow had the greatest overall number of private alleles (69) while the Santa Cruz River had 38. The next greatest was Monkey and Cottonwood Springs (16) while all other lineages had fewer than 8 private alleles each. Bylas Springs and Sonoita Creek (Red Rock Canyon) lineages had a single private allele each. When private alleles were evaluated according to sampling location, Yaqui Topminnow again had the greatest number (41) while the Santa Cruz River had 27. The remaining localities had a combined 28 private alleles, with the captive ASU Parker Canyon population being the greatest among them (5). Most localities (N=24) had no private alleles.

Population Structure

The program Structure confirmed the ancestry of most sampling localities when the dataset was evaluated for eight genetic clusters (Figure 2). However, there were five localities that did not conform to suspected histories. The ASU captive Monkey Spring population indicated admixture with Yaqui Topminnow, whereas this was previously assumed to be a pure Monkey Spring population. 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 clustered with other Cienega Creek sites. Finally, there was difficulty classifying some populations from the Santa Cruz River and Sharp Spring. Captive samples from Parker Canyon (suspected Sharp Spring lineage) clustered with those from the Santa Cruz River-Nogales, rather than with other Sharp Spring lineage sites. Furthermore, this is an odd pairing due to the large number of unique alleles found exclusively at the Santa Cruz River-Nogales locality.

Despite these 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 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 ($F_{CT} = 0.43$, $P < 0.001$) compared to differences among sampling sites within lineages ($F_{SC} = 0.16$, $P < 0.001$). The greatest source of genetic variation was found within localities (48.0%) rather than among lineages (42.64%), 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 (9.36%).

The patterns revealed through AMOVA were evident when tests for population substructure were conducted in the program Structure (Figure 3). These tests indicated that some lineages can be further subdivided into discrete entities, while others are homogenous. Evaluation of –log likelihood values was used to identify the best K for each group. This yielded K values of 8, 3, and 7 for Groups 1, 2, and 3 respectively. In Group 1, Structure indicated population substructure among Monkey and Cottonwood Springs sites, with the 11 representatives of this lineage being split among six genetic clusters (Figure 3A). However, Structure failed to differentiate among the Bylas Springs and Cienega Creek representatives that were included in this group for the purpose of evaluating the mixed ancestry of Scottsdale Community College fish. This may result from Group 1 being dominated by Monkey and Cotttonwood Springs lineage samples, since Structure is known to struggle with such cases of unbalanced population representation (Wang 2017).

Structure did not find any additional population substructure in Group 2 (Figure 3B) for Bylas Springs, Sonoita Creek (Red Rock Canyon) and Sharp Spring lineages. Results matched those from analysis of the full dataset. Sonoita Creek (Fresno / Coal Mine Canyon) and Cienega Creek lineages were homogenous in Group 3 (Figure 3C), with the exception of Deer Valley High School. The Santa Cruz River lineage in Group 3 was split into three genetic clusters, represented by the ASU captive lineage of Parker Canyon, Phoenix Zoo, and the Santa Cruz River-Nogales.

Effective Population Size

Estimates of N_E were highly variable among sites, ranging from 3 to infinity (Table 7). Eleven sampling events yielded $N_E < 50$, while an additional six had an $N_E < 100$. Eleven sites had an $N_E > 500$, however nine 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 1 to 275, with 23 sampling

events 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 7 to infinity, with one site having upper 95% CI N_E < 50, and an additional two with upper 95% CI N_E < 100. Overall, 12 sampling events yielded upper 95% CI N_E < 500. Twenty-six of the 29 values above 500 (89.7%) 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, 2018 Lousy Canyon, and Robbins Butte Wildlife Area tested positive for a genetic bottleneck. In contrast, the two-phase (TPM) and stepwise (SMM) mutation models 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, as well as Robbins Butte Wildlife Area.

The Garza-Williamson (G-W) *M* Ratio test (Table 7) indicated several more potential bottlenecks, with nine sampling sites (22%) showing evidence of a recent bottleneck. The modified G-W test, which uses fixed loci in its calculations in contrast to all other bottleneck detection methods employed here, showed that every site has experienced a bottleneck event.

Hybridization

NewHybrids revealed the presence of multigenerational Gila by Yaqui Topminnow hybrids among the ASU captive Monkey Spring fish. 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 individuals were assigned with a high posterior probability (Pr > 0.94) of belonging to their respective classifications.

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 an example from the Monkey and Cottonwood Springs lineage 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 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 have faced native fishes 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 basin (Olden and Poff 2005). The AMOVA results revealed that the greatest source of genetic variation was found within localities (48.0%), however variation among lineages remained high (42.64%) and variation among localities was low. Such numbers are close to the partitioning of within-species diversity observed among isolated endorheic basins (Mussmann 2018) 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 lineages: Table 4) and the probable loss of diversity in contemporary time. For example, Sharp Spring lineage is now fixed for allele 159 at locus *Pooc-G49*, and Sonoita Creek (Red Rock Canyon) is fixed for allele 149 at *Pooc-OO56*. Previously these lineages had one additional allele detected at each 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 serve as a proxy for evaluating evolutionary or adaptive potential (Holderegger et al. 2006), but can elucidate population dynamics that may have impacted such 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, and we have not assessed loci that may contribute to the adaptive potential of the species, but the documented continued decline of neutral diversity remains a cause for concern.

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 various 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 results (Table 4; Figure 2). 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 representing 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 located upstream of these sites in 2005. The holding tank had most recently been stocked with 1,000 fish in 1982, in addition to previous stocking events of unrecorded size during the 1970s. Unnamed Drainage was populated by fish from this unintentional release, and documented population histories suggest Tortilla Creek resulted from natural dispersal of fish from Unnamed Drainage. This is also suggested by genetic data, as the population substructure analysis indicated these two sites form a unique genetic cluster within the Monkey and Cottonwood Springs lineage (Figure 3A).

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 last 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 fixed for the same allele at 15/29 microsatellite loci, and Tortilla

Creek is fixed for 16/29 loci. Neither locality exhibits any private alleles. Both have exceptionally low measures of genetic diversity (H_E = 0.188 for both; D68B A_R = 1.83; TTC A_R = 1.72). Likewise, N_E estimates are low for these sites. Unnamed Drainage had an infinite N_E, and the confidence interval for each site encompassed infinity. However, this is likely a consequence of reduced power resulting from the high number of fixed loci at each population (i.e., NeEstimator ignores fixed loci which effectively reduces the number of loci from which NE estimates are derived: Do et al. 2014). However, in these cases the lower bound of the confidence interval can be compared to draw some inference of effective size (Waples and Do 2010). These estimates are again low for each site (lower 95% CI N_E for D68B = 46.9; TTC = 21.2).

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 apparent in all other Gila Topminnow lineages as well, where each lineage contains a base number of fixed loci (≥ 8: 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).

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 viability of a population. In other words, a minimum population size of 50 is required to minimize inbreeding, and 500 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 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 the correction could by chance be greater than estimator for which it is trying to compensate (Waples and Do 2010). This method can also have 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 if adequate data is 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; however, the number of alleles present at each locus was

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 29 total microsatellite loci used for genotyping.

