Taxonomic Uncertainty and Phylogenomics: Rescuing a Contentious Species Complex from the Anomaly Zone (6438)

Tyler K. Chafin[1]

Marlis R. Douglas[1]

Max R. Bangs[1,2]

Steven M. Mussmann[1]

Michael E. Douglas[1]

[1]Department of Biological Sciences, University of Arkansas, Fayetteville, Arkansas 72701, USA; Email: (TKC) tkchafin@uark.edu, send reprint requests to this address; (MRD) mrd1@uark.edu; (SMM) smussmann@gmail.com; (MED) med1@uark.edu.

Current address:

[2]Department of Biological Science, Florida State University, Tallahassee, Florida, 32306, USA; Email: mbangs@bio.fsu.edu (present address)

Running head: *Phylogenomics in the anomaly zone*

Keywords: phylogenetics; conservation genetics; RADseq; anomaly zone

Disclosure statement: Authors have nothing to disclose.

Abstract (242/250)

Species are an indisputable unit for biodiversity conservation, yet their delimitation is fraught with both conceptual and methodological difficulties. A classic example is the taxonomic controversy surrounding the *Gila robusta* complex in the lower Colorado River of southwestern North America. Three species were originally defined according to subtle morphological differences with weak diagnostic power, with traditional genetic analyses failing to support the nominal species. Consequently, the complex was re-defined as a single polytypic unit, with the proposed 'threatened' status of two putative species withdrawn at the federal level. Here, we utilized dense spatial and genomic sampling (N=387 and >22k loci) to re-evaluate the status of the complex, based on SNP-based coalescent and polymorphism-aware phylogenetic models. In doing so, all three species were supported as evolutionarily independent lineages, despite widespread phylogenetic discordance. To understand this discrepancy with past studies, we categorized evolutionary mechanisms driving discordance. We tested (and subsequently rejected) prior hypotheses suggesting that phylogenetic discord in the complex was hybridization-driven. Instead, we found the *G. robusta* complex to have diverged within the 'anomaly zone' of tree space and, as such, have accumulated inconsistent patterns of diversity which have confounded prior studies. After extending these analyses with phylogeographic modeling, we propose that this is reflective of a rapid radiation promoted by Plio-Pleistocene tectonism. Our results not only support resurrection of the three species as distinct entities, but also offer an empirical example of how phylogenetic discordance can be categorized in other recalcitrant taxa.

Introduction (1260)

Complex evolutionary histories remain consistently difficult to disentangle, despite a recent paradigm shift towards the development of increasingly comprehensive datasets (e.g. Edwards 2009; Giarla and Esselstyn 2015). Regardless of these efforts, phylogenetic uncertainty is still prevalent, and with wide-ranging consequences on the study of macroevolutionary patterns (Stadler et al. 2016; Pereira and Schrago 2018), trait evolution (Hahn and Nakhleh 2016; Mendes et al. 2016; Wu et al. 2018), and ecological and biogeographic processes (Rangel et al. 2015; McVay et al. 2017).

Importantly, phylogenetic uncertainty also translates to taxonomic uncertainty. This is because modern systematic taxonomy fundamentally describes homology [i.e. Darwin's (1859) ‘propinquity of descent’ (Simpson 1961)], which, by definition, requires a phylogenetic context. Phylogenetic uncertainty in this sense can manifest itself as a soft polytomy (= ‘honest’ uncertainty), the erroneous promotion of non-monophyletic clades, or controversial ‘splitting’ versus ‘lumping.’ Incomplete or biased sampling is often a driver of this disparity (Ahrens et al. 2016; Reddy et al. 2017). Here, narrow taxon sampling may introduce substantial ascertainment bias (=systematic deviations due to sampling). On the other hand, a broader yet sparse sampling regime often fails to sample cryptic lineages (Heath et al. 2008) — with subsequent impacts on both the delimitation of species (Pante et al. 2015; Linck et al. 2019) and study of their traits (Beaulieu and O’Meara 2018).

These sources of uncertainty culminate in topologies that often fluctuate with regard to sampling designs or methodologies, and this results in taxonomic uncertainty [e.g. Ctenophora versus Porifera as sister to all other animals (Pisani et al. 2015; Whelan et al. 2015; Simion et al. 2017)]. Access to genome-scale data has alleviated some of these issues by offering a level of

precision not possible with single-gene phylogenies (Philippe et al. 2005). However, their inherent complexity and heterogeneity introduces new problems, and consequently, additional sources of phylogenetic uncertainty.

Gene tree heterogeneity is a ubiquitous source of discordance in genomic data, and "noise" as a source of this variance must consequently be partitioned from "signal" (where "noise" is broadly categorized as systematic or stochastic error). Large genomic datasets can reduce stochastic error (Kumar et al. 2012), yet it still remains a prevalent issue when individual genes are examined (Springer and Gatesy 2016). On the other hand, systematic error in phylogenomics may represent a probabilistic bias towards incongruence that is inherent to the evolutionary process itself (Maddison 1997). This, in turn, exemplifies the complications introduced by genomic data: As genomic resolution increases, so also does the probability of sampling unmodeled processes (Rannala and Yang 2008; Lemmon and Lemmon 2013). This potential (i.e., simultaneously decreasing stochastic error as systematic error increases) produces the very real possibility of building a highly supported tree that is ultimately incorrect.

Certain demographic histories are more predisposed to systematic error than others. For instance, when effective population sizes are large and speciation events exceptionally rapid, time between divergence events may be insufficient to sort ancestral variation, such that the most probable gene topology will conflict with the underlying species branching pattern. This results in what has been coined an "anomaly zone" of tree space (i.e., dominated by anomalous gene trees (AGTs); Degnan and Rosenberg 2006). Inferring species trees is demonstrably difficult in this region (Liu and Edwards 2009), and exceedingly so if additional sources of phylogenetic discordance, such as hybridization, are also apparent (Bangs et al. 2018).

In clades with such complex histories, it is often unclear where the source of poor support and/or topological conflict resides. Yet, to analytically account for gene tree conflict, it is necessary to categorize these sources and select approaches accordingly. Failure to do so promotes a false confidence in an erroneous topology, as driven by model misspecification (Philippe et al. 2011). The overwhelmingly parametric nature of modern phylogenetics insures that imperative issues will revolve around the processes being modeled, and what they actually allow us to ask from our data (Sullivan and Joyce 2005). However, the selection of methods that model processes of interest requires an *a priori* hypothesis that delimits which processes are involved. Yet, diagnosing prominent processes is difficult in that a phylogenetic context is required from which to build hypotheses. Fortunately, a wealth of information can be parsed from otherwise "non-phylogenetic" signal (*sensu* Philippe et al. 2005). For example, many statistical tests diagnose hybridization via its characteristic signature on the distribution of discordant topologies (e.g. Pease and Hahn 2015). Theoretical predictions regarding AGTs and the parameters under which they are generated are also well characterized (Degnan and Salter 2005; Degnan and Rosenberg 2009). Thus, by applying appropriate analytical approaches that sample many independently segregating regions of the genome, empiricists can still derive biologically meaningful phylogenies, despite the presence of complicated species-histories (McCormack et al. 2009; Kumar et al. 2012).

Here, we demonstrate an empirical approach that infers species-histories and sources of subtree discordance when conflict originates not only from anomaly zone divergences but also hybridization. To do so, we used SNP-based coalescent and polymorphism-aware phylogenetic methods (Chifman and Kubatko 2014; Leache et al. 2014; De Maio et al. 2015) that bypass the necessity of fully-resolved gene trees. We combine coalescent predictions, phylogenetic network

inference (Solís-Lemus and Ané 2016), and novel coalescent phylogeographic methods (Oaks 2018) to diagnose the sources of phylogenetic discordance and, by so doing, resolve a rather conflicted complex of study-species. We then contextualize our results to demonstrate the downstream implications of 'problematic' tree-space for threatened and endangered taxa, as represented by our study complex.

Gila

Few freshwater taxa have proven as problematic in recent years as the *Gila robusta* complex (Cyprinoidea: Leuciscidae) endemic to the Gila River basin of southwestern North America (Fig. 1). The taxonomic debate surrounding this complex exemplifies an inherent conflict between the traditional rigidity of systematic taxonomy versus the urgency of decision-making for conservation and management (Forest et al. 2015). Our study system is the Gila River, a primary tributary of the lower basin Colorado River that drains the majority of Arizona and ~11% of New Mexico. The critical shortage of water in this region (Sabo et al. 2010) is a major geopolitical driver for the taxonomic conflict surrounding our study species. As an example, the lower Colorado basin is responsible for approximately half of the total municipal and agricultural water requirements of the state of Arizona, and nearly two-thirds of its total gross state product (GSP) (Bureau of Reclamation 2012; James et al. 2014). This disproportionate regional reliance creates tension between the governance of a resource and its usage (e.g. Huckleberry and Potts 2019) which in turn magnifies the stakes involved in conservation policy (Minckley 1979; Carlson and Muth 1989; Minckley et al. 2006).

We focused on three species (Roundtail chub, *Gila robusta*; Gila chub, *G. intermedia*; and Headwater chub, *G. nigra*) that comprise a substantial proportion of the endemic

ichthyfaunal diversity of the Gila Basin [=20% of 15 extant native species (excluding extirpated *Gila elegans* and *Xyrauchen texanus*); Minckley and Marsh 2009]. Historically, the focal taxa have been subjected to numerous taxonomic rearrangements (Fig. 1). Until recently, the consensus was defined by Minckley and DeMarais (2000) on the basis of morphometric and meristic characters. These have since proven of limited diagnostic capacity in the field, thus provoking numerous attempts to re-define morphological delimitations (Brandenburg et al. 2015; Moran et al. 2017; Carter et al. 2018). Genetic approaches have to date been unproductive (Schwemm 2006; Copus et al. 2018), leading to a recent taxonomic recommendation that subsequently collapsed the complex into a single polytypic species (Page et al. 2016, 2017). For the sake of clarity, we employ herein the nomenclature of Minckley and DeMarais (2000).

Methods (1514)

Taxonomic sampling

A representative panel of *N*=386 individuals (Table S1; Fig. 2) was chosen from existing collections (Douglas et al. 2001; Douglas and Douglas 2007), to include broad geographic sampling of the complex as well as congeners. For the sake of clarity, we retained species-level nomenclature for all members of the *Gila robusta* complex, and additionally discriminate between *G. robusta* from the upper and lower basins of the Colorado River ecosystem (Chafin et al. 2019).

One species, *Gila elegans*, is extirpated from the wild, thus samples were provided by the Southwestern Native Aquatic Resources and Recovery Center (Dexter, NM). The genus *Ptychocheilus* served to root the *Gila* clade within the broader context of western leuciscids (Schönhuth et al. 2012, 2014, 2018).

Reduced-representation sequencing

Genomic DNA was extracted using either PureGene® or DNeasy® kits (Qiagen Inc.) and quantified via fluorometer (Qubit™; Thermo-Fisher Scientific). Library preparations followed the published ddRAD protocol (Peterson et al. 2012). Restriction enzyme and size-selection ranges were first screened using an *in silico* procedure (Chafin et al. 2018), with the target fragment sizes further optimized by quantifying digests for 15 representative samples on an Agilent 2200 TapeStation. Final library preparations were double-digested using a high-fidelity *PstI* (5'-CTGCAG-3') and *MspI* (5'-CCGG-3') following manufacturer's protocols (New England Biosciences). Digests were purified using bead purification (Ampure XP; Beckman-Coulter Inc.), and standardized at 100 ng per sample. Samples were then ligated with customized adapters containing unique in-line barcodes, pooled in sets of 48, and size-selected at 250-350bp (not including adapter length), using a Pippin Prep automated gel extraction instrument (Sage Sciences). Adapters were then extended in a 12-cycle PCR using Phusion high-fidelity DNA polymerase (New England Biosciences Inc.), completing adapters for Illumina sequencing and adding an i7 index. Libraries were pooled to N=96 samples per lane (i.e., 2 sets of 48) for 100bp single-end sequencing on an Illumina HiSeq 2500 at the University of Wisconsin Biotechnology Center (Madison, WI).

Data processing and assembly

Raw Illumina reads were demultiplexed and filtered using the PYRAD pipeline (Eaton 2014). We removed reads containing >1 mismatch in the barcode sequence, or >5 low-quality base-calls (Phred Q<20). Homologs assembly was then performed using *de novo* clustering in

VSEARCH (Rognes et al. 2016) using an 80% mismatch threshold. Loci were excluded according to following criteria: >5 ambiguous nucleotides; >10 heterozygous sites in the alignment; >2 haplotypes per individual; <20X and >500X sequencing depth per individual; >70% heterozygosity per-site among individuals.

Our ddRAD approach generated 22,768 loci containing a total of 173,719 variable sites, of which 21,717 were sampled (=1/ locus). Mean per-individual depth of coverage across all retained loci was 79X. All relevant scripts for post-assembly filtering and data conversion are available as open-source (github.com/tkchafin/scripts).

Phylogenetic inference

We formulated two simple hypotheses with regards to independent evolutionary sub-units. If populations represented a single polytypic species, then phylogenetic clustering should reflect intraspecific processes (e.g. structured according to stream heirarchy; Meffe and Vrijenhoek 1988). However, if *a priori* taxon assignments are evolutionarily independent, then they should be recapitulated in the phylogeny. Given well-known issues associated with application of supermatrix/ concatenation approaches (Degnan and Rosenberg 2006; Edwards et al. 2016) and pervasive gene-tree uncertainty associated with short loci (Leaché and Oaks 2017), we also employed SNP-based methods that bypassed the derivation of gene trees (Leaché and Oaks 2017).

