1   **Title:** Complex hybridization between deeply diverged fish species in a disturbed ecosystem

2

3   **Authors:**

4

5   Shreya M. Banerjee[1,2,3*+], Daniel L. Powell[1,2*], Benjamin M. Moran[1,2], Wilson F. Ramírez-
6   Duarte[2,4], Quinn K. Langdon[1,2], Theresa R. Gunn[1,2], Gaby Vazquez[2], Chelsea Rochman[4], Molly
7   Schumer[1,2,5+]

8

9   [1]Department of Biology, Stanford University
10   [2]Centro de Investigaciones Científicas de las Huastecas "Aguazarca", A.C.
11   [3]Center for Population Biology, University of California, Davis
12   [4]Department of Ecology and Evolutionary Biology, University of Toronto
13   [5]Hanna H. Gray Fellow, Howard Hughes Medical Institutes

14

15   *co-first authorship
16   +Correspondence to: smbanerj@ucdavis.edu and schumer@stanford.edu

17

18   **Abstract**

19

20   Over the past two decades researchers have documented the extent of natural hybridization
21   between closely related species using genomic tools. Many species across the tree of life show
22   evidence of past hybridization with their evolutionary relatives. In some cases, this hybridization
23   is complex – involving gene flow between more than two species. While hybridization is
24   common over evolutionary timescales, some researchers have proposed that it may be even more
25   common in contemporary populations where anthropogenic disturbance has modified myriad
26   aspects of the environments in which organisms live and reproduce. Here, we develop a flexible
27   tool for local ancestry inference in hybrids derived from three source populations and describe a
28   complex, recent hybridization event between distantly related swordtail fish lineages
29   (*Xiphophorus*) and its potential links to anthropogenic disturbance.

30

31

Impact Summary

As sequencing tools have advanced, we have found that barriers between animal species are more porous than once thought. Researchers have found evidence for hybridization between species throughout many branches of the tree of life. In some cases, these hybridization events can involve more than two species. Here, we develop a flexible and user-friendly tool that can be used to identify three-way hybrids and report the discovery of hybrids with ancestry from three swordtail (*Xiphophorus*) species from an anthropogenically impacted site on the Río Calnali in Hidalgo, Mexico. Researchers have studied hybrids between two *Xiphophorus* species along this river for decades, but this is the first documented case of hybridization involving three species. We explore hypotheses for what drove this hybridization event, including anthropogenic pollutants and reduced water quality.

Introduction

Hybridization, or genetic exchange between species, is common in diverse organisms across the tree of life, and can have important evolutionary consequences (Moran *et al.*, 2021). The genetic, ecological, and evolutionary outcomes of hybridization are varied, from facilitating rapid adaptation to exposing genetic incompatibilities. Examples from the recent literature include introgression as a source of genetic rescue (Oziolor *et al.*, 2019) and hybridization resulting in decreased tolerance of thermal stressors (Payne *et al.*, 2022). While evidence of ancient introgression in the genomes of diverse taxa suggests that hybridization is common in the evolutionary history of many species (Moran *et al.*, 2021), a growing number of studies point towards anthropogenic disturbance as contributing to the formation of new hybrid zones (Fisher *et al*., 2006; Kelly *et al*., 2010; Pampoulie *et al*., 2020). As humans modify habitats, there are numerous mechanisms by which anthropogenic environmental disturbance can cause hybridization. These include phenological changes (Chunco, 2014; Vallejo-Marín & Hiscock, 2016), species introductions (Oziolor *et al.*, 2019), habitat alterations that lead to new contact zones (Kelly *et al.* 2010), and decreased encounter rates of conspecifics (Willis *et al.*, 2011). Environmental disturbance, e.g., pollution with urban effluents and/or reduced water quality, can also directly impact sensory communication, disrupting signals used in mate choice (Seehausen *et al.*, 1997; Fisher *et al.*, 2006; Powell *et al.,* 2022).

Hybridization between pairs of species has been intensively studied for several decades, but a growing body of literature highlights that hybridization events can be complex, involving three or more species (Heliconius Genome Consortium 2012; Toews *et al.* 2018; Langdon *et al.* 2019; Grant and Grant 2020, Natola *et al.* 2022). These types of complex hybridization events are likely to be more common in groups where many species are interfertile and have overlapping ranges. The evolutionary consequences of these events are not as well understood. One possible outcome is "conduit" introgression, where genetic exchange can occur between species that are not in geographic contact through hybridization with a third species (Langdon *et al.* 2019; Grant and Grant 2020; Natola *et al.* 2022). Such dynamics could explain observations in the empirical literature such as cases where gene flow is inferred between geographically isolated species (Cui *et al.*, 2013); although there are other potential causes of these patterns including introgression from a now extinct lineage (Ottenburghs, 2020).

Xiphophorus species and their hybrids from the Sierra Madre Oriental of eastern Mexico have been intensively studied over two decades. Much of this work has focused on hybridization between two sister-species from the "Northern swordtail" clade, *X. birchmanni* and *X. malinche* (Fig. 1). Throughout its range *X. birchmanni* is sympatric with a distantly related species in the "platyfish" clade, *X. variatus* (Fig. 1). *X. variatus* is also common at many sites where *X. birchmanni* x *X. malinche* hybrids (i.e. Northern swordtail hybrids) are found, but does not reach the high elevations inhabited by pure *X. malinche*. Genetic and historical estimates indicate that hybridization has been occurring between *X. birchmanni* and *X. malinche* in the Río Calnali for more than 40 generations (Rosenthal *et al.*, 2003; Schumer *et al.*, 2014, 2017). However, despite extensive collections in regions where they co-occur over the last two decades, no hybrids between *X. variatus* and *X. birchmanni* or *X. variatus* and *X. birchmanni* x *X. malinche* have been reported.