Hybridization

Two populations displayed unique allele frequencies that warranted investigation to determine if they had mixed with non-*P. occidentalis* populations. These were the ASU captive Monkey Spring population, and the wild population in the Santa Cruz River-Nogales. NewHybrids was used to verify that all but four Monkey Spring fish resulted from a hybridization event that occurred 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 fishes, with the exception of two samples in other populations which were identified as F2 hybrids.

Unfortunately the status of the Santa Cruz River samples cannot be determined at this time. Interestingly, this population had just one locus fixed for a single allele, whereas other Gila Topminnow localities were fixed at several additional loci (with the known hybrid ASU Monkey Spring population being excluded from this consideration). Furthermore, this site exhibited a high number of private alleles relative to other Gila Topminnow sites (N=27 compared to N=28 for all 38 other sites combined). Most loci (22 of 29; 75.9%) deviated from Hardy-Weinberg equilibrium, and 145 of 406 locus pairs (35.7%) exhibited linkage disequilibrium. These deviations were exceptionally rare for other Gila Topminnow populations, and were never as extensive when present. Such deviations could be caused by recent hybridization with other *Poeciliopsis* species, or mixing of pure members of another species or subspecies into the population pool (i.e., the Wahlund effect) (Law et al. 2003). Neighboring drainages of the Santa Cruz in Mexico (e.g., Rio Concepción; Rio Sonora) are known to contain a hybrid species (*P. monacha-occidentalis*) which reproduces by hybridogenesis: 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 (Quattro et al. 1992). A hemiclonal species such as *P. monacha-occidentalis* could deviate from Hardy-Weinberg equilibrium due to violations of its underlying assumptions (i.e., sexual reproduction), and show evidence of linkage disequilibrium due to a lack of recombination (Barton and Otto 2005). However, a lack of available *P. monacha* or *P. monacha-occidentalis* reference samples currently prevents the testing of these hypotheses concerning the Santa Cruz River-Nogales sample origins.

Conclusions and Management Implications

Additional analysis remains to be conducted on the existing data, however two points have become apparent that should be considered for future management of the captive and wild populations of this species. The first is that the captive Monkey and Cottonwood Springs lineage at ASU may no longer be representative of its lineage. Methods for hybrid analysis revealed that 92% of the samples from this population are of mixed ancestry with Yaqui Topminnow. As a consequence, this captive population should not be utilized for any management actions unless a method can be determined to identify and isolate these fish from the remainder of the population, with any remaining 'pure' population being genetically reassessed.

Secondly, analyses indicate that genetic drift has played a 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.

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.

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

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

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.

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

Do C, Waples RS, Peel D, et al (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.

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

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.

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.

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

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.

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.

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.

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

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.

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

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

Mussmann SM (2018) Diversification Across a Dynamic Landscape: Phylogeography and Riverscape Genetics of Speckled Dace (*Rhinichthys osculus*) in Western North America. Doctoral Dissertation, University of Arkansas

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 89:348.

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

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.

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

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

Wang J (2017) The computer program structure for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources* 17:981–990.

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.

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.

Table 1. Sampling localities (sites) for Gila Topminnow (*Poeciliopsis occidentalis*) and Yaqui Topminnow (*P. o. sonorensis*) evaluated in this report. The number of samples (N), captive (C) or wild (W) status, two-digit collection year, site locality code, and suspected lineage are also provided for each locality.

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

Table 2. Novel microsatellite loci developed for Gila Topminnow (*Poeciliopsis occidentalis*) and Yaqui Topminnow (*P. o. sonorensis*). The forward (F) and reverse (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 both observed (H_O) and expected (H_E) heterozygosity.

Loci		Primer Sequence	Size (BP)	N_A	H_O	H_E
Pocc02	F	CTAACCGAAGTCAGTGGCAAGC	187-207	5	0.206	0.207
	R	GGCTGAACAATGACAGAGGAGG				
Pocc03	F	GGAAGCAAGTCTAGAATTGACGC	177-196	6	0.159	0.152
	R	AGGAACAGTGTGCTCTTGTTAAGG				
Pocc07	F	AACATGAAGGCCTCACTTGC	165-177	3	0.030	0.024
	R	CCACAGGAGTAAGATGAAGAGAGC				
Pocc09	F	TATGTCGTTGTAGGTGCTCTGG	174-224	7	0.200	0.184
	R	ACCGGATGATCATTACAACAGG				
Pocc15	F	CAGGAACAGGAGCACTCATTGG	123-169	7	0.302	0.297
	R	TTCCAGGACAGAAGCGCTAACC				
Pocc16	F	CACAGGCACACGCATTAGAAGG	125-190	12	0.349	0.345
	R	AAGTAGGATGGCCTGGCAACG				
Pocc18	F	ATGCTCACAGCGCATCTGG	275-329	13	0.480	0.507
	R	TGCGCACCTGTTATTATTCCGC				
Pocc21	F	ACCAAGGCTCAGAATACACACC	115-131	4	0.037	0.029
	R	CAGAATCTGCGCCAGTCTGC				
Pocc25	F	GGAGGTGGCAGCATCATTACG	324-388	16	0.488	0.474
	R	CGCCACATCGTTACATAATAGACG				
Pocc26	F	GCTGCTTGCTTAAGAGTGCG	180-252	16	0.583	0.584
	R	ACCTTGCTATAACCTTGTTGCGC				
Pocc27	F	TCACCTTCAGTGTGAGTTCTCC	124-186	15	0.407	0.396
	R	CGGCTCCATCCTACTCCTATCC				
Pocc28	F	AGTCATCACATCACTGCTGGC	207-251	8	0.165	0.174
	R	ACTTGAGAATGAGCTATGCATGC				
Pocc29	F	ATTGACGACGATGGAACAAGCC	181-228	8	0.187	0.188
	R	AGACCGGTGACCTGTTCATGG				
Pocc31	F	CACACACAAGCCTCAACTTCTACC	122-173	8	0.147	0.140
	R	ACGTAAGAGGAGGAACACAGCC				
Pocc34	F	GAGTCACCGCTTCTCCACATCG	233-415	20	0.490	0.482
	R	GTCGTACAACAGGAGCCAGAGG				
Pocc36	F	TCCTACTCCGACACTGTGTACC	219-362	20	0.501	0.481
	R	GTGAGATATTCCAGGCATGTCGC				
Pocc38	F	TTGAGTGTATGTATGTGTCCATCC	182-265	21	0.498	0.500
	R	ATCAAGAGCATCAGAAGACCGG				
Pocc40	F	ATGATTACTGACATTCTGACCAGG	176-208	5	0.052	0.053
	R	ATAACGGCAATTCAAGAGCTGC				
Pocc43	F	AAGATGCTCAGCAATCACCACG	189-243	4	0.029	0.024
	R	TCTTCCTCTCCTCGCTGAGC				
Pocc44	F	GCAGTTATCACAGTTGCTTGTGC	263-403	24	0.544	0.529
	R	GCTGATAGACACGAGCAACTCC				
Pocc45	F	TGTGGAAGTAGAACCAACAACAGG	158-256	11	0.047	0.038
	R	GCAGTGAAGGTTCAACTCCAGC				
Pocc47	F	AGTTGCTTGTGGTTCTGACAGC	195-274	15	0.386	0.380
	R	GGAGACTGACTCATGACTGTCCG				

Table 3. Genetic diversity estimates for all Gila Topminnow (*Poeciliopsis occidentalis*) and Yaqui Topminnow (*P. o. sonorensis*) sampling 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, observed (H_O) and expected (H_E) heterozygosity, fixation index (F), allelic richness (A_R) and private alleles are also provided for each sampled population.