We first explored population trees in SVDQUARTETS (Chifman and Kubatko 2014, 2015; as implemented in PAUP*, Swofford 2002) across 12 variably filtered datasets using four differing occupancy thresholds per SNP locus (i.e., 10, 25, 50, and 75%), along with three differing thresholds per individual (10, 25, and 50%). These filtered datasets ranged from 7357–

21007 SNPs, with 8.48–43.65% missing data and 256–347 individuals. SVDQUARTETS eases computation by inferring coalescent trees from randomly sampled quartets of species (i.e. optimizing among 3 possible unrooted topologies). It then generates a population tree with conflicts among quartet trees minimized via implementation of a quartet-assembly algorithm (Snir and Rao 2012). Given run-time constraints (the longest was 180 days on 44 cores), all runs sampled $\binom{N_{tips}}{4}/2$ quartets and were evaluated across 100 bootstrap pseudo-replicates.

We also used a polymorphism-aware method (POMO; Schrempf et al. 2016) in IQ-TREE (Nguyen et al. 2014). POMO considers allele frequencies rather than single nucleotides, thus allowing evaluation of change due to both substitution and drift. To provide POMO with empirical estimates of polymorphism, we used the entire alignment, to include non-variable sequences. We filtered liberally using individual occupancy thresholds of 10% per-locus so as to maximize individual retention and per-population sample sizes. We then deleted populations that contained <2 individuals, and loci with >=90% missing data per-population. This yielded a dataset of 281,613 nucleotides and 40 tips. Non-focal outgroups were excluded due to their disproportionate effect on missing data.

We also calculated concordance factors (CFs) using a Bayesian concordance analysis in BUCKY (Larget et al. 2010), parallelized across all quartets via an adaptation of the TICR pipeline (Stenz et al. 2015). To prepare these data, we sampled all non-monomorphic full gene alignments for which at least 1 diploid genotype could be sampled per population. We excluded outgroups and non-focal *Gila* so as to maximize number of loci retained. This yielded 3,449 genes across 31 sampled tips. Gene-tree priors were generated using MRBAYES v.3.2.6 (Ronquist et al. 2012) with 4 independent chains, each of which was sampled every 10,000 iterations, with a total chain length of 100,000,000 iterations and 50% discarded as burn-in. BUCKY was then

run in parallel to generate quartet CFs across 31,465 quartets, using a chain length of 10,000,000, again with 50% burn-in. Quartet topologies were used to generate a population tree using QUARTETMAXCUT (Snir and Rao 2012), using the *get-pop-tree.pl* script from TICR (Stenz et al. 2015; https://github.com/nstenz/TICR).

Comparing phylogenies and estimating site-wise conflict

To evaluate the performance of SVDQUARTETS, TICR, and PoMo, we first computed site-wise log-likelihood scores (*SLS*) for each topology by performing a constrained ML search in IQ-TREE. For comparison, we also generated an unconstrained concatenated tree. All ML analyses employed a GTR model with empirical base frequencies and gamma-distributed rates, and were assessed across 1,000 bootstrap pseudoreplicates. Analyses were also reduced to a subset of tips common across all variably filtered datasets. We quantified the phylogenetic signal supporting each resolution as the difference in site-wise log-likelihood scores (ΔSLS) between each population tree and the concatenation tree (Shen et al. 2017). We then calculated site-wise concordance factors (s*CF*) as an additional support metric (Minh et al. 2018).

Tests of hybridization and deep-time reticulation

D-statistics (Green et al. 2010; Eaton and Ree 2013) were calculated using COMP-D (Mussmann et al. 2019). To further test hypotheses of reticulation, we used quartet CFs as input for phylogenetic network inference using the SNAQ algorithm in PHYLONETWORKS (Solís-Lemus and Ané 2016; Solís-Lemus et al. 2017). The network was estimated under models of 0–5 hybrid nodes (*h*). Models were evaluated using 48 independent replicates, with the best-fit model

being that which maximized change in pseudolikelihood. Given the computational constraints of network inference, we reduced the dataset to N=2 populations per focal species (=12 total tips).

Anomaly zone detection

Coalescent theory characterizes the boundaries of the anomaly zone in terms of branch lengths in coalescent units (Degnan and Rosenberg 2006). To test if contentious relationships in our tree fell within the anomaly zone, we first transformed branch lengths using quartet CFs (Stenz et al. 2015, equation 1), then tested if internode branch lengths fell within the theoretical boundary for the anomaly zone (Linkem et al. 2016, equation 1). Code for these calculations are modified from Linkem et al. (2016) and are available as open-source (github.com/tkchafin/anomaly_zone).

Tests of co-divergence

The contemporary course of the Colorado River resulted from the Pliocene erosion of the Grand Canyon (McKee et al. 1967; Minckley 1986). *Gila* in the lower Colorado River basin then differentiated following one or more colonization events (e.g. Rinne 1976). Subsequent work (Douglas et al. 1999) supported this conclusion by examining contemporary phenotypic variation among all three species as a function of historical drainage connectivity, with the conclusion that body shape was most readily explained by Pliocene hydrography.

We tested if divergences were best explained by a model of *in situ* diversification following a single colonization event, or instead by multiple, successive colonizations. We compared divergence models using a Bayesian approach (program ECOEVOLITY, Oaks 2018) that used a coalescent model (Bryant et al. 2012) to update a prior expectation for the number of

evolutionary events across independent comparisons. Four independent MCMC chains were run with recommended settings and a burn-in that maximized effective sample sizes. Event models followed a Dirichlet process, with the concentration parameter exploring four alternative gamma distributed priors (i.e. α=2.0, β=5.70; α=0.5, β=8.7; α=1.0, β=0.45; and α=2.0, β=2.18).

We randomly sampled 2000 full-locus alignments, then examined potential co-divergences in the lower-basin complex by selecting a series of pairwise comparisons: *Gila elegans* x *G. robusta* (lower); *G. seminuda* x *G. robusta* (lower); *G. jordani* x *G. robusta* (lower); *G. intermedia* x *G. robusta* (lower); and *G. intermedia* x *G. nigra* (lower). These targeted nodes represent H, G, F, I, and N in the SVDQUARTETS topology (Fig. 3A).

Results (1147)

Phylogenetic conflict in Gila

We found tree reconstructions across all three population methods relatively congruent (SVDQUARTETS = Fig. 3; TICR, and POMO = Fig. 4). The concatenated supermatrix tree (Fig. S1) was also largely congruent with the population trees, but with two major disparities (discussed below). Bootstrap support was variable and declined with decreasing nodel depth in the SVDQUARTETS analysis (Fig. 3), whereas the vast majority of nodes in POMO were supported at 100% (Fig. 4A).

All analyses consistently supported the monophyly of a clade consisting of *G. intermedia*, *G. nigra*, and lower basin *G. robusta* (hereafter the 'lower basin complex'). This clade had high bootstrap support in both SVDQUARTETS and POMO, and was universally placed as sister to *G. jordani*. *Gila robusta* was unequivocally polyphyletic in all analyses, forming two distinct clades separated by the Grand Canyon. The lower basin clade was monophyletic in all cases, save the

concatenated tree, where it was paraphyletic (Fig. S1). It was also consistently recovered as sister to a monophyletic *G. nigra* + *G. intermedia*, with the exclusion of a single sample site (Aravaipa Creek) that nested within *G. intermedia* in the PoMo tree. Of note, this population had been previously diagnosed as trending towards *G. intermedia* in terms of morphology (Rinne 1976; Demarais 1986), although hybridization was not supported by *D*-statistics (Table 1).

Topology within the *G. nigra* + *G. intermedia* clade was less consistent. Both were reciprocally monophyletic in the SVDquartets tree (albeit with low support; Fig. 3), whereas PoMo yielded a monophyletic *G. intermedia*, with but one population (Spring Creek) contained within *G. nigra* (Fig. 4A). The PoMo tree also conflicted with the other methods in its paraphyletic placement of upper basin *G. robusta*. We suspect this represents an artefact of well-known hybridization with sympatric *G. cypha* (Dowling and DeMarais 1993; Gerber et al. 2001; Douglas and Douglas 2007; Chafin et al. 2019).

Discriminating among sources of phylogenetic conflict

Phylogenetic conflict was variably attributable to either hybridization or rapid divergence. We found support for a single reticulation event connecting *G. seminuda* and *G. elegans*, an hypothesis consistent with prior interpretations (DeMarais et al. 1992). This particular model (i.e., h=1) was selected as that which maximized both the first [$L'(h) = L(h) - L(h\text{-}1)$] and second order [$L''(h) = L'(h+1) - L'(h)$] rate of change in pseudolikelihood (Fig. S2; following Evanno et al. 2005). Of note, introgression between *G. elegans* and *G. seminuda* was supported by elevated values of h, and by *D*-statistics ($\bar{D}$ = 0.302 across 86,400 tests; Table 1). Introgression between upper basin *G. robusta* and *G. cypha* was also supported ($\bar{D}$ = -0.236 across 45,056 tests). No other introgressions were noted, thus rejecting the hypothesized hybrid

origins for both *G. jordani* (Dowling and DeMarais 1993; Dowling and Secor 1997) and *G. nigra* (Demarais 1986; Minckley and DeMarais 2000).

Multiple internode pairs were observed in the anomaly zone (Fig. 5). In all cases, internode branches separating *G. nigra* and *G. intermedia*, and those separating their constituent lineages, reflected coalescent lengths that would yield anomalous gene trees. Not surprisingly, the internode separating *G. jordani* from the lower basin complex, and that of *G. robusta* from *G. intermedia*/ *G. nigra* (Fig. 5C; tan branches) also fell within the anomaly zone, per TICR and concatenated topology results.

Relative performance of species-tree methods

Change in site-likelihoods among constrained and unconstrained IQ-TREE searches in all cases suggested that our recovered species-trees were supported by a minority of sites (Fig. S3), an observation consistent with tree regions being in the anomaly zone. Several discrepancies also reflected idiosyncrasies among the different approaches. For example, the POMO topology has a paraphyletic upper basin *G. robusta* within which *G. elegans*, *G. cypha*, *G. seminuda*, *G. jordani*, and the lower basin complex were subsumed (Fig. 4A). However, only ~10% of SNPs supported this resolution (Fig. S4), a value far below the theoretical minimum s*CF* derived from completely random data (Minh et al. 2018). Of note, hybridization is a well-known artefact when a bifurcating tree is inferred from reticulated species (Sosef 1997; Schmidt-Lebuhn 2012), with concatenation or binning approaches using genomic data being demonstrably vulnerable (Bangs et al. 2018). Thus, we tentatively attribute the observed paraphyly as an artefact of documented hybridization between *G. cypha* and *G. robusta* (Chafin et al. 2019), and the inability of POMO to

model hybridization. Hybridization also potentially drives the lack of monophyly in *G. seminuda*, per TICR and the concatenation tree (Fig. S1).

We also explored the impact of matrix occupancy filters on SVDQUARTETS, and bootstrap support and overall topological consistency declining with increasingly stringent filters (Fig. 3b). This corroborates prior evaluations with regard to the impacts of over-filtering RADseq data (Eaton et al. 2017). In all cases, site-wide concordance was significantly predicted by subtending branch lengths, but not by node depths (Fig. S5). This suggests that site-wise concordance was unbiased in our analyses at either shallower or deeper timescales, but was affected instead by the extent of time separating divergences. Some bioinformatic biases such as ortholog misidentification or lineage-specific locus dropout will disproportionally affect deeper nodes (Eaton 2017). However, we interpret the lack of correlation between node depth and site-wise concordance as an indication that these processes lack substantial bias.

Biogeographic hypotheses and co-divergence

ECOEVOLITY used four independent MCMC chains that sampled posterior parameter space over 300,000 iterations. Burn-in was determined as the number of samples removed to optimize effective sample sizes (ESS>500 in all cases following 10-30% burn-in). Model selection was not found to be vulnerable to alternative event priors (Fig. S6). The best-fitting model across all priors consistently demonstrated co-divergence of *G. jordani* with the lower basin complex (*G. robusta* x *G. intermedia* and *G. intermedia* x *G. nigra*; Fig. 6). The divergence of *G. elegans* and *G. seminuda* from a theoretical lower basin ancestor pre-dates this putatively rapid radiation, although it is unclear if these estimates were impacted by the aforementioned introgression between *G. seminuda* and *G. elegans*.

Posterior effective population size (N_e) estimates were large (e.g. >20,000) and consistent with previous estimates (Garrigan et al. 2002). *Gila jordani* was an exception, with a mean posterior N_e=6,062. We find this discrepancy rather unsurprising, given the extremely narrow endemism of this species (Tuttle and Scoppettone 1990), and its recent bottleneck (Hardy 1982), although this is still a rather large estimate given the latter. Posterior divergence time estimates suggested a late-Miocene/ early-Pliocene origin of *G. elegans*. Results for *G. seminuda* and the lower basin radiation indicated Pliocene and early Pleistocene divergences, respectively. These results are supported in the fossil record (Uyeno 1960; Uyeno and Miller 1963), although we note paleontological evaluations of *Gila* have been sparse. Thus, we hesitate to interpret these as absolute dates, given our fixed mutation rate for these analyses and an uncertainty regarding the capacity of RADseq methods to yield an unbiased sampling of genome-wide mutation rate variation (e.g. Cariou et al. 2016).

Discussion (1962 + 313 = 2275)

The goal of our study was to determine if extensive geographic and genomic sampling could resolve the taxonomically recalcitrant *G. robusta* complex. We applied diverse phylogenetic models and tests of hybridization and predictions of parameter space within the anomaly zone to diagnose sources of subtree discordance. In so doing, we also tested multiple hypothesized hybrid speciation events. We detected a single reticulation (*G. seminuda*), although other events with a lower component of genomic introgression may have also occurred. We documented rapid co-divergence of lower basin taxa within the anomaly zone and were able to resolve these despite the prevalence of incomplete lineage sorting. This scenario (as outlined

below) is consistent with the geomorphology of the region and seemingly represents an adaptive radiation by our study complex, as facilitated by drainage evolution.

Methodological artefacts and conflicting phylogenetic hypotheses for Gila

Increased geographic and genomic sampling revealed the presence of diagnosable lineages within the *G. robusta* complex, with both rapid and reticulate divergences influencing inter-locus conflict. Phylogenetic hypotheses for our focal group had previously been generated using allozymes (Dowling and DeMarais 1993), Sanger sequencing (Schwemm 2006; Schönhuth et al. 2014), microsatellites (Dowling et al. 2015), and more recently RADseq (Copus et al. 2018). None could resolve relationships within the lower basin complex. To explain these contrasts, we argue that prior studies suffered from systematic artefacts and ascertainment biases that were overcome, at least in part, by our approach.

Incomplete or biased sampling is a familiar problem for biologists (e.g. Hillis 1998; Schwartz and McKelvey 2009; Ahrens et al. 2016), and we suggest it served as a major stumbling block for delineating the evolutionary history of *Gila*. Although insufficient sampling is common in studies of threatened and endangered species, its repercussions are severe with regard to phylogenetic inference (Hillis 1998). This fact is substantiated by the many examples in which increasingly comprehensive geographic sampling spurred a revision of phylogenetic hypotheses (e.g. Oakey et al. 2004; Linck et al. 2019). Likewise, incomplete sampling of genome-wide topological variation (e.g. Maddison 1997; Degnan and Rosenberg 2009) is an additional source of bias, especially when a very small number of markers are sampled. These issues alone may explain the variation among prior studies. For example, Schwemm (2006) sampled extensively, including nearly all of the sites included in this study, but was only able to

examine a handful of genes. Because anomalous gene trees are most probable under a scenario of rapid radiation (as documented herein), the reduced number of loci used by Schwemm (2006) could not recover a consistent species tree. Copus et al. (2016, 2018) examined a dataset containing 6,658 genomic SNP loci (across 1,292 RAD contigs), but only did so across a sparse sample of 19 individuals. A bioinformatic acquisition bias also likely impacted this study, in the form of strict filtering that disproportionately excluded loci with higher mutation rates (Huang and Knowles 2016).

A necessary consideration when validating phylogenetic hypotheses across methods (and datasets) is to gauge compatibility between the underlying evolutionary processes and those actually being modeled. In this sense, the consideration of statistical support metrics alone can be not only misleading, but also promote false conclusions. For example, bootstrapping is by far the most prevalent method of evaluating support in phylogenetic datasets (Felsenstein 1985). While bootstrap concordances may be appropriate for moderately-sized sequence alignments (e.g. Efron et al. 1996), they can be meaningless when applied to sufficiently large datasets (Gadagkar et al. 2005; Kumar et al. 2012). This is apparent in the high bootstrap support displayed for anomalous relationships in our own analysis (Fig. S1). Phylogenetic signal also varies among loci, such that in many instances, relatively few loci drive contentious relationships (Shen et al. 2017). Likewise, not all methods are equal with respect to their simplifying assumptions. Given this, we feel it imperative to consider the biases and imperfections in both our data, and the models we apply.

Complex evolution and biogeography of the Colorado River

The taxonomic instability in *Gila* is not uncommon for fishes of western North America, where confusing patterns of diversity were generated by tectonism and vulcanism (Minckley et al. 1986; Spencer et al. 2008). This issue is particularly dramatized when viewed through the lens of modern drainage connections (Douglas et al. 1999). Historic patterns of drainage isolation and intermittent fluvial connectivity not only support our genomic conclusions but also summarize the paleohistory of the Colorado River over temporal and spatial scales.

The earliest record of fossil *Gila* from the ancestral Colorado River is mid-Miocene (Uyeno and Miller 1963), with subsequent Pliocene fossils representing typical 'big river' morphologies now associated with *G. elegans*, *G. cypha*, and *G. robusta* (Uyeno and Miller 1965). The modern Grand Canyon region lacked any fluvial connection at the Miocene-Pliocene transition, due largely to regional tectonic uplifts that subsequently diverted the Colorado River (Spencer et al. 2001; House et al. 2005). Flows initiated in early Pliocene (c.a. 4.9 mya; Sarna-Wojcicki et al. 2011), and subsequently formed a chain of downstream lakes associated with the Bouse Formation (Lucchitta 1972; Spencer and Patchett 2002). Evidence suggests 'spillover' by a successive string of Bouse Basin paleolakes was episodic, and culminated in mid-Pliocene (House et al. 2008), with an eventual marine connection via the Salton Trough to the Gulf of California (Dorsey et al. 2007). Prior to this, the Gila River also drained into the Gulf (Eberly and Stanley 1978), and sedimentary evidence indicated that it was isolated from the Colorado until at least mid-Pliocene by a northward extension of the Gulf (Helenes and Carreno 2014). This geomorphology is reflected in a broader phylogeographic pattern that underscores marked differences between resident fish communities in the upper and lower basins (Hubbs and Miller 1948).

Intra-basin diversification also occurred as an addendum to hydrologic evolution. Although the course of the pluvial White River is now generally dry, it may have been a Pliocene-early Pleistocene tributary of a paleolake system when the proto-Colorado first extended into the modern-day lower basin (Dickinson 2013). This may represent an initial colonization opportunity for upper basin fishes, an hypothesis that coincidentally aligns well with our rudimentary age estimate for Virgin River chub, *G. seminuda* (Fig. 6). This early isolation, as well as the continued contrast between the spring-fed habitats therein, and the high flows of the ancestral Colorado River, provide an explanation for the unique assemblage of *Gila* and other fishes therein (Hubbs and Miller 1948).

Phylogenetic signatures of the anomaly zone (Fig. 5) coupled with co-divergence modeling (Fig. 6) suggest the diversification of lower basin *Gila* occurred rapidly post-colonization. Late Pliocene integration of the two basins provided an opportunity for dispersal into the lower basin tributaries. The Plio-Pleistocene climate of the region was quite different, with a relatively mesic Pliocene as precursor to a protracted monsoonal period extending through early Pleistocene (Thompson 1991; Smith et al. 1993). The latter, in turn, may have resulted in relatively unstable drainage connections (Huckleberry 1996). The potential for climate-driven instability, and the complex history of intra-drainage integration of Gila tributaries during the Plio-Pleistocene (Dickinson 2015), lends support to the ‘cyclical-vicariance’ model proposed by Douglas et al. (1999). These periods of isolation may have promoted an accumulation of ecological divergences that persisted post-contact, and were sufficient to maintain species boundaries despite contemporary sympatric distributions and weak morphological differentiation. This hypothesis is also supported by the non-random mating found among *G. robusta* and *G. nigra*, despite anthropogenically-induced contact (Marsh et al. 2017).

Management implications

A request by the Arizona Game and Fish Department to review the taxonomy of the *Gila robusta* complex prompted the American Fisheries Society (AFS) and the American Society of Ichthyology and Herpetology (ASIH) to recommend the synonymization of *G. intermedia* and *G. nigra* with *G. robusta*, owing in part to their morphological ambiguity and an imprecise taxonomic key (Carter et al. 2018). Given this, a proposal to extend protection to lower basin *G. robusta* and *G. nigra* at the federal level was subsequently withdrawn (USFWS 2017; Fig. 1). Prior to this, *G. intermedia* alone was classified as endangered (USFWS 2005) under the Endangered Species Act (ESA 1973; 16 U.S.C. § 1531 et seq).

This study provides a much needed resolution to this debate by defining several aspects: First, our study reinforced the recognition of *G. robusta* as demonstrably polyphyletic, with two discrete, allopatric clades corresponding to the upper and lower basins of the Colorado River (Dowling and DeMarais 1993; Schönhuth et al. 2014). These data, together with the geomorphic history of the region that promoted endemic fish diversification (as above), clearly reject '*G. robusta*' as a descriptor of contemporary diversity. This underscores a major discrepancy in the taxonomic recommendations for the lower basin complex (Page et al. 2016). Given that the type locality of *G. robusta* is in the upper basin (i.e., the Little Colorado River), we note a pressing need either to determine taxonomic precedence for the lower basin '*G. robusta*,' or to provide a novel designation. The potential resurrection of a synonym is a possibility, necessitating a detailed examinations of the type specimens prior to a formal recommendation. This may be appropriately adjudicated by the AFS-ASIH Names of Fishes panel, as a follow-up to their earlier involvement.

The situation with *G. intermedia* and *G. nigra* is slightly more ambiguous. The short internodes and anomaly zone divergences identified herein explain previous patterns found in population-level studies, with elevated among-population divergence but scant signal uniting species (Dowling et al. 2015). We also unequivocally rejected the previous hypothesis of hybrid speciation for *G. nigra* (Minckley and DeMarais 2000; Dowling et al. 2015).

Rather, intermediacy in the body shape of *G. nigra* reflects differences accumulated during historic isolation (Douglas et al. 1999) and/ or the retention of an adaptive ecomorphology (Douglas and Matthews 1992). These hypotheses warrant further exploration, with provisional results employed in future management decisions (Forest et al. 2015). With regards to taxonomy, we confidently recommend that *G. intermedia* be resurrected, and that additional studies be implemented to dissect the potential distinctiveness of *G. nigra*. For management purposes, we echo a conservative, population-centric approach (previously argued for by Dowling et al. 2015; Marsh et al. 2017).

Three primary components of a 'Darwinian shortfall' in biodiversity conservation are recognized (Diniz-Filho et al. 2013): (i) The lack of comprehensive phylogenies; (ii) Uncertain branch lengths and divergence times; and (iii) insufficient models linking phylogenies with ecological and life-history traits. Taxonomic uncertainty in *Gila* is severely impacted by the first two of these, with taxonomic resolution prevented by the comingling of sparse phylogenetic coverage with temporal uncertainty. We must now address the relationships between ecology, life history, and phylogeny in *Gila*, so as to understand the manner by which phylogenetic groupings (identified herein) are appropriate as a surrogate for adaptive/ functional diversity. For example: To what degree are *Gila* in the lower basin ecologically non-exchangeable? How do

they vary in their respective life histories? Is reproductive segregation maintained in sympatry (as in Marsh et al. 2018), and if so, by what mechanism?

Conclusion (313)

The intractable phylogenetic relationships in *Gila* were resolved herein through improved spatial and genomic sampling. Our data, coupled with polymorphism-aware methods and contemporary approaches that infer trees, yielded a revised taxonomy for *Gila* in the lower Colorado basin. The geomorphic history of the Colorado River served to explain many anomalous patterns seen in this and previous studies, wherein opportunities for contact and colonization were driven by the tectonism that is characteristic of the region. The signal of rapid diversification is quite clear in our data, as interpreted from patterns inherent to phylogenetic discord. We emphasize that discordance in this sense does not necessarily represent measurement error or uncertainty, but rather an intrinsic component of phylogenetic variance that is not only expected within genomes (Maddison 1997), but also a necessary component from which to build hypotheses regarding the underlying evolutionary process (Hahn and Nakhleh 2016). Ignoring this variance in pursuit of a ‘resolved phylogeny’ can lead to incorrect inferences driven by systematic error. Similarly, insufficient spatial or genomic sampling may also promote a false confidence in anomalous relationships, particularly when character sampling is particularly dense, whereas taxon sampling is sparse.

We reiterate that phylogenetic hypotheses, by their very nature, cannot exhaustively capture the underlying evolutionary process. One approach is to categorize phylogenetic (and “non-phylogenetic”) signals in those regions of the tree that are refractive to certain models (as done herein). We acknowledge that attempting to reconstruct the past using contemporary

observations is a battle against uncertainty and bias, with the evolution of phylogenetic/ taxonomic hypotheses expected as additiona data are accrued. As such, we urge empiricists that engage in taxonomic controversies (such as this one) to interrogate their results for transparency. The task of sorting through conflicting recommendations invariably falls to managers, with unreported biases (be they methodological or geopolitical) only confounding those efforts.

Literature Cited

Ahrens D., Fujisawa T., Krammer H.J., Eberle J., Fabrizi S., Vogler A.P. 2016. Rarity and incomplete sampling in DNA-based species delimitation. Syst. Biol. 65:478–494.

Bangs M.R., Douglas M.R., Mussmann S.M., Douglas M.E. 2018. Unraveling historical introgression and resolving phylogenetic discord within *Catostomus* (Osteichthys□: Catostomidae). BMC Evol. Biol. 18:86.

Beaulieu J.M., O'Meara B.C. 2018. Can we build it? Yes we can, but should we use it? Assessing the quality and value of a very large phylogeny of campanulid angiosperms. Am. J. Bot. 105:417–432.

Brandenburg W.H., Kennedy J.L., Farrington M.A. 2015. Determining the historical distribution of the Gila robusta complex (Gila chub, *Gila intermedia*, Headwater Chub, *Gila nigra*, and Roundtail Chub, *Gila robusta*) in the Gila River Basin, New Mexico, using morphological analysis. Final Rep. to New Mex. Dep. Game Fish.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N.A., RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Bureau of Reclamation. 2012. Colorado River Basin water supply and demand study. Executive Summary.

Cariou M., Duret L., Charlat S. 2016. How and how much does RAD-seq bias genetic diversity estimates? BMC Evol. Biol. 16:1–8.

Carlson C.A., Muth R.T. 1989. The Colorado River: Lifeline of the American Southwest. Can. Spec. Publ. Fish. Aquat. Sci. 106:220–239.

Carter J.M., Clement M.J., Makinster A.S., Crowder C.D., Hickerson B.T. 2018. Classification success of species within the *Gila robusta* complex using morphometric and meristic characters—A re-examination. Copeia. 106:279–291.