Population	Monomorphic	H_O	H_E	F	A_R	Private
C18BYL	17	0.121	0.124	0.005	1.65	0
C18CNG	13	0.239	0.236	-0.031	2.66	1
C18MNK	0	0.562	0.568	0.004	4.37	3
C18PRK	8	0.414	0.399	-0.054	3.27	5
C18RRK	17	0.186	0.175	-0.079	1.79	0
C18SHP	9	0.310	0.302	-0.025	2.57	0
W17D68B	15	0.193	0.188	-0.037	1.83	0
W17LBC	11	0.311	0.305	-0.037	2.72	1
W17SCN	1	0.541	0.488	0.022	3.89	27
W17TTC	16	0.190	0.188	-0.036	1.72	0
W18ADW	11	0.260	0.274	0.046	2.37	2
W18CCLC	14	0.263	0.258	-0.027	2.62	1
W18CCPC	9	0.231	0.246	0.159	3.09	3
W18CSE	13	0.312	0.298	-0.047	2.72	2
W18CWS	20	0.152	0.142	-0.085	1.5	0
W18DVHS	15	0.149	0.151	-0.011	1.93	0
W18LCN	14	0.234	0.239	-0.047	1.72	1
W18PZ	8	0.365	0.372	0.001	3.5	2
W18RSAC	9	0.344	0.341	-0.005	2.9	3
W18SBC	13	0.259	0.252	-0.035	2.43	0
W18SPR	17	0.090	0.096	0.029	1.79	0
W18WS20	11	0.333	0.322	-0.053	3.5	1
W18WS392	15	0.175	0.179	-0.002	1.91	1
W19ASDM	11	0.236	0.250	0.185	2.74	0
W19BSP	11	0.278	0.275	-0.024	2.34	0
W19BYL	18	0.073	0.070	-0.026	1.61	0
W19CMC	9	0.263	0.262	-0.016	2.31	1
W19FCN	11	0.259	0.272	0.025	2.45	1
W19LCN	14	0.169	0.192	0.025	1.66	0
W19LMC	12	0.232	0.232	0.000	2.17	0
W19LYC	12	0.240	0.239	0.017	1.93	0
W19MDS	13	0.240	0.228	-0.056	2.23	0
W19MNK	11	0.315	0.314	-0.025	3.09	0
W19RBW	11	0.275	0.279	0.011	2.02	0
W19RCT	12	0.247	0.254	0.014	2.6	0
W19SCC	9	0.285	0.298	0.019	2.64	0
W19SSP	10	0.305	0.300	-0.024	2.72	0
W19SWC	19	0.079	0.081	0.044	1.56	0
W19TMB	15	0.202	0.197	-0.035	2.01	0
W19TUC	11	0.303	0.306	0.021	2.77	0
Yaqui Topminnow	2	0.432	0.423	-0.033	3.9	41

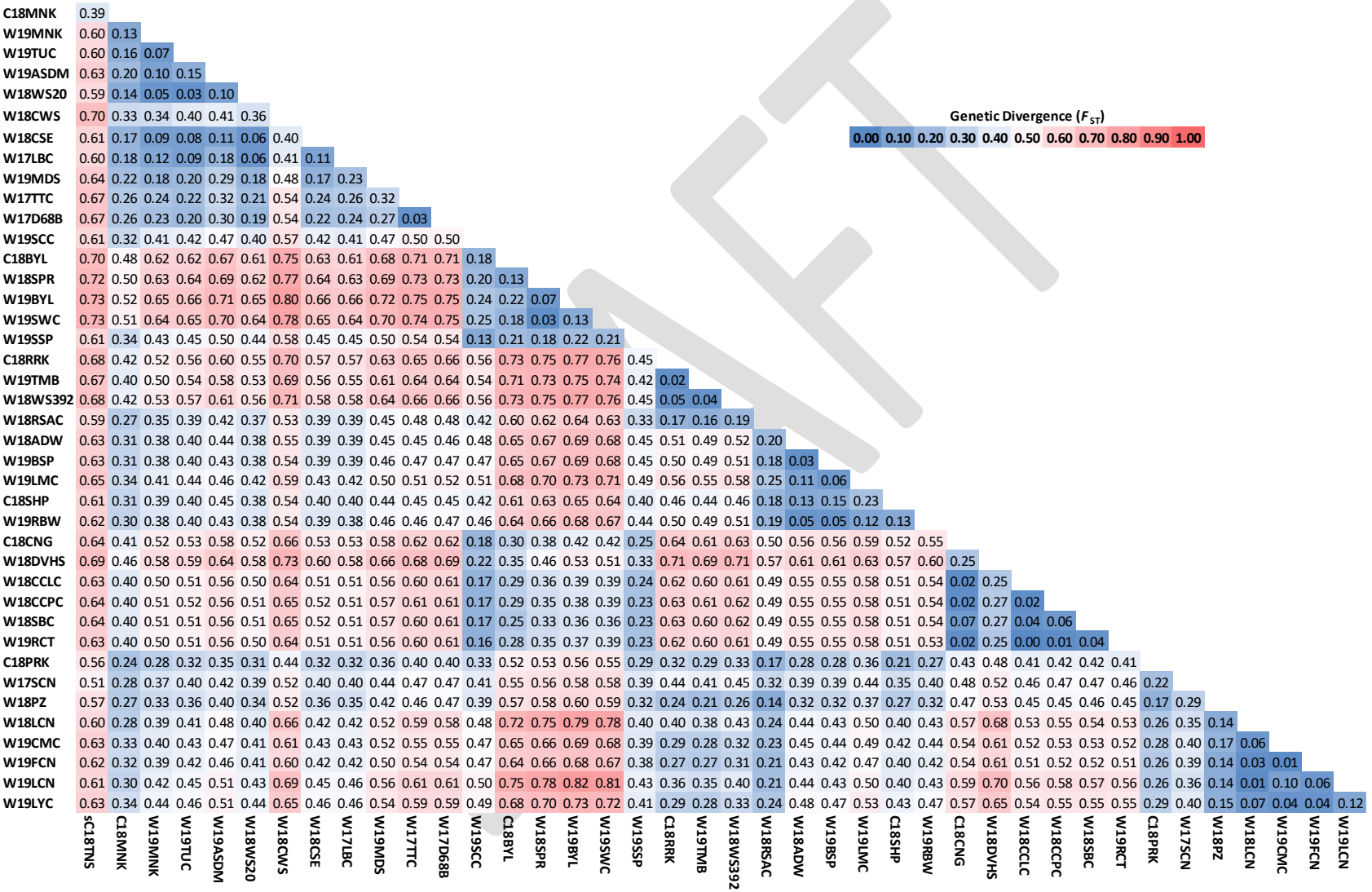
Table 4. Microsatellite loci that are fixed for specific alleles in each of the eight Gila Topminnow (*Poeciliopsis occidentalis*) genetic lineages evaluated in this report. The number associated with each locus for each lineage represents the allele (measured in base pairs) that is fixed in that lineage. Yaqui Topminnow were fixed at *Pocc07* and *Pooc-OO56*, however these loci rarely amplified in this species. 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. Seven Santa Cruz River loci (†) polymorphic in the population north of Nogales, possibly as a result of still undetermined hybridization or inclusion of another species in sampling for that locality.