Carter J.M., Crowder C.D., Leavitt D.J., S. M.A., Hickerson B. 2016. Classification success of

species within the *Gila robusta* complex using morphological and meristics characters—A re-examination. Rep. to Arizona Game Fish Dep.

Chafin T.K., Douglas M.R., Martin B.T., Douglas M.E. 2019. Hybridization drives genetic erosion in sympatric desert fishes of western North America. bioRxiv Prepr. 644641.

Chafin T.K., Martin B.T., Mussmann S.M., Douglas M.R., Douglas M.E. 2018. FRAGMATIC: in silico locus prediction and its utility in optimizing ddRADseq projects. Conserv. Genet. Resour. 10:325–328.

Chifman J., Kubatko L. 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics. 30:3317–3324.

Chifman J., Kubatko L. 2015. Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes. J. Theor. Biol. 374:35–47.

Copus J.M., Montgomery W.L., Forsman Z.H., Bowen B.W., Toonen R.J. 2018. Geopolitical species revisited: genomic and morphological data indicate that the roundtail chub *Gila robusta* species complex (Teleostei, Cyprinidae) is a single species. PeerJ. 6:e5605.

Darwin C. 1859. On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life. London: John Murray.

Degnan J.H., Rosenberg N.A. 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2:e68.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Degnan J.H., Salter L.A. 2005. Gene tree distributions under the coalescent process. Evolution. 59:24–37.

Demarais B.D. 1986. Morphological variation in *Gila* (Pisces: Cyprinidae) and geological history: lower Colorado River basin. Thesis.

DeMarais B.D., Dowling T.E., Douglas M.E., Minckley W.L., Marsh P.C. 1992. Origin of Gila seminuda (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. Proc. Natl. Acad. Sci. U. S. A. 89:2747–2751.

Dickinson W.R. 2013. Rejection of the lake spillover model for initial incision of the Grand Canyon, and discussion of alternatives. Geosphere. 9:1–20.

Dickinson W.R. 2015. Integration of the Gila River drainage system through the Basin and Range province of southern Arizona and southwestern New Mexico (USA). Geomorphology. 236:1–24.

Diniz-Filho J.A.F., Loyola R.D., Raia P., Mooers A.O., Bini L.M. 2013. Darwinian shortfalls in biodiversity conservation. Trends Ecol. Evol. 28:689–695.

Dorsey R.J., Fluette A., McDougall K., Housen B.A., Janecke S.U., Axen G.J., Shirvell C.R. 2007. Chronology of Miocene-Pliocene deposits at Split Mountain Gorge, Southern California: A record of regional tectonics and Colorado River evolution. Geology. 35:57–60.

Douglas M., Minckley W.L., DeMarais B.D. 1999. Did vicariance mold phenotypes of western North American fishes? Evidence from Gila River cyprinids. Evolution (N. Y). 53:238–246.

Douglas M.E., Matthews W.J. 1992. Does morphology predict ecology? Hypothesis testing within a freshwater stream fish assemblage. Oikos. 65:213.

Douglas M.R., Douglas M.E. 2007. Genetic structure of humpback chub *Gila cypha* and roundtail chub *G. robusta* in the Colorado River ecosystem. Rep. to Gd. Canyon Monit. Reserach Center, U.S. Geol. Surv.:99.

Dowling T.E., Anderson C.D., Marsh P.C., Rosenberg M.S. 2015. Population structure in the Roundtail chub (*Gila robusta* complex) of the Gila River Basin as determined by microsatellites: Evolutionary and conservation implications. PLoS One. 10:e0139832.

Dowling T.E., DeMarais B.D. 1993. Evolutionary significance of introgressive hybridization in cyprinid fishes. Nature. 362:444–446.

Dowling T.E., Secor C.L. 1997. The role of hybridization and introgression in the diversification of animals. Annu. Rev. Ecol. Syst. 28:593–619.

Eaton D.A.R. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics. 30:1844–1849.

Eaton D.A.R., Ree R.H. 2013. Inferring phylogeny and introgression using RADseq data: an example from glowering plants (Pedicularis: Orobanchaceae). Syst. Biol. 62:689–706.

Eaton D.A.R., Spriggs E.L., Park B., Donoghue M.J. 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. Syst. Biol. 66:399–412.

Eberly L.D., Stanley T.B. 1978. Cenozoic stratigraphy and geologic history of southwestern Arizona. Geol. Soc. Am. Bull. 89:921–940.

Edwards S. V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

Edwards S. V., Xi Z., Janke A., Faircloth B.C., McCormack J.E., Glenn T.C., Zhong B., Wu S., Lemmon E.M., Lemmon A.R., Leaché A.D., Liu L., Davis C.C. 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. Evol. 94:447–462.

Efron B., Halloran E., Holmes S. 1996. Bootstrap confidence levels for phylogenetic trees. Proc. Natl. Acad. Sci. 93:13429–13429.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14:2611–2620.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution. 39:783–791.

Forest F., Crandal K.A., Chase M.W., Faith D.P. 2015. Phylogeny, extinction and conservation: Embracing uncertainties in a time of urgency. Philos. Trans. R. Soc. B Biol. Sci. 370:1–8.

Gadagkar S.R., Rosenberg M.S., Kumar S. 2005. Inferring species phylogenies from multiple genes: Concatenated sequence tree versus consensus gene tree. J. Exp. Zool. Part B Mol. Dev. Evol. 304:64–74.

Garrigan D., Marsh P.C., Dowling T.E. 2002. Long-term effective population size of three endangered Colorado River fishes. Anim. Conserv. 5:95–102.

Gerber A.S., Tibbets C.A., Dowling T.E. 2001. The role of introgressive hybridization in the evolution of the *Gila robusta* complex (Teleostei: Cyprinidae). Evolution. 55:2028–2039.

Giarla T.C., Esselstyn J.A. 2015. The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. Syst. Biol. 64:727–740.

Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai W., Fritz M.H.-Y., Hansen N.F., Durand E.Y., Malaspinas A.-S., Jensen J.D., Marques-Bonet T., Alkan C., Prüfer K., Meyer M., Burbano H.A., Good J.M., Schultz R., Aximu-Petri A., Butthof A., Höber B., Höffner B., Siegemund M., Weihmann A., Nusbaum C., Lander E.S., Russ C., Novod N., Affourtit J., Egholm M., Verna C., Rudan P., Brajkovic D., Kucan Ž., Gušic I., Doronichev V.B., Golovanova L. V, Lalueza-Fox C., de la Rasilla M., Fortea J., Rosas A., Schmitz R.W., Johnson P.L.F., Eichler E.E., Falush D., Birney E., Mullikin J.C., Slatkin M., Nielsen R., Kelso J., Lachmann M., Reich D., Pääbo S. 2010. A draft sequence of the Neandertal genome. Science. 328:710–722.

Hahn M.W., Nakhleh L. 2016. Irrational exuberance for resolved species trees. Evolution. 70:7–17.

Hardy T.B. 1982. Ecological interactions ofthe introduced and native fishes in the outflow of Ash Springs, Lincoln County, Nevada. Thesis.

Heath T.A., Hedtke S.M., Hillis D.M. 2008. Taxon sampling and the accuracy of phylogenetic analyses. J. Syst. Evol. 46:239–257.

Helenes J., Carreno A.L. 2014. Neogene sedimentary record of the Gulf of California: towards a highly biodiverse scenario. In: Wehncke E. V., Lara-lAra J.R., Alvarez-Borrego S., Ezcurra E., editors. Conservation Science in Mexico’s Northwest: Ecosystem Status and Trends in the Gulf of Calfiornia. .

Hillis D.M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47:3–8.

House P.K., Pearthree P.A., Howard K.A., Bell J.W., Perkins M.E., Faulds J.E., Brock A.L. 2005. Birth of the lower Colorado River—Stratigraphic and geomorphic evidence for its inception near the conjunction of Nevada, Arizona, and California. GSA Field Guide 6: Interior Western United States. .

House P.K., Pearthree P.A., Perkins M.E. 2008. Stratigraphic evidence for the role of lake spillover in the inception of the lower Colorado River in southern Nevada and western Arizona. Spec. Pap. 439 Late Cenozoic Drain. Hist. Southwest. Gt. Basin Low. Color. River Reg. Geol. Biot. Perspect. 2439:335–353.

Huang H., Knowles L.L. 2016. Unforeseen consequences of excluding missing data from next-generation sequences: Simulation study of RAD sequences. Syst. Biol. 65:357–65.

Hubbs C.L., Miller R.R. 1948. The zoological evidence: correlation between fish distribution and hydrogeographic history in the desert basin of western North America. Bull. Univ. Utah. 10:17–166.

Huckleberry G. 1996. Historical geomorphology of the Gila River. Rep. to Arizona Geol. Surv.

Huckleberry J.K., Potts M.D. 2019. Constraints to implementing the food-energy-water nexus concept: Governance in the Lower Colorado River Basin. Environ. Sci. Policy. 92:289–298.

James T., Evans A., Madly E., Kelly C. 2014. The economic importance of the Colorado River to the Basin region. L William Seidman Res. Institute, W. P. Carey Sch. Business, Arizona State Univ.

Kumar S., Filipski A.J., Battistuzzi F.U., Kosakovsky Pond S.L., Tamura K. 2012. Statistics and truth in phylogenomics. Mol. Biol. Evol. 29:457–472.

Larget B.R., Kotha S.K., Dewey C.N., Ané C. 2010. BUCKy: Gene tree reconciliation with concordance. Bioinformatics. 26:2910–2911.

Leache A.D., Fujita M.K., Minin V.N., Bouckaert R.R. 2014. Species delimitation using genome-wide SNP data. Syst. Biol. 63:534–542.

Leaché A.D., Oaks J.R. 2017. The utility of single nucleotide polymorphism (SNP) data in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 48:69–84.

Lemmon E.M., Lemmon A.R. 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44:99–121.

Linck E., Epperly K., van Els P., Spellman G.M., Bryson Jr R.W., McCormack J.E., Canales-del-Castillo R., Klicka J. 2019. Dense geographic and genomic sampling reveals paraphyly and a cryptic lineage in a classic sibling species complex. Syst. Biol. Early Acce:syz027.

Linkem C.W.C., Minin V.N.V., Leache A., Leaché A.D. 2016. Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). Syst. Biol. 65:465–477.

Liu L., Edwards S. V. 2009. Phylogenetic analysis in the anomaly zone. Syst. Biol. 58:452–460.

Lucchitta I. 1972. Early history of the Colorado River in the Basin and Range province. Geol. Scoeity Am. Buttetin. 83:1933–1948.

Maddison W. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

De Maio N., Schrempf D., Kosiol C. 2015. PoMo: An allele frequency-based approach for species tree estimation. Syst. Biol. 64:1018–1031.

Marsh P.C., Clarkson R.W., Dowling T.E. 2017. Molecular genetics informs spatial segregation of two desert stream *Gila* species. Trans. Am. Fish. Soc. 146:47–59.

McCormack J.E., Huang H., Knowles L.L. 2009. Maximum likelihood estimates of species trees: How accuracy of phylogenetic inference depends upon the divergence history and sampling design. Syst. Biol. 58:501–508.

McKee E.D., Wislon R.F., Breed W.J., Breed C.S. 1967. Evolution of the Colorado River in Arizona- An hypothesis developed at the Symposium on Cenozoic Geology of the Colorado Plateau in Arizona, August 1964. Museum of Northern Arizona Bulletin.

McVay J.D., Hipp A.L., Manos P.S. 2017. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. Proc. R. Soc. B Biol. Sci. 284.

Meffe G.K., Vrijenhoek R.C. 1988. Conservation genetics in the management of desert fishes. Conserv. Biol. 2:157–169.

Mendes F.K., Hahn Y., Hahn M.W. 2016. Gene tree discordance can generate patterns of diminishing convergence over time. Mol. Biol. Evol. 33:3299–3307.

Minckley W.L. 1979. Aquatic habitats and fishes of the lower Colorado River, southwestern United States. Rep. to U.S. Bur. Reclam.

Minckley W.L., Henderson D.A., Bond C.E. 1986. Geography of western North American freshwater fishes: description and relationships to intracontinental tectonism. Zoogeography North Am. Freshw. Fishes.:519–613.

Minckley W.L., DeMarais B.D. 2000. Taxonomy of chubs (Telostei, Cyprinidae, genus *Gila*) in the American Southwest with comments on conservation. Copeia. 2000:251–256.

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

Minckley W.L., Marsh P.C., Deacon J.E., Dowling T.E., Hedrick P.W., Matthews W.J., Mueller G. 2006. A conservation plan for native fishes of the lower Colorado River. Bioscience. 53:219.

Minh B.Q., Hahn M., Lanfear R. 2018. New methods to calculate concordance factors for

phylogenomic datasets. bioRxiv.

Moran C.J., O'Neill M.W., Armbruster J.W., Gibb A.C. 2017. Can members of the south-western *Gila robusta* species complex be distinguished by morphological features? J. Fish Biol. 91:302–316.

Mussmann S.M., Douglas M.R., Bangs M.R., Douglas M.E. 2019. Comp-D: a program for comprehensive computation of D-statistics and population summaries of reticulated evolution. Conserv. Genet. Resour. 0:0.

Nguyen L.-T., Schmidt H.A., von Haeseler A., Minh B.Q. 2014. IQ-TREE: A fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. Mol. Biol. Evol. 32:268–274.

Oakey D.D., Douglas M.E., Douglas M.R. 2004. Small Fish in a Large Landscape: Diversification of *Rhinichthys osculus* (Cyprinidae) in Western North America. Copeia. 2004:207–221.

Oaks J.R. 2018. Full Bayesian comparative phylogeography from genomic data. Syst. Biol. 0:1–25.

Page L.M., Baldwin C.C., Espinosa-Pérez H., Findley L.T., Gilbert C.R., Hartel K.E., Lea R.N., Mandrak N.E., Schmitter-Soto J.J., Walker H.J. 2017. Taxonomy of *Gila* in the Lower Colorado River Basin of Arizona and New Mexico. Fisheries. 42:456–460.

Page L.M., Baldwin C.C., Espinosa-Pérez H., Gilbert C.R., Hartel K.E., Lea R.N., Mandrak N.E., Schmitter-Soto J.J., Walker H.J. 2016. Final report of the AFS/ASIH Joint Committee on the Names of Fishes on the taxonomy of *Gila* in the Lower Colorado River basin of Arizona and New Mexico. Rep. to Arizona Game Fish Dep.

Pante E., Puillandre N., Viricel A., Arnaud-Haond S., Aurelle D., Castelin M., Chenuil A., Destombe C., Forcioli D., Valero M., Viard F., Samadi S. 2015. Species are hypotheses: Avoid connectivity assessments based on pillars of sand. Mol. Ecol. 24:525–544.

Pease J.B., Hahn M.W. 2015. Detection and polarization of introgression in a five-taxon phylogeny. Syst. Biol. 64:651–662.

Pereira A.G., Schrago C.G. 2018. Incomplete lineage sorting impacts the inference of macroevolutionary regimes from molecular phylogenies when concatenation is employed: An analysis based on Cetacea. Ecol. Evol. 8:6965–6971.

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

Philippe H., Brinkmann H., Lavrov D. V., Littlewood D.T.J., Manuel M., Wörheide G., Baurain D. 2011. Resolving difficult phylogenetic questions: Why more sequences are not enough. PLoS Biol. 9:e1000602.

Philippe H., Delsuc F., Brinkmann H., Lartillot N. 2005. Phylogenomics. Annu. Rev. Ecol. Evol.

Syst. 36:541–562.

Pisani D., Pett W., Dohrmann M., Feuda R., Rota-Stabelli O., Philippe H., Lartillot N., Wörheide G. 2015. Genomic data do not support comb jellies as the sister group to all other animals. Proc. Natl. Acad. Sci. 112:15402–15407.

Rangel T.F., Colwell R.K., Graves G.R., Fučíková K., Rahbek C., Diniz-Filho J.A.F. 2015. Phylogenetic uncertainty revisited: Implications for ecological analyses. Evolution (N. Y). 69:1301–1312.

Rannala B., Yang Z. 2008. Phylogenetic inference using whole genomes. Annu. Rev. Genomics Hum. Genet. 9:217–231.

Reddy S., Kimball R.T., Pandey A., Hosner P.A., Braun M.J., Hackett S.J., Han K.L., Harshman J., Huddleston C.J., Kingston S., Marks B.D., Miglia K.J., Moore W.S., Sheldon F.H., Witt C.C., Yuri T., Braun E.L. 2017. Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. Syst. Biol. 66:857–879.

Rinne J.N. 1976. Cyprinid fishes of the genus *Gila* from the lower Colorado River basin. Wasmann J. Biol. 34(1):65-107.

Rognes T., Flouri T., Nichols B., Quince C., Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ Prepr. 4:e2409v1.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Hohna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–542.

Sabo J. L., Sinha T., Bowling L.C., Schoups G.H.W., Wallender W.W., Campana M.E., Cherkauer K.A., Fuller P. , Graf W.L., Hopmans J.W., Kominoski J.S., Taylor C., Trimble S.W., Webb R.H., Wohl E.E.. 2010. Reclaiming freshwater sustainability in the Cadillac Desert. Proc. Nat. Acad. Sci. 107(50):21256-21262.

Sarna-Wojcicki A.M., Deino A.L., Fleck R.J., McLaughlin R.J., Wagner D., Wan E., Wahl D., Hillhouse J.W., Perkins M. 2011. Age, composition, and areal distribution of the Pliocene Lawlor Tuff, and three younger Pliocene tuffs, California and Nevada. Geosphere. 7:599–628.

Schmidt-Lebuhn A.N. 2012. Fallacies and false premises-a critical assessment of the arguments for the recognition of paraphyletic taxa in botany. Cladistics. 28:174–187.

Schönhuth S., Perdices A., Lozano-Vilano L., García-de-León F.J., Espinosa H., Mayden R.L. 2014. Phylogenetic relationships of North American western chubs of the genus Gila (Cyprinidae, Teleostei), with emphasis on southern species. Mol. Phylogenet. Evol. 70:210–230.

Schönhuth S., Shiozawa D.K., Dowling T.E., Mayden R.L. 2012. Molecular systematics of western

North American cyprinids (Cypriniformes: Cyprinidae). Zootaxa. 303:281–303.

Schönhuth S., Vukić J., Šanda R., Yang L., Mayden R.L. 2018. Phylogenetic relationships and classification of the Holarctic family Leuciscidae (Cypriniformes: Cyprinoidei). Mol. Phylogenet. Evol. 127:781–799.

Schrempf D., Minh B.Q., De Maio N., von Haeseler A., Kosiol C. 2016. Reversible polymorphism-aware phylogenetic models and their application to tree inference. J. Theor. Biol. 407:362–370.

Schwartz M.K., McKelvey K.S. 2009. Why sampling scheme matters: The effect of sampling scheme on landscape genetic results. Conserv. Genet. 10:441–452.

Schwemm M.R. 2006. Genetic variation in the *Gila robusta* complex (Teleostei: Cyprinidae) in the lower Colorado River. Thesis.

Shen X.X., Hittinger C.T., Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. 1:1–10.

Simion P., Philippe H., Baurain D., Jager M., Richter D.J., Di Franco A., Roure B., Satoh N., Quéinnec É., Ereskovsky A., Lapébie P., Corre E., Delsuc F., King N., Wörheide G., Manuel M. 2017. A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. Curr. Biol. 27:958–967.

Simpson G.G. 1961. Principles of animal taxonomy. New York: Columbia University Press.

Smith G.A., Yang Wang, Cerling T.E., Geissman J.W. 1993. Comparison of a paleosol-carbonate isotope record to other records of Pliocene-early Pleistocene climate in the western United States. Geology. 21:691–694.

Snir S., Rao S. 2012. Quartet MaxCut: A fast algorithm for amalgamating quartet trees. Mol. Phylogenet. Evol. 61:1–8.

Solís-Lemus C., Ané C. 2016. Inferring phylogenetic networks with Maximum Pseudolikelihood under incomplete lineage sorting. PLoS Genet. 12:1–21.

Solís-Lemus C., Bastide P., Ané C. 2017. PhyloNetworks: A package for phylogenetic networks. Mol. Biol. Evol. 34:3292–3298.

Sosef M.S.M. 1997. Hierarchical models, reticulate evolution and the inevitability of paraphyletic supraspecific taxa. Taxon. 46:75–85.

Spencer J.E., Patchett P.J. 2002. Sr isotope evidence for a lacustrine origin for the upper Miocene to Pliocene Bouse Formation, lower Colorado River trough, and implications for timing of Colorado Plateau uplift. Geol. Soc. Am. Bull. 109:767–778.

Spencer J.E., Peters L., McIntosh W.C., Patchett P.J., Young R.A., Spamer E.E. 2001. 40Ar/39Ar geochronology of the Hualapai Limestone and Bouse Formation and implications for the

age of the lower Colorado River. Color. River Orig. Evol. Gd. Canyon, Arizona, Gd. Canyon Assoc.:89–91.

Spencer J.E., Smith G.R., Dowling T.E. 2008. Middle to late Cenozoic geology, hydrography, and fish evolution in the American Southwest. Geol. Soc. Am. 80301.

Springer M.S., Gatesy J. 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94:1–33.

Stadler T., Degnan J.H., Rosenberg N.A. 2016. Does gene tree discordance explain the mismatch between macroevolutionary models and empirical patterns of tree shape and branching times? Syst. Biol. 65:628–639.

Stenz N.W.M., Larget B., Baum D.A., Ané C. 2015. Exploring tree-like and non-tree-like patterns using genome sequences: An example using the inbreeding plant species *Arabidopsis thaliana (L.) heynh*. Syst. Biol. 64:809–823.

Sullivan J., Joyce P. 2005. Model selection in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 36:445–466.

Thompson R.S. 1991. Pliocene environments and climates in the western United States. Quat. Sci. Rev. 10:115–132.

Tuttle P., Scoppettone G. 1990. Status and life history of Pahranagat River fishes. Rep. to Nevada Dep. Wildl.

USFWS. 2005. Endangered and Threatened Wildlife and Plants: Listing Gila chub as endangered with critical habitat. Fed. Regist. 70:66664–66721.

USFWS. 2017. Endangered and Threatened Wildlife and Plants: Threatened species status for the Headwater chub and Roundtail chub distinct population segment. Fed. Regist. 82:16981–16988.

Uyeno T. 1960. Osteology and phylogeny of the American cyprinid fishes allied to the genus *Gila*. Thesis.

Uyeno T., Miller R.R. 1963. Summary of late Cenozoic freshwater fish records for North America. Occas. Pap. Museum Zool. Univ. Michigan. 631:1–34.

Uyeno T., Miller R.R. 1965. Middle Pliocene cyprinid fishes from the Bidahochi Formation. 1965(1):28–41.

Whelan N. V., Kocot K.M., Moroz L.L., Halanych K.M. 2015. Error, signal, and the placement of Ctenophora sister to all other animals. Proc. Natl. Acad. Sci. 112:5773–5778.

Wu M., Kostyun J.L., Hahn M.W., Moyle L.C. 2018. Dissecting the basis of novel trait evolution in a radiation with widespread phylogenetic discordance. Mol. Ecol. 27:3301–3316.

Figures and Tables

Table 1: Four-taxon *D*-statistic tests for admixture. Results are reported across *N* separate quartet samples per four-taxon test, randomly sampled without replacement. Significance is reported as the proportion of tests at $p<0.05$ (nSig/*N*) using chi-squared (χ^2), *Z*-test[1], and *Z*-test with Bonferroni correction[2]. Positive and negative values of *D* suggest introgression of the P3 lineage with either P2 or P1, respectively. Results in bold were also supported by the phylogenetic network.

P3	P2	P1	Mean *D*	*N*	nSig/*N* (χ^2)	nSig/*N* (Z^1)	nSig/*N* (Z^2)
cypha	*jordani*	lower basin	0.175	86400	0.033	0.072	0.001
cypha	*seminuda*	lower basin	0.099	86400	0.102	0.130	0.002
elegans	*jordani*	lower basin	-0.063	84800	0.029	0.050	0.000
elegans	*robusta* (lower)	*nigra/int.*	-0.026	413600	0.014	0.047	0.001
elegans	*robusta* (upper)	*cypha*	-0.236	45056	0.380	0.415	0.045
elegans	***seminuda***	lower basin	**0.302**	**86400**	**0.654**	**0.674**	**0.251**
jordani	*robusta* (lower)	*nigra/int.*	0.087	601600	0.042	0.072	0.001
jordani	*robusta* (lower)	*nigra/int.*	0.091	212800	0.041	0.067	0.005
nigra	*int.* (Salt)	*int.* (Verde)	0.086	126976	0.057	0.082	0.001
robusta (lower)	*intermedia*	*nigra*	0.041	793600	0.001	0.002	0.000
robusta (upper)	*jordani*	*robusta* (lower)	0.165	168000	0.050	0.081	0.001
robusta (upper)	*robusta* (lower)	*nigra/int.*	-0.009	601600	0.011	0.031	0.000
robusta (upper)	*seminuda*	lower basin	-0.017	180800	0.030	0.053	0.004
seminuda	*jordani*	lower basin	-0.204	81920	0.107	0.152	0.000
seminuda	*robusta* (lower)	*nigra/int.*	0.054	212800	0.011	0.031	0.001
atraria	*robusta* (upper)	*cypha*	0.082	57344	0.064	0.095	0.033
nigrescens	*robusta* (lower)	*nigra/int.*	-0.075	485472	0.023	0.079	0.002
nigrescens	*robusta* (upper)	*cypha*	-0.039	53248	0.040	0.066	0.005
pandora	*robusta* (lower)	*nigra/int.*	-0.123	225600	0.012	0.105	0.010
pandora	*robusta* (upper)	*cypha*	-0.047	24576	0.031	0.057	0.003

Figure 1: Timeline of the conservation status of *Gila* species endemic to the lower Colorado River basin [*See Copus et al (2018) for a detailed overview of taxonomic synonymies; †'The Center' refers to the Center for Biological Diversity (501c3), Tuscon, AZ; ‡'DPS' = Distinct Population Segment as referenced in the Endangered Species Act (ESA 1973; 16 U.S.C. § 1531 et seq), here referring specifically to a lower basin sub-unit of *G. robusta*]. Note that timeline is not to scale.