Locus	Monkey and Cottonwood Springs	Bylas Springs	Sonoita Creek (Red Rock Canyon)	Sharp Spring	Cienega Creek	Santa Cruz River	Sonoita Creek (Fresno / Coal Mine Canyon)	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	246		207			
Pocc29								
Pocc31		157	157					
Pocc34								
Pocc36		312						
Pocc38								
Pocc40		192	192		192	192†		
Pocc43	189	189	189	189	189	189	189	
Pocc44								
Pocc45	162	162	162	162		162†	162	
Pocc47								
Pooc-4-44	107	107	107	107	107	107†	107	
Pooc-G49	159		159	159			159	
Pooc-G53	101	101	101	101	101	101†	101	
Pooc-OO56		149	149			149†		149*
Pooc-C15								
Pooc-G10	192	192	192	192	192	192†	192	180
Pooc-LL53								
Total Monomorphic	**9**	**14**	**15**	**8**	**8**	**8**	**9**	**3**

Table 5. The number of private alleles detected in each Gila Topminnow (*Poeciliopsis occidentalis*) as well as Yaqui Topminnow (*P. o. sonorensis*). The number of alleles unique to a lineage at each microsatellite locus are shown. Lineages of known mixed ancestries (e.g., captive Monkey Spring lineage at Arizona State University) were not included in calculations.

Locus	Monkey and Cottonwood Springs	Bylas Springs	Sonoita Creek (Red Rock Canyon)	Sharp Spring	Cienega Creek	Santa Cruz River	Sonoita Creek (Fresno / Coal Mine Canyon)	Yaqui Topminnow
Pocc02						1		1
Pocc03	1					1		1
Pocc07						1		1
Pocc09						1		2
Pocc15						1		2
Pocc16	1					1		2
Pocc18						2		2
Pocc21						1		1
Pocc25						3		3
Pocc26	1				1	1		4
Pocc27	1			1	1	1		
Pocc28						1	2	2
Pocc29								1
Pocc31	1					1		
Pocc34	1					2		5
Pocc36	1				1	1	1	7
Pocc38								5
Pocc40	1					1		1
Pocc43								3
Pocc44	2		1	1		2		2
Pocc45					1	2		6
Pocc47	2					2		3
Pooc-4-44								2
Pooc-G49						1		
Pooc-G53						2		2
Pooc-OO56					1		1	
Pooc-C15	2	1		2	1	3	1	6
Pooc-G10						1		1
Pooc-LL53	2			1	1	5		4
Total Private	**16**	**1**	**1**	**5**	**7**	**38**	**5**	**69**

Table 6. Pairwise estimates of genetic divergence (F_{ST}) calculated from 29 microsatellite loci. Colors range from low divergence (Blue = 0.0) to high (Red = 1.0). Locality names are represented by a combination of status, two-digit collection year, and locality code from Table 1.



	sC18TNS	C18MNK	W19MNK	W19TUC	W19ASDM	W18WS20	W18CWS	W18CSE	W17LBC	W19MDS	W17TTC	W17D68B	W19SCC	C18BYL	W18SPR	W19BYL	W19SWC	W19SSP	C18RRK	W19TMB	W18WS392	W18RSAC	W18ADW	W19BSP	W19LMC	C18SHP	W19RBW	C18CNG	W18DVHS	W18CCLC	W18CCPC	W18SBC	W19RCT	C18PRK	W17SCN	W18PZ	W18LCN	W19CMC	W19FCN	W19LCN
C18MNK	0.39																																								
W19MNK	0.60	0.13																																							
W19TUC	0.60	0.16	0.07																																						
W19ASDM	0.63	0.20	0.10	0.15																																					
W18WS20	0.59	0.14	0.05	0.03	0.10																																				
W18CWS	0.70	0.33	0.34	0.40	0.41	0.36																																			
W18CSE	0.61	0.17	0.09	0.08	0.11	0.06	0.40																																		
W17LBC	0.60	0.18	0.12	0.09	0.18	0.06	0.41	0.11																																	
W19MDS	0.64	0.22	0.18	0.20	0.29	0.18	0.48	0.17	0.23																																
W17TTC	0.67	0.26	0.24	0.22	0.32	0.21	0.54	0.24	0.26	0.32																															
W17D68B	0.67	0.26	0.23	0.20	0.30	0.19	0.54	0.22	0.24	0.27	0.03																														
W19SCC	0.61	0.32	0.41	0.42	0.47	0.40	0.57	0.42	0.41	0.47	0.50	0.50																													
C18BYL	0.70	0.48	0.62	0.62	0.67	0.61	0.75	0.63	0.61	0.68	0.71	0.71	0.18																												
W18SPR	0.72	0.50	0.63	0.64	0.69	0.62	0.77	0.64	0.63	0.69	0.73	0.73	0.20	0.13																											
W19BYL	0.73	0.52	0.65	0.66	0.71	0.65	0.80	0.66	0.66	0.72	0.75	0.75	0.24	0.22	0.07																										
W19SWC	0.73	0.51	0.64	0.65	0.70	0.64	0.78	0.65	0.64	0.70	0.74	0.75	0.25	0.18	0.03	0.13																									
W19SSP	0.61	0.34	0.43	0.45	0.50	0.44	0.58	0.45	0.45	0.50	0.54	0.54	0.13	0.21	0.18	0.22	0.21																								
C18RRK	0.68	0.42	0.52	0.56	0.60	0.55	0.70	0.57	0.57	0.63	0.65	0.66	0.56	0.73	0.75	0.77	0.76	0.45																							
W19TMB	0.67	0.40	0.50	0.54	0.58	0.53	0.69	0.56	0.55	0.61	0.64	0.64	0.54	0.71	0.73	0.75	0.74	0.42	0.02																						
W18WS392	0.68	0.42	0.53	0.57	0.61	0.56	0.71	0.58	0.58	0.64	0.66	0.66	0.56	0.73	0.75	0.77	0.76	0.45	0.05	0.04																					
W18RSAC	0.59	0.27	0.35	0.39	0.42	0.37	0.53	0.39	0.39	0.45	0.48	0.48	0.42	0.60	0.62	0.64	0.63	0.33	0.17	0.16	0.19																				
W18ADW	0.63	0.31	0.38	0.40	0.44	0.38	0.55	0.39	0.39	0.45	0.45	0.46	0.48	0.65	0.67	0.69	0.68	0.45	0.51	0.49	0.52	0.20																			
W19BSP	0.63	0.31	0.38	0.40	0.43	0.38	0.54	0.39	0.39	0.46	0.47	0.47	0.47	0.65	0.67	0.69	0.68	0.45	0.50	0.49	0.51	0.18	0.03																		
W19LMC	0.65	0.34	0.41	0.44	0.46	0.42	0.59	0.43	0.42	0.50	0.51	0.52	0.51	0.68	0.70	0.73	0.71	0.49	0.56	0.55	0.58	0.25	0.11	0.06																	
C18SHP	0.61	0.31	0.39	0.40	0.45	0.38	0.54	0.40	0.40	0.44	0.45	0.45	0.42	0.61	0.63	0.65	0.64	0.40	0.46	0.44	0.46	0.18	0.13	0.15	0.23																
W19RBW	0.62	0.30	0.38	0.40	0.43	0.38	0.54	0.39	0.38	0.46	0.46	0.47	0.46	0.64	0.66	0.68	0.67	0.44	0.50	0.49	0.51	0.19	0.05	0.05	0.12	0.13															
C18CNG	0.64	0.41	0.52	0.53	0.58	0.52	0.66	0.53	0.53	0.58	0.62	0.62	0.18	0.30	0.38	0.42	0.42	0.25	0.64	0.61	0.63	0.50	0.56	0.56	0.59	0.52	0.55														
W18DVHS	0.69	0.46	0.58	0.59	0.64	0.58	0.73	0.60	0.58	0.66	0.68	0.69	0.22	0.35	0.46	0.53	0.51	0.33	0.71	0.69	0.71	0.57	0.61	0.61	0.63	0.57	0.60	0.25												
W18CCLC	0.63	0.40	0.50	0.51	0.56	0.50	0.64	0.51	0.51	0.56	0.60	0.61	0.17	0.29	0.36	0.39	0.39	0.24	0.62	0.60	0.61	0.49	0.55	0.55	0.58	0.51	0.54	0.02	0.25											
W18CCPC	0.64	0.40	0.51	0.52	0.56	0.51	0.65	0.52	0.51	0.57	0.61	0.61	0.17	0.29	0.35	0.38	0.39	0.23	0.63	0.61	0.62	0.49	0.55	0.55	0.58	0.51	0.54	0.02	0.27	0.02										
W18SBC	0.64	0.40	0.51	0.51	0.56	0.51	0.65	0.52	0.51	0.57	0.60	0.61	0.17	0.25	0.33	0.36	0.36	0.23	0.63	0.60	0.62	0.49	0.55	0.55	0.58	0.51	0.54	0.07	0.27	0.04	0.06									
W19RCT	0.63	0.40	0.50	0.51	0.56	0.50	0.64	0.51	0.51	0.56	0.60	0.61	0.16	0.28	0.35	0.37	0.39	0.23	0.62	0.60	0.61	0.49	0.55	0.55	0.58	0.51	0.53	0.02	0.25	0.00	0.01	0.04								
C18PRK	0.56	0.24	0.28	0.32	0.35	0.31	0.44	0.32	0.32	0.36	0.40	0.40	0.33	0.52	0.53	0.56	0.55	0.29	0.32	0.29	0.33	0.17	0.28	0.28	0.36	0.21	0.27	0.43	0.48	0.41	0.42	0.42	0.41							
W17SCN	0.51	0.28	0.37	0.40	0.42	0.39	0.52	0.40	0.40	0.44	0.47	0.47	0.41	0.55	0.56	0.58	0.58	0.39	0.44	0.41	0.45	0.32	0.39	0.39	0.44	0.35	0.40	0.48	0.52	0.46	0.47	0.47	0.46	0.22						
W18PZ	0.57	0.27	0.33	0.36	0.40	0.34	0.52	0.36	0.35	0.42	0.46	0.47	0.39	0.57	0.58	0.60	0.59	0.32	0.24	0.21	0.26	0.14	0.32	0.32	0.37	0.27	0.32	0.47	0.53	0.45	0.45	0.46	0.45	0.17	0.29					
W18LCN	0.60	0.28	0.39	0.41	0.48	0.40	0.66	0.42	0.42	0.52	0.59	0.58	0.48	0.72	0.75	0.79	0.78	0.40	0.40	0.38	0.43	0.24	0.44	0.43	0.50	0.40	0.43	0.57	0.68	0.53	0.55	0.54	0.53	0.26	0.35	0.14				
W19CMC	0.63	0.33	0.40	0.43	0.47	0.41	0.61	0.43	0.43	0.52	0.55	0.55	0.47	0.65	0.66	0.69	0.68	0.39	0.29	0.28	0.32	0.23	0.45	0.44	0.49	0.42	0.44	0.54	0.61	0.52	0.53	0.53	0.52	0.28	0.40	0.17	0.06			
W19FCN	0.62	0.32	0.39	0.42	0.46	0.41	0.60	0.42	0.42	0.50	0.54	0.54	0.47	0.64	0.66	0.68	0.67	0.38	0.27	0.27	0.31	0.21	0.43	0.42	0.47	0.40	0.42	0.54	0.61	0.51	0.52	0.52	0.51	0.26	0.39	0.14	0.03	0.01		
W19LCN	0.61	0.30	0.42	0.45	0.51	0.43	0.69	0.45	0.46	0.56	0.61	0.61	0.50	0.75	0.78	0.82	0.81	0.43	0.36	0.35	0.40	0.21	0.44	0.43	0.50	0.40	0.43	0.59	0.70	0.56	0.58	0.57	0.56	0.26	0.36	0.14	0.01	0.10	0.06	
W19LYC	0.63	0.34	0.44	0.46	0.51	0.44	0.65	0.46	0.46	0.54	0.59	0.59	0.49	0.68	0.70	0.73	0.72	0.41	0.29	0.28	0.33	0.24	0.48	0.47	0.53	0.43	0.47	0.57	0.65	0.54	0.55	0.55	0.55	0.29	0.40	0.15	0.07	0.04	0.04	0.12

Table 7. Results of effective population size (N_E) and genetic bottleneck tests for Gila Topminnow (*Poeciliopsis occidentalis*) populations. The 95% jackknife confidence interval (95% CI) is displayed for N_E estimates. 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 two-tailed P-value that was Bonferroni-adjusted (α = 0.0012). Assessment of a shift in the mode of allele frequencies was also conducted (shifted = indication of a genetic bottleneck). The Garza-Williamson *M* Ratio (G-W) and the modified version (G-W Mod) were also calculated. Values < 0.68 are significant. Sample locality names are represented by a combination of status, two-digit collection year, and locality code from Table 1. Significant values are bolded.