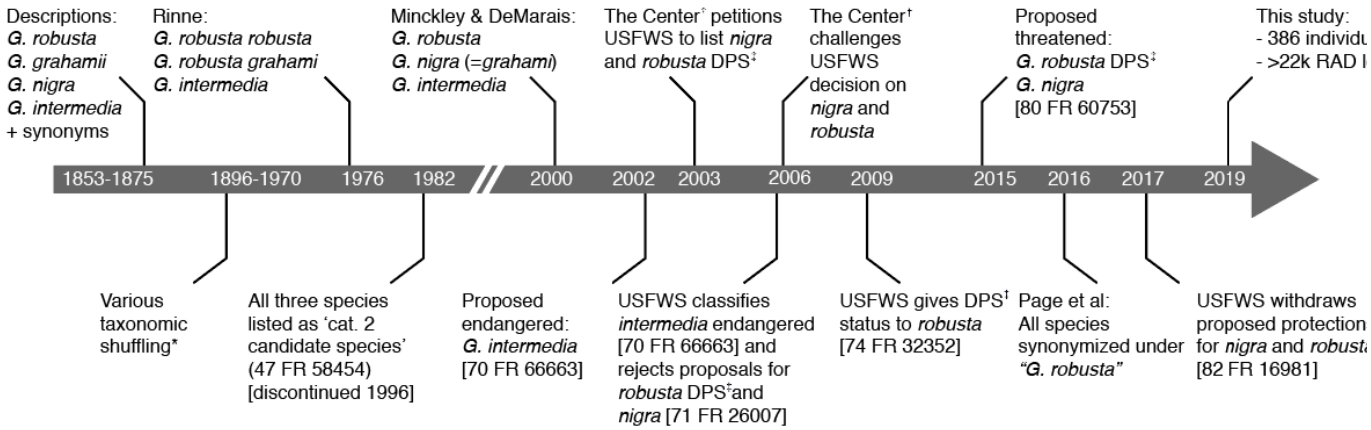

Figure 2: Sampling localities for *Gila* within the Colorado River Basin, southwestern North America. Locality codes are defined in Table S1. Sympatric locations (R14 and C2) are slightly offset for visibility purposes. Map insert increases the viewing scale for sampling sites within the lower basin 'complex' (Bill Williams and Gila Rivers).

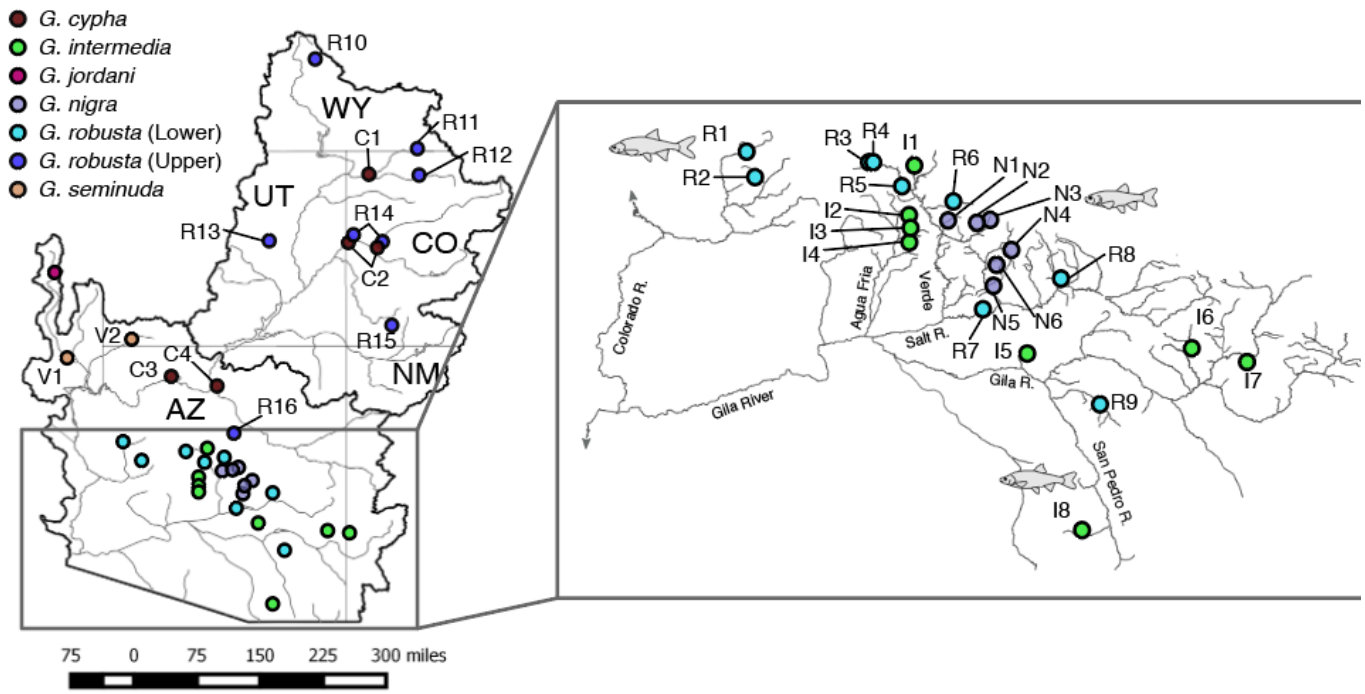

Figure 3: (A) Majority-rule consensus cladogram of SVDQUARTETS across 12 variably filtered datasets. (B) Binned bootstrap concordance values are reported for each dataset, coded by the matrix occupancy threshold per individual ("i") and per column ("c"; e.g. i50_c50 = 50% occupancy required per individual and per column). Dashed terminal branches indicate positions for taxa missing from >50% of datasets.

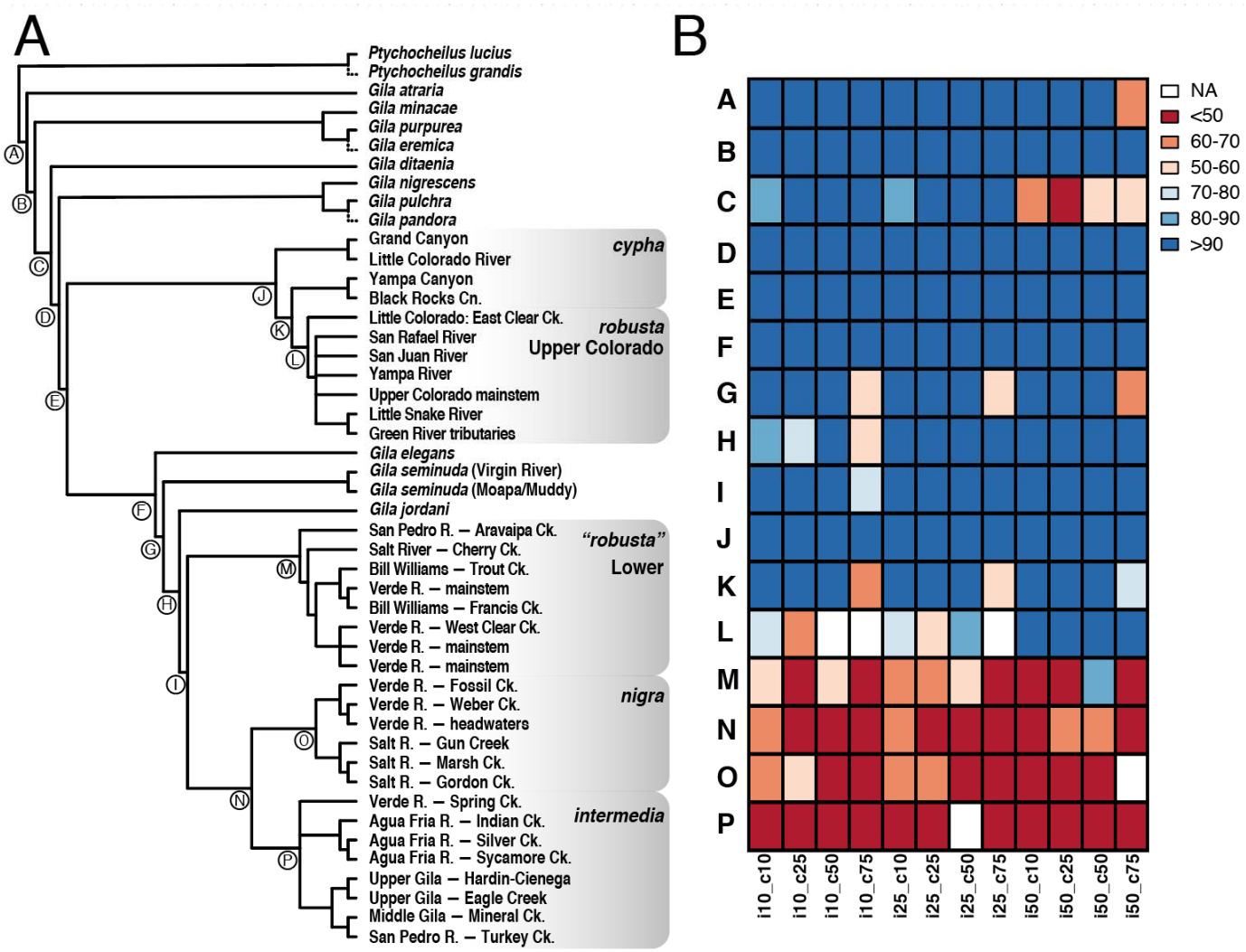

Figure 4: (A) PoMo phylogram with branch lengths as the number of substitutions *and* inferred number of drift events per site, with branch supports (as values <100%) representing concordance among 1000 bootstrap replicates; (B) TICR phylogram reporting branch lengths in coalescent units.

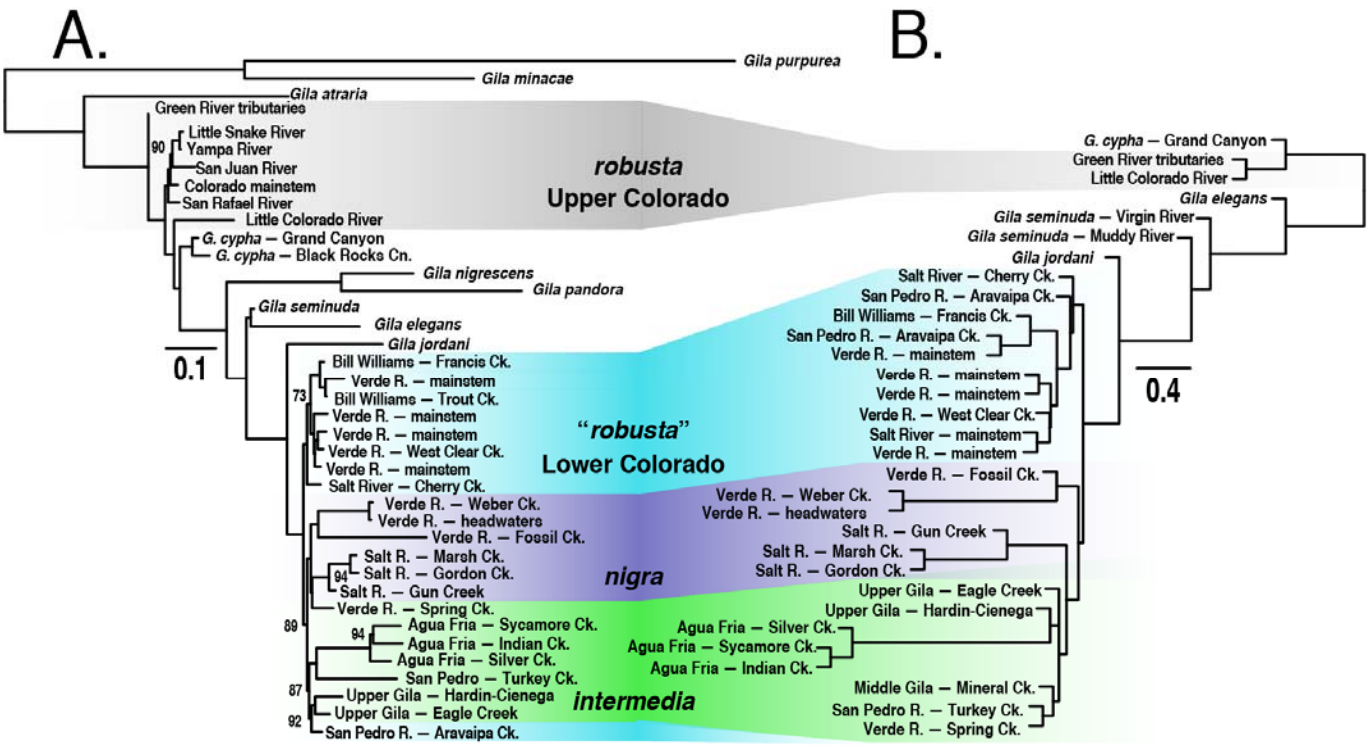

Figure 5: Internode pairs within the anomaly zone, as determined using coalescent-unit transformed branch lengths mapped onto the (A) SVDQUARTETS, (B) POMO, (C) TICR, and (D) concatenated trees (displayed here as cladograms). Paired internodes are color-coded, with those bicolored indicating multiple anomalous divergences.

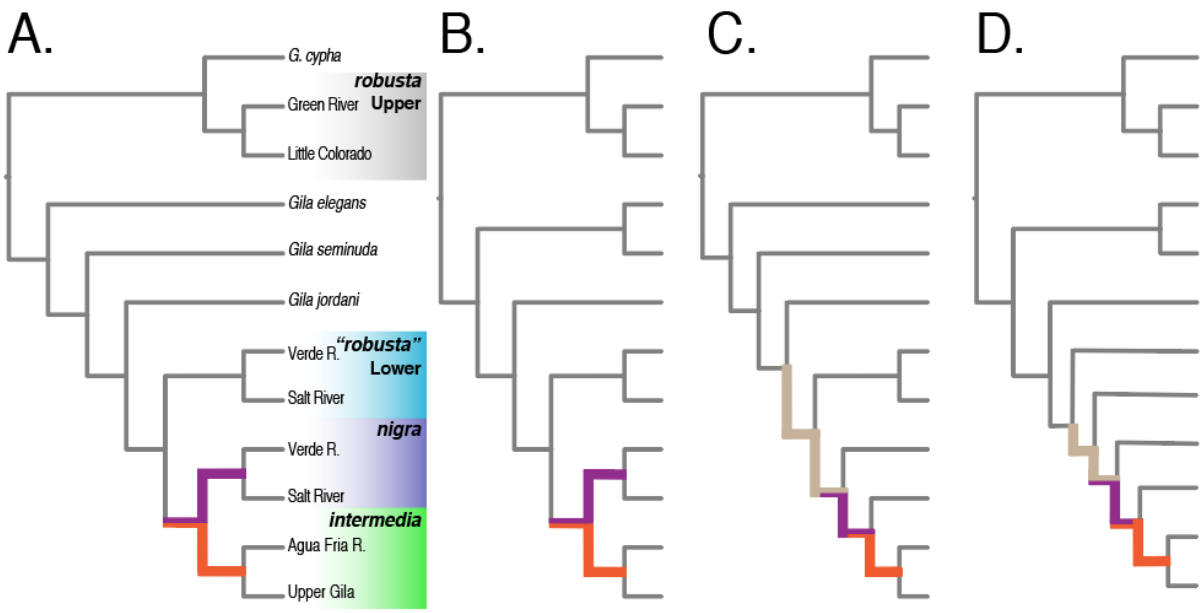

Figure 6: Posterior estimates for divergences times and effective populations sizes (N_e) derived from ECOEVOLITY. Branches are annotated with mean (std. dev.) N_e and posterior probabilities for divergence times are plotted on corresponding nodes. Units are in years, using a static mutation rate of 1.2 e^{-08} substitutions per year. Posterior probabilities for divergence models (insert) suggest the co-divergence of *Gila jordani*, *G. robusta*, *G. nigra*, and *G. intermedia*.