Population	N_E	95% CI	IAM	TPM	SMM	Mode-Shift	G-W	G-W Mod
C18BYL	446	22.1 - ∞	0.677	0.791	0.733		0.708	**0.187**
C18CNG	83	42 - 367.9	0.860	0.021	0.008		0.770	**0.260**
C18MNK	51	29.9 - 109	**0.000**	0.004	0.051		0.727	**0.453**
C18PRK	34	21.1 - 60.7	**0.000**	0.320	0.919		0.724	**0.324**
C18RRK	44	18.3 - 280	0.008	0.850	0.970		0.694	**0.200**
C18SHP	120	56.5 - 1,460.9	0.009	0.869	0.898		0.695	**0.262**
W17D68B	∞	46.9 - ∞	0.030	0.583	0.855		0.804	**0.197**
W17LBC	222	69.3 - ∞	0.007	0.551	0.966		0.847	**0.261**
W17SCN	4	2.7 - 7	0.006	0.831	0.305		**0.649**	**0.454**
W17TTC	73	21.2 - ∞	0.002	0.048	0.094		0.778	**0.189**
W18ADW	41	17.5 - 194.7	0.006	0.832	0.832		0.696	**0.242**
W18CCLC	138	51.8 - ∞	0.277	0.095	0.064		0.782	**0.255**
W18CCPC	11	1.3 - 1,352.9	0.083	**0.000**	**0.000**		0.719	**0.305**
W18CSE	245	77 - ∞	0.013	1.000	0.495		0.821	**0.259**
W18CWS	726	23.8 - ∞	0.004	0.010	0.010		**0.623**	**0.171**
W18DVHS	∞	50.2 - ∞	1.000	0.135	0.091		0.716	**0.218**
W18LCN	5	1.1 - ∞	**0.001**	0.035	0.095	**Shifted**	0.728	**0.205**
W18PZ	402	119.2 - ∞	0.019	0.157	0.070		0.774	**0.347**
W18RSAC	∞	149.5 - ∞	0.004	0.277	0.729		**0.672**	**0.278**
W18SBC	∞	274.5 - ∞	0.074	0.323	0.211		0.761	**0.239**
W18SPR	147	22.7 - ∞	0.077	0.002	0.002		0.797	**0.192**
W18WS20	640	114 - ∞	0.048	0.081	0.024		0.838	**0.311**
W18WS392	48	20.9 - 268.4	0.358	0.463	0.358		0.732	**0.208**
W19ASDM	19	3.6 - 162.7	1.000	0.009	0.002		**0.645**	**0.275**
W19BSP	57	29.8 - 172.2	0.002	0.551	0.832		**0.649**	**0.235**
W19BYL	∞	10.7 - ∞	0.083	0.009	0.009		0.772	**0.182**
W19CMC	257	50.2 - ∞	0.154	0.522	0.330		0.819	**0.261**
W19FCN	261	51.3 - ∞	0.067	0.325	0.167		0.754	**0.267**
W19LCN	3	1.3 - 99.3	0.421	0.639	0.639	**Shifted**	0.709	**0.196**
W19LMC	∞	125 - ∞	0.057	0.517	0.378		**0.664**	**0.233**
W19LYC	46	17.1 - ∞	0.013	0.225	0.263		0.734	**0.213**
W19MDS	94	35.2 - ∞	0.083	0.404	0.231		0.801	**0.232**
W19MNK	179	63.2 - ∞	0.030	1.000	0.523		0.853	**0.285**
W19RBW	367	67.1 - ∞	**0.000**	0.010	0.024	**Shifted**	**0.621**	**0.216**
W19RCT	∞	101.2 - ∞	0.404	0.145	0.089		0.791	**0.259**
W19SCC	135	60.7 - 190,783.6	0.003	0.841	0.452		**0.622**	**0.263**
W19SSP	326	83.7 - ∞	0.087	0.374	0.182		**0.656**	**0.273**
W19SWC	∞	33.4 - ∞	0.557	0.131	0.105		0.793	**0.174**
W19TMB	∞	201.7 - ∞	0.091	0.715	0.463		0.731	**0.212**
W19TUC	48	28.1 - 108	0.006	0.832	0.369		0.846	**0.271**
Yaqui Topminnow	58	36 - 116.2	0.194	0.017	0.002		0.755	**0.447**

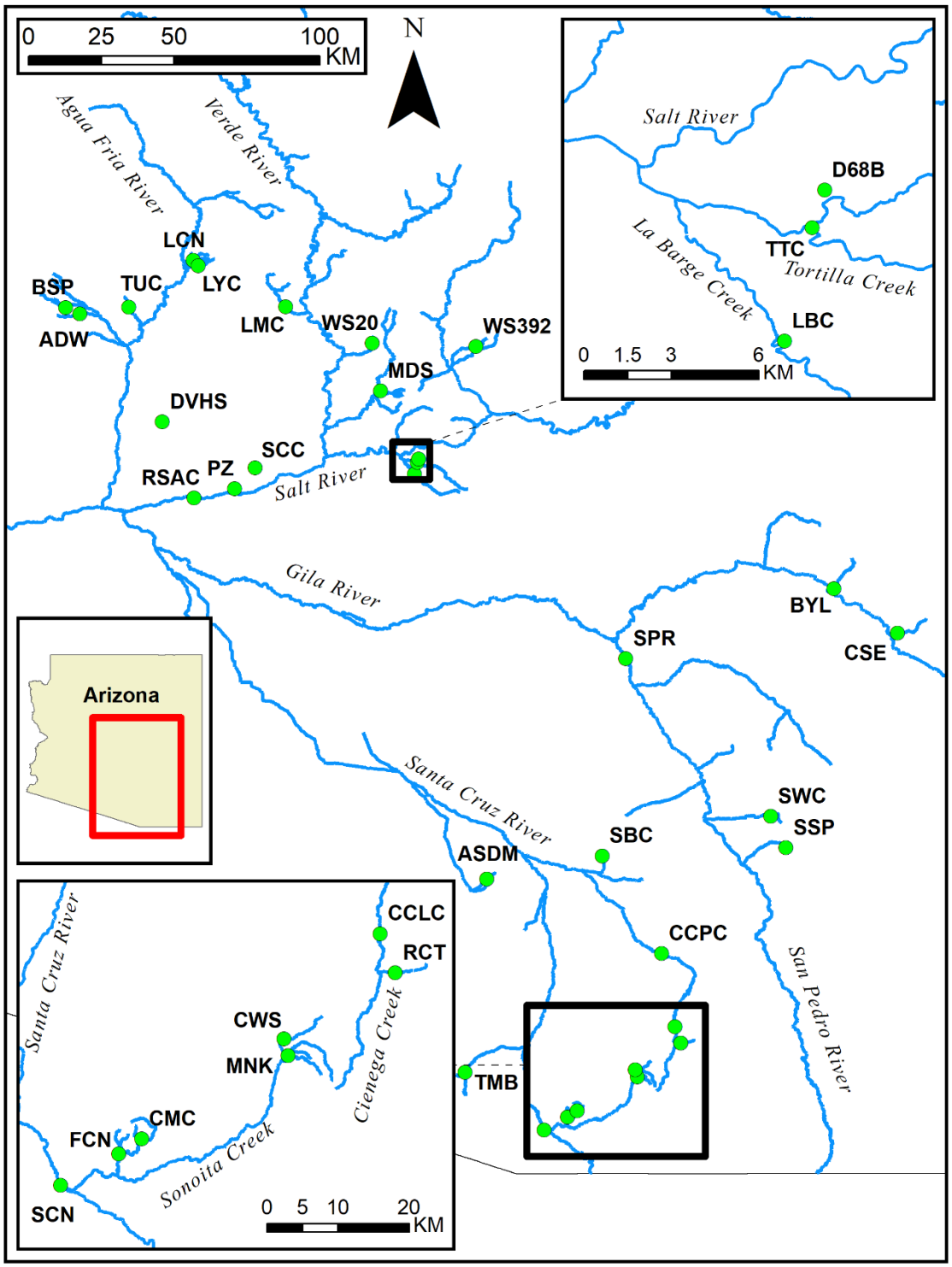


Figure 1. Gila Topminnow (*Poeciliopsis occidentalis*) sampling localities located in Arizona (center left inset map). The Upper right inset shows detail within the La Barge and Tortilla Creek drainages. The lower left inset shows detail within the Santa Cruz River, Sonoita Creek, and Cienega Creek drainages.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

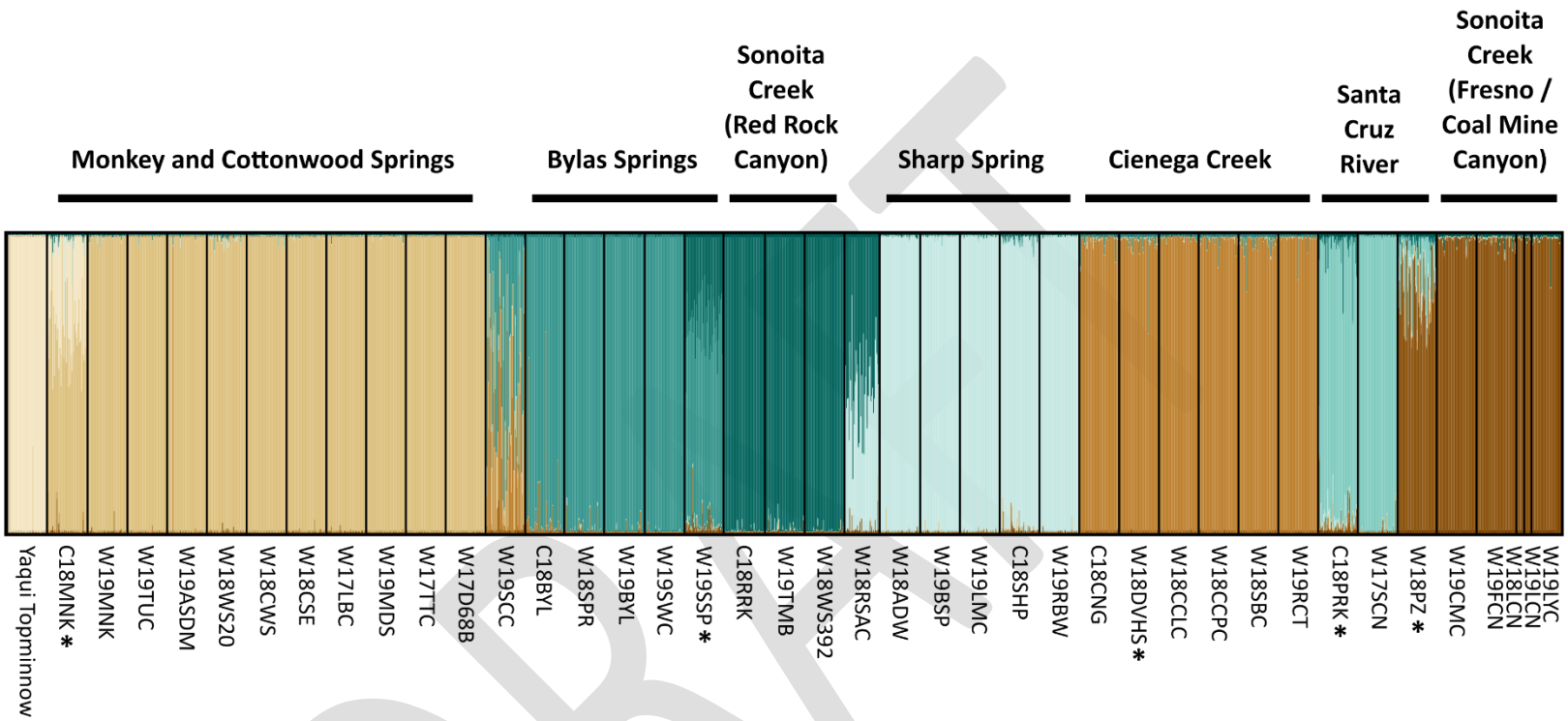


Figure 2. Population structure results showing the assignment of each sampling locality to each of eight genetic clusters representing Gila Topminnow (*Poeciliopsis occidentalis*) lineages defined in earlier works. Each genetic cluster is represented by a color, and each individual sample is represented 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 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. Those denoted with an asterisk (*) were either assigned to a cluster that conflicted with management records, or showed evidence of admixture with another lineage.