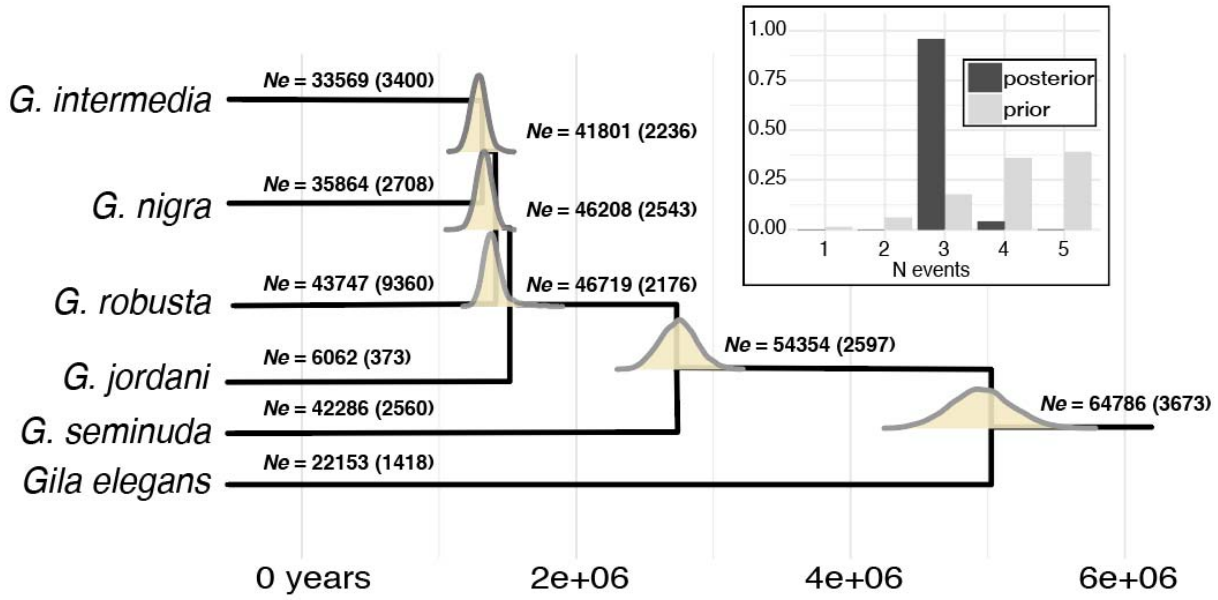

Table S1: Sampling locations and drainages listed for *N*=386 *Gila* individuals and outgroups.

Site	Major Drainage	Location	*N*
Gila atraria	Bonneville, Snake	Multiple localities	14
Gila cypha			
C1	Yampa	Yampa Canyon, Moffat Co., CO	1
C2	Colorado	Black Rocks & Westwater Canyons, Colorado	2
C3	Colorado	Grand Canyon, Coconino Co., AZ	9
C4	Little Colorado	Little Colorado R. mainstem, Coconino Co., AZ	4
Gila ditaenia	de la Concepcion	Rio Magdalena, Sonora, MX	2
Gila elegans	Hatchery	SNARRC, Dexter, NM	11
Gila eremica	Rio Sonora	Rio Sonora, Sonora, MX	2
Gila intermedia			
I1	Verde	Spring Creek, Yavapai Co., AZ	10
I2	Agua Fria	Silver Creek, Yavapai Co., AZ	8
I3	Agua Fria	Indian Creek, Yavapai Co., AZ	9
I4	Agua Fria	Sycamore Creek, Yavapai Co., AZ	10
I5	Gila	Mineral Creek, Pinal Co., AZ	10
I6	Gila	Eagle Creek, Greenlee Co., AZ	10
I7	Gila	Hardin-Cienega Creek, Greenlee Co., AZ	10
I8	San Pedro	Turkey Creek, Santa Cruz Co., AZ	10
Gila jordani	Pahranagat/ White	Key-Pittman WMA refuge population (brood stock sourced from Ash Springs), NV	21
Gila minacae	Rio Yaqui	Rio Bavispe, Chihuahua, MX	7
Gila nigra			
N1	Verde	Fossil Creek, Yavapai Co., AZ	11
N2	Verde	Weber Creek, Gila Co., AZ	8
N3	Verde	Verde River headwaters, Gila Co., AZ	5
N4	Salt	Gordon Creek, Gila Co., AZ	10
N5	Salt	Gun Creek, Gila Co., AZ	11
N6	Salt	Marsh Creek, Gila Co., AZ	10
Gila nigrescens	Multiple	Multiple localities	5
Gila pandora	Rio Grande	Palomas Creek, Sierra Co., NM	6
Gila pulchra	Rio Yaqui	Rio Tomochic, Chihuahua, MX	5
Gila purpurea	Rio Yaqui	San Bernadino NWR, Cochise Co., AZ	2
Gila robusta			
R1†	Bill Williams	Trout Creek, Mohave Co., AZ	9
R2†	Bill Williams	Francis Creek, Yavapai Co., AZ	8
R3†	Verde	Verde River mainstem, Yavapai Co., AZ	8
R4†	Verde	Verde River mainstem, Yavapai Co., AZ	4
R5†	Verde	Verde River mainstem, Yavapai Co., AZ	3
R6†	Verde	West Clear Creek, Yavapai Co., AZ	5
R7†	Salt	Salt River mainstem, Maricopa Co., AZ	5
R8†	Salt	Cherry Creek, Gila Co., AZ	8
R9†	San Pedro	Aravaipa Creek, Pinal Co., AZ	10
R10‡	Green	Upper Green River tributaries, Wyoming	21
R11‡	Yampa	Little Snake River, Wyoming	5
R12‡	Yampa	Upper Yampa River tributaries, Moffat Co., CO	5
R13‡	Green	San Rafael River, Utah	4
R14‡	Colorado	Upper Colorado River mainstem, Colorado	4
R15‡	San Juan	Navajo River, Colorado	2
R16‡	Little Colorado	East Clear Creek, Coconino Co., AZ	16
V1	Virgin	Muddy (Moapa) River, Clark Co., NV	19

V2	Virgin	Virgin River mainstem, Washington Co., UT	17
Ptychocheilus			
P. grandis	Eel River	South Fork Eel River, Humboldt Co., CA	2
P. lucius	Colorado	Yampa River, Moffat Co., Colorado	8

[†]Lower Colorado River basin (below Grand Canyon)
[‡]Upper Colorado River basin (above Grand Canyon)

Figure S1: Phylogram showing results from an unconstrained search using concatenated SNP data in IQ-TREE. Focal nodes are annotated with bootstrap support (values for shallow nodes omitted for clarity).

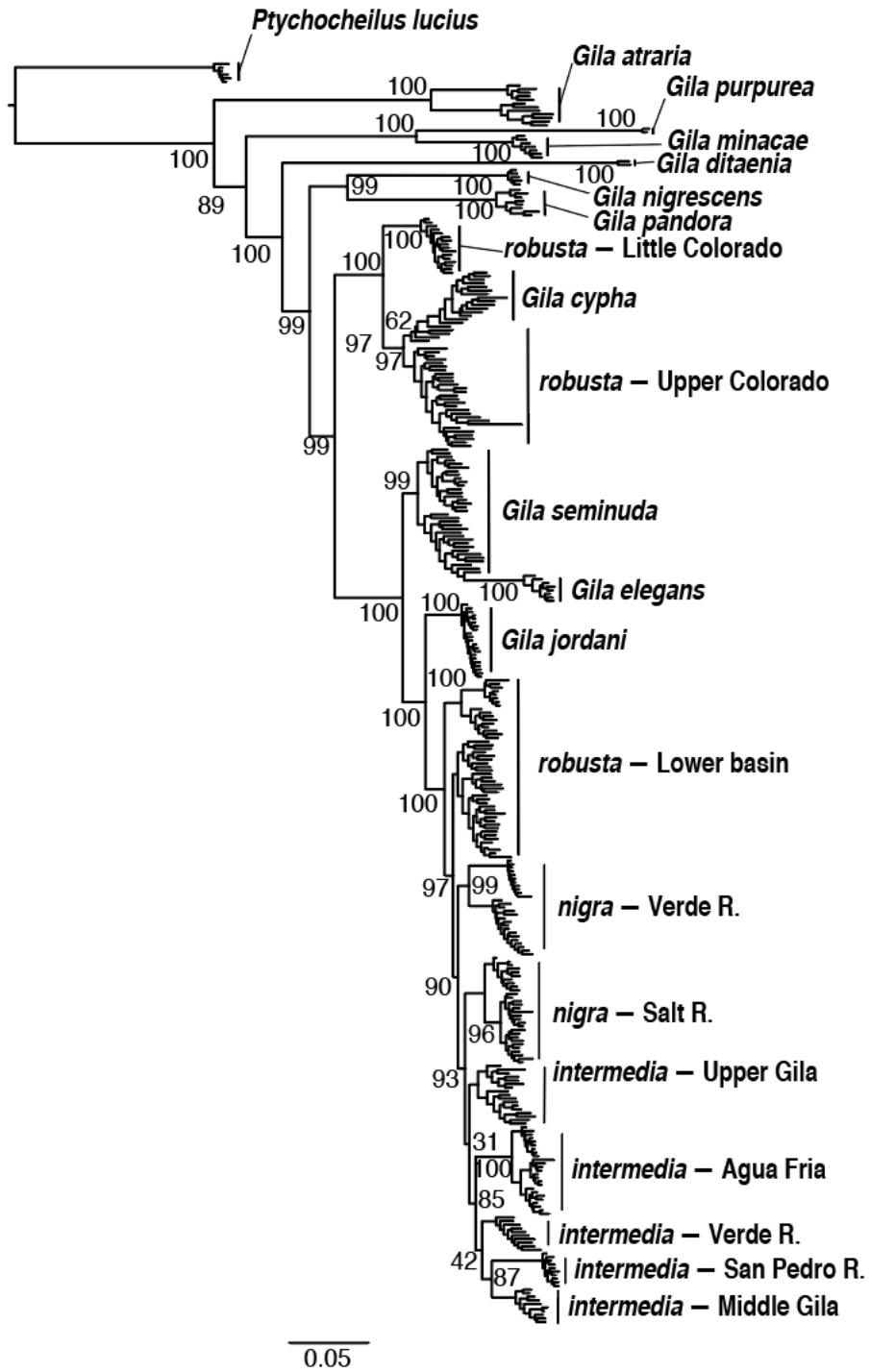

Figure S2: Model selection results for SNAQ/ PHYLONETWORKS; *h*=maximum number of hybrid edges allowed per model; (A) L(*h*) = -log likelihood for the best network of *N*=48 replicate runs per value of *h*; (B) L'(*h*) = 1st order change in L(*h*) = L(*h*) – L(*h*-1); (C) L''(*h*) = 2nd order change in L(*h*) = L'(*h*+1) – L'(*h*); and (D) Δ*h* = L''(*h*) / s(*h*) where s(*h*) is the standard deviation in L(*h*) among replicates.