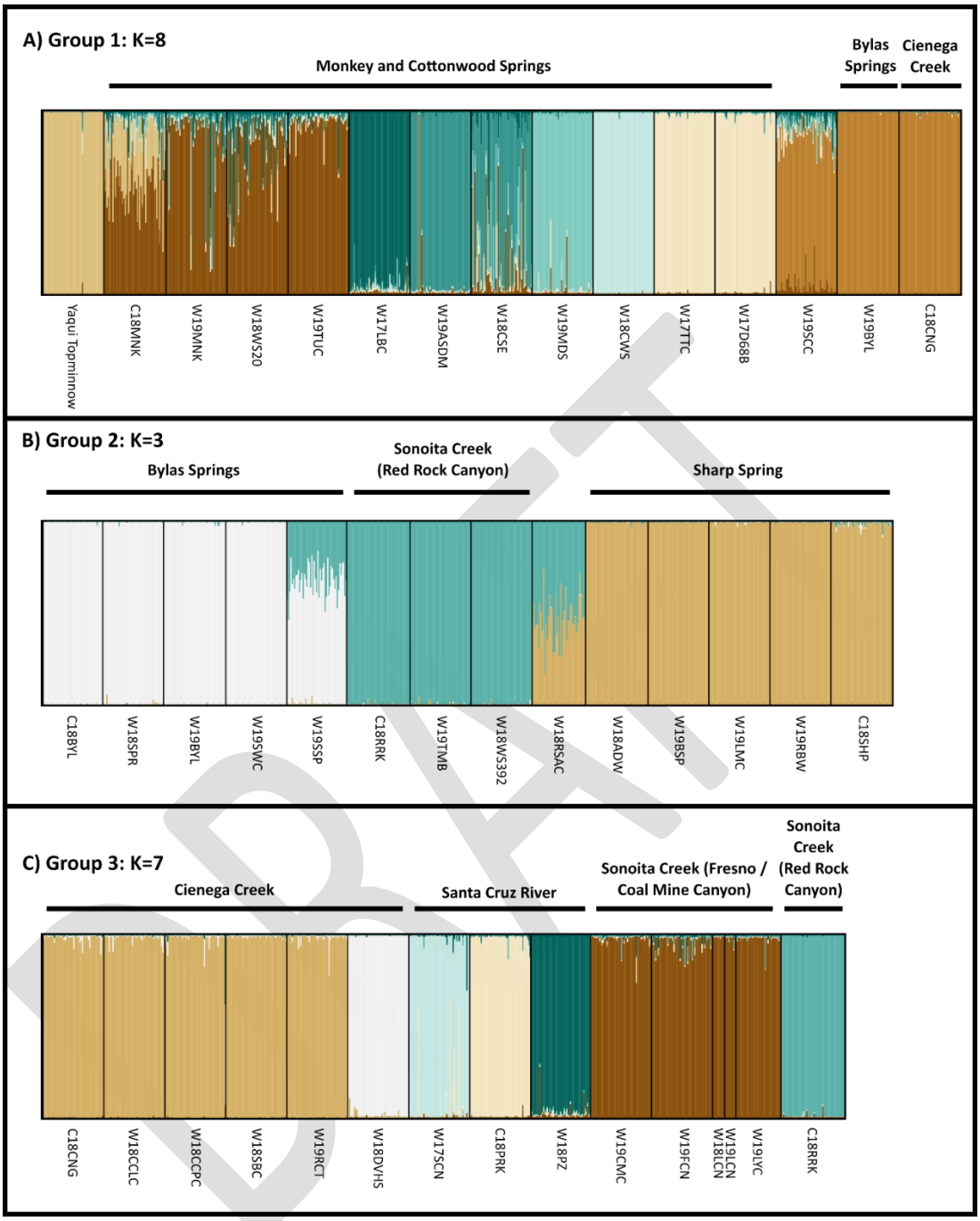


Figure 3. Subpopulation structure for the eight genetic lineages shown in Figure 2. Each Gila Topminnow (*Poeciliopsis occidentalis*) lineage was evaluated for within-lineage population structuring. Each genetic cluster is represented by a color, and each individual sample is represented by a vertical 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 refugia, meaning a total of 16 genetic clusters were detected across all eight lineages previously outlined in Figure 2.

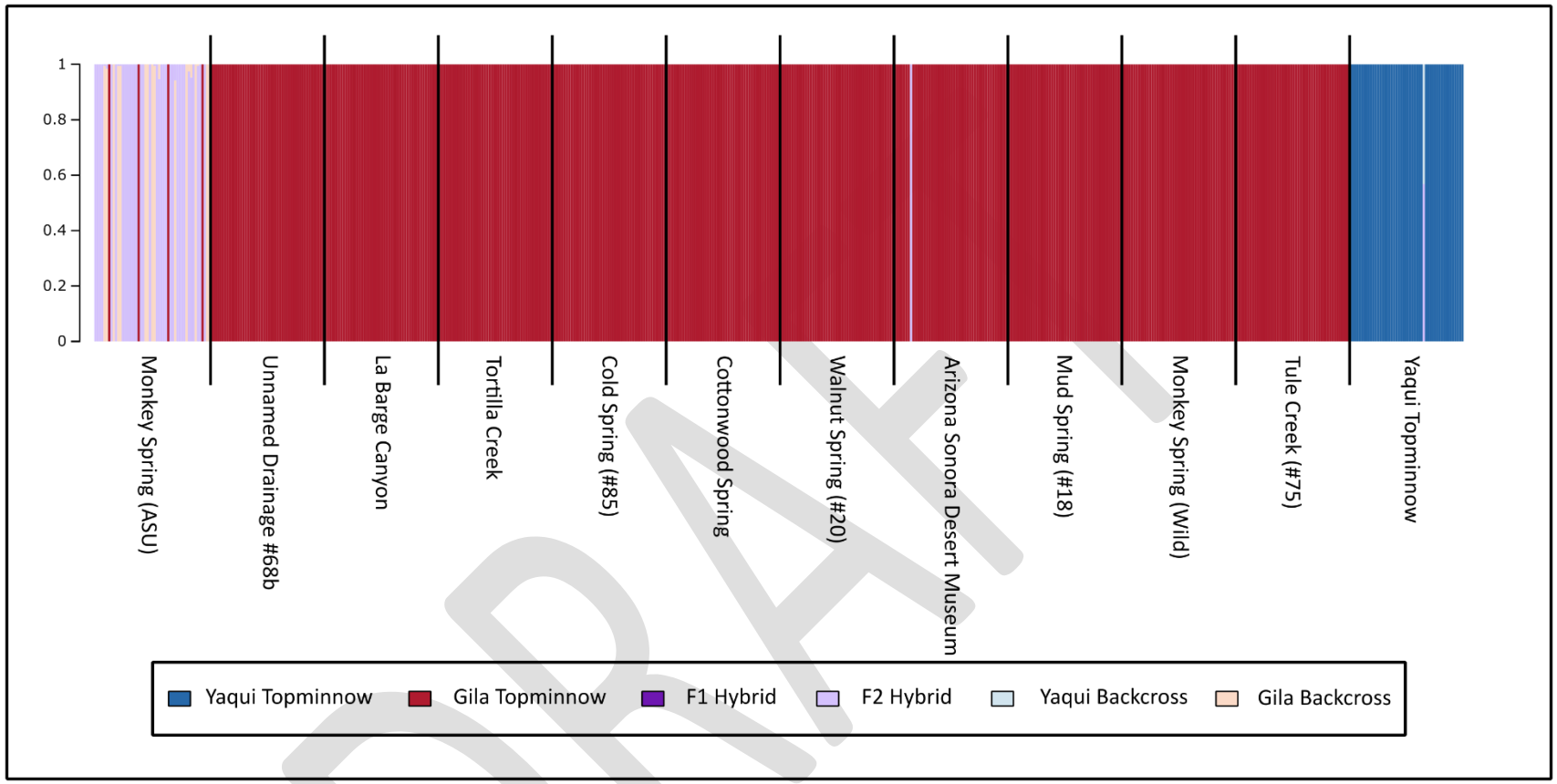


Figure 4. Hybrid analysis of the Monkey and Cottonwood Springs lineage verified that the captive Monkey Spring population at Arizona State University (ASU) represents a hybrid group of Gila Topminnow (*Poeciliopsis occidentalis*) and Yaqui Topminnow (*P. o. sonorensis*). Each vertical bar in the figure represents an individual sample (N=601). 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 Monkey Spring fish are mostly second-generation hybrids (F2 hybrid) and Gila Topminnow backcrosses [i.e., offspring of a first-generation (F1) hybrid by pure Gila Topminnow]. All assignments were at a high posterior probability (Pr > 0.94).