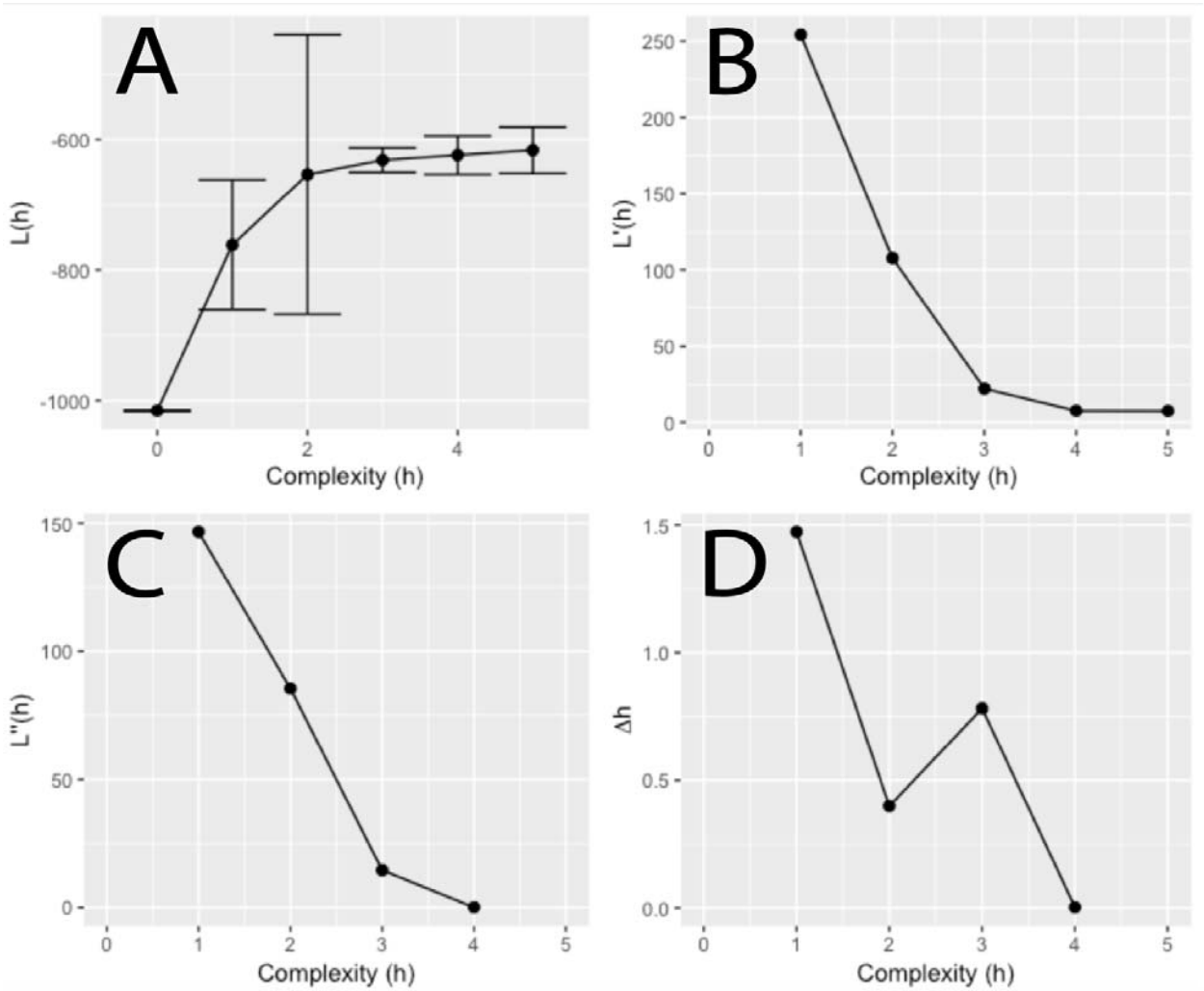

Figure S3: Site-wise log-likelihood differences (ΔSLS) for (A) SVDQUARTETS, (B) POMO, and (C) TICR topologies as compared to an unconstrained concatenated tree. ΔSLS values are transformed as signed square-roots, with positive values indicating increased site-likelihood under the constrained model, and negative values having increased likelihood under the unconstrained concatenated model.

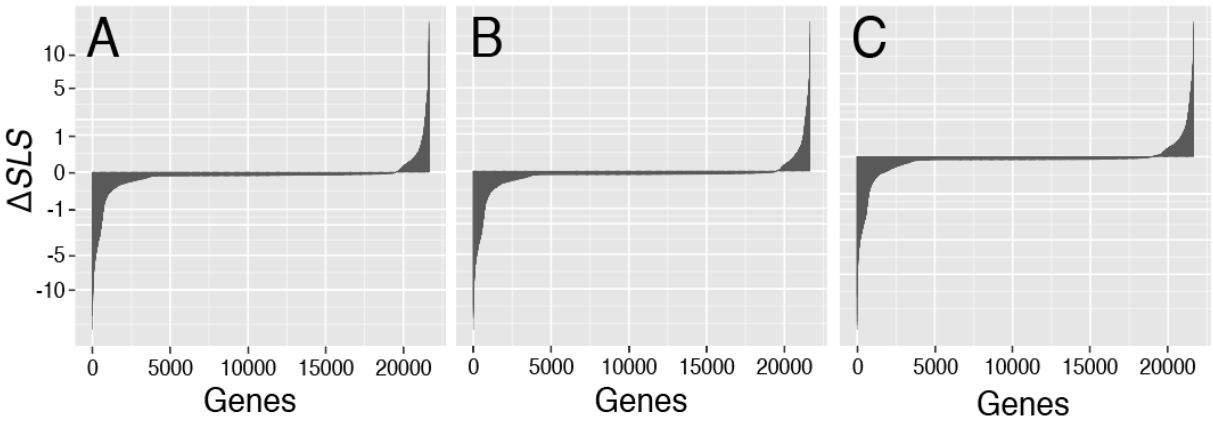

Figure S4: Site-wise concordance factors (s*CF*) for lineage trees produced in IQ-TREE under topological constraints for the (A) SVDQUARTETS, (B) POMO, and (C) TICR results.

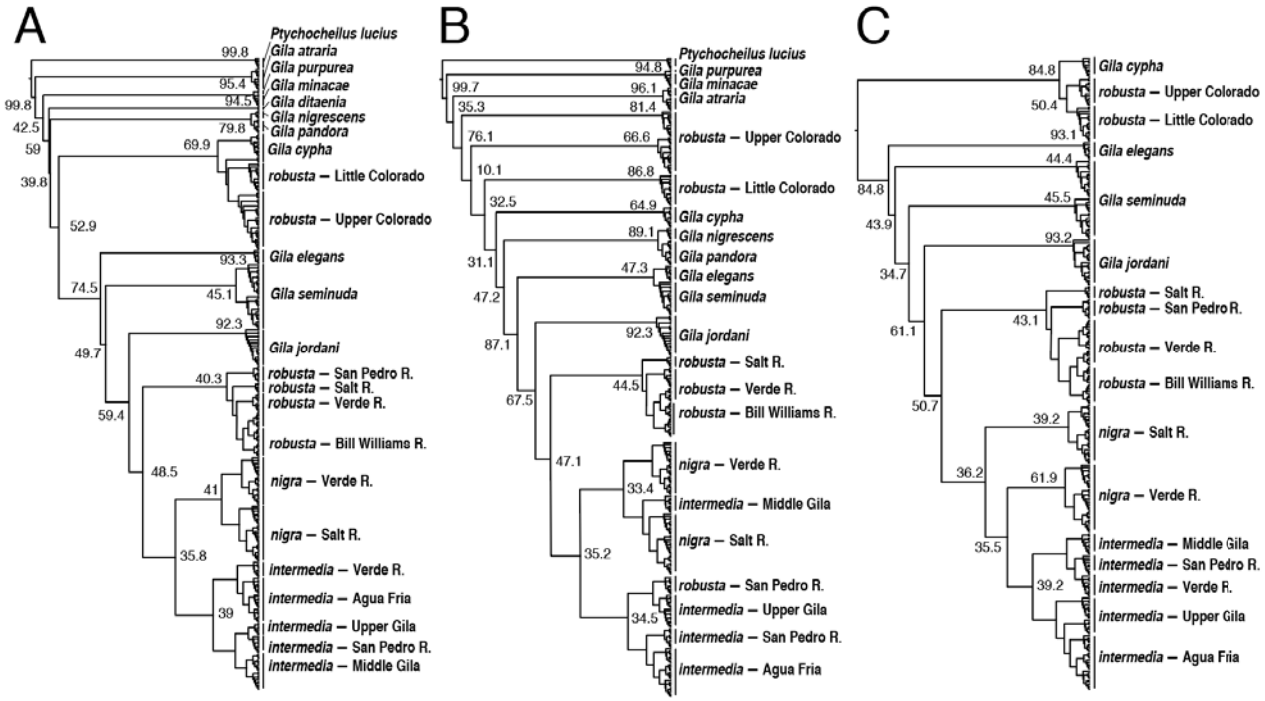

Figure S5: Characterization of site-wise concordance (*s*CF) factors for SVDQUARTETS, POMO, and TICR phylogenies. Panels show (left to right): Linear regression of subtending branch lengths (log-transformed) with *s*CF; node height (cumulative branch lengths from root to focal node); and densities of *s*CF across nodes as compared to the discordance factors for the two conflicting quartet resolutions (*s*DF1 and *s*DF2).

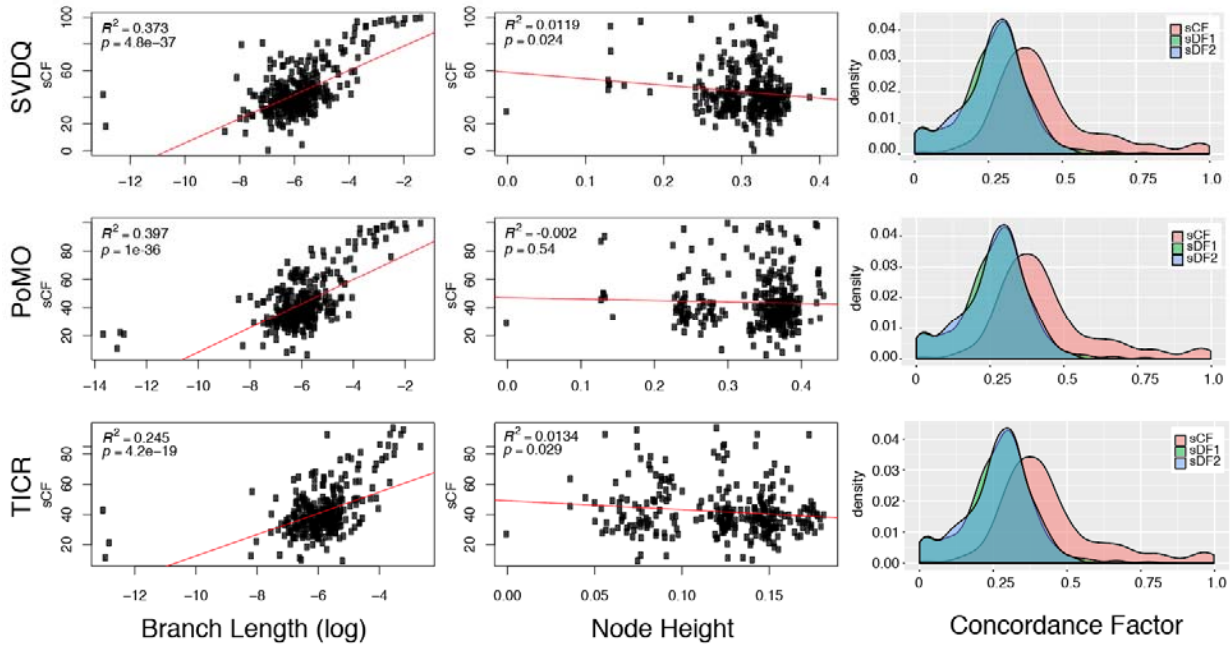
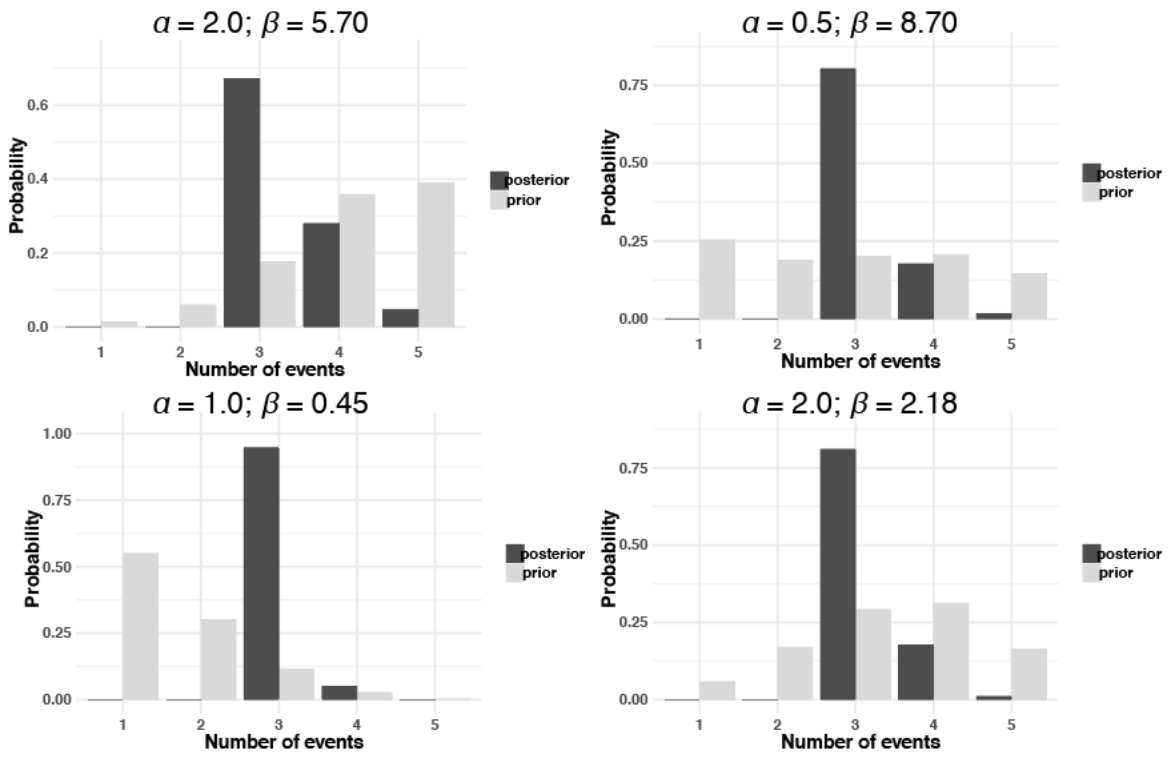
Figure S6: Prior and posterior probabilities for number of independent divergence events in ECOEVOLITY co-divergence models for *Gila*. Parameters across all runs were identical, except for the shape (α) and scale (β) of the gamma-distributed prior on the Dirichlet process concentration.