Here, we characterize a newly discovered three-way hybridization event involving *X. birchmanni* x *X. malinche* hybrids and *X. variatus* at an anthropogenically disturbed site on the Río Calnali (hereafter the Tlalica site; Fig. 1). To facilitate this analysis, we develop an easy to

91 use and accurate extension of the *ancestryinfer* pipeline (Schumer *et al.*, 2020) that enables local
92 ancestry inference of individual hybrids formed from three source populations. We initially
93 identified three-way hybrids based on morphology, and confirmed our observations with whole
94 genome-sequencing and local ancestry inference. We also characterized water quality and
95 chemistry (relevant to the visual and olfactory environment) at Tlalica and other sites along the
96 Río Calnali to explore relationships between environmental disturbance and hybridization. Our
97 results hint at a connection between anthropogenic disturbance and hybridization in these deeply
98 diverged species with a long history of reproductive isolation in sympatry.
99
100

Methods

Morphological evidence of hybridization between X. variatus, X. birchmanni, and X. malinche

Three-way hybrids were first identified at the Tlalica site based on their unusual phenotype combinations. Male *X. variatus*, *X. birchmanni*, and *X. malinche* differ in several traits (females are phenotypically similar in many *Xiphophorus* species). *X. malinche* has a modification of the caudal fin known as the "sword" which is absent in the two other species and *X. birchmanni* has a much larger dorsal fin compared to the other two species (Fig. 1). *X. variatus* is characterized by a distinctive diamond body shape (Fig. 1A, 2A), and two horizontal stripes composed of melanophores bracketing the lateral line. By comparison, *X. malinche*, *X. birchmanni*, and their hybrids are more elongated (Fig. 2A) and have a single, broader horizontal stripe. *X. variatus* have polymorphic melanophore tailspot patterns (described in Borowsky, 1980; Culumber & Rosenthal, 2013) that are distinct from the polymorphic melanophore patterns present in some *X. birchmanni* and *X. birchmanni* x *X. malinche* hybrid individuals (Rauchenberger *et al.*, 1990; Culumber, 2014; Powell *et al.,* 2020).

We noticed that two adult males sampled from Tlalica in June 2021 appeared to have *X. variatus*-like characteristics, such as a diamond body shape and dual horizontal stripes, *X. birchmanni*-like characteristics, such as large body size and large rounded dorsal fins, and *X. malinche*-like characteristics, such as short sword extensions (Fig. 2). For these fish and additional male three-way hybrids identified in subsequent collections (n = 12) we measured standard length, body depth, peduncle depth, caudal fin length, dorsal fin width, dorsal fin height, and sword length from photographs of anaesthetized adult male fish using ImageJ (Fig. S1; Schneider, Rasband, & Eliceiri, 2012). We included phenotypes of *X. birchmanni* x *X. malinche* hybrids from a nearby population (Calnali Low; n = 9) and of pure parental species individuals from sites where hybrids have not previously been reported (*X. variatus* from Coacuilco: n = 27, *X. birchmanni* from Coacuilco: n = 24; *X. malinche* from Capac: n = 5) (Fig. 1C). We performed principal component analyses to assess morphological differences between groups.

Genomic libraries of putative hybrids

For the two male individuals collected in June of 2021 that were morphologically identified as likely three-way hybrids, we produced high-coverage whole genome data (Table S1) following Schumer *et al.* (2016, 2018). Briefly, we extracted DNA from fin clips using half reactions of the Agencourt DNAdvance kit. DNA was sheared by sonication and end-repaired with dNTPs, T4 DNA polymerase, Klenow DNA polymerase, and T4 polynucleotide kinase, then A-tailed with Klenow exonuclease and dATP. Universal Illumina adapters were ligated onto the A-tailed sample using DNA ligase. Samples were purified with the Qiagen PCR Purification kit between steps. Samples were PCR amplified for 12 cycles using the Phusion PCR kit. After amplification, the final PCR product was purified with 18% SPRI beads and sent to Admera Health (South Plainfield, NJ, USA) for sequencing on an Illumina HiSeq 4000.

Preliminary investigation of three-way hybrids with sppIDer

We used the competitive mapping and read depth analysis pipeline, sppIDer (Langdon *et al.*, 2018) as an initial approach to investigate the potential genetic contributions to the two male fish sampled in June of 2021 from Tlalica. We created a combined fasta file by concatenating *X. birchmanni, X. malinche,* and *X. variatus* reference genomes (described in Powell *et al.*, 2021).

147 Reads from the high coverage Tlalica males were mapped to this combination reference genome
148 and uniquely mapped reads were used by sppIDer to estimate the proportion of the genome
149 derived from each species (see Langdon *et al.*, 2018).

151 *PSMC and analysis of whole genome sequences*
152   Raw Illumina reads from an *X. variatus* individual from the Coacuilco population
153 (previously sequenced by Powell *et al.*, 2020) were aligned to a *de novo* assembly derived from
154 *X. variatus* that was scaffolded with cactus (Armstrong *et al.*, 2020) to a chromosome-level *X.*
155 *birchmanni* assembly (Powell *et al.*, 2020). Alignment of reads was performed using *bwa* (Li &
156 Durban, 2009). We used PicardTools and *GATK* (Van der Auwera & O'Connor, 2020) to realign
157 mapped reads around indels and call variant sites in a gvcf format. Individual jobs were run for
158 each chromosome for indel realignment and variant calling. We combined gvcf files for all
159 chromosomes using *bcftools* (Danecek *et al.*, 2021). We filtered variant and invariant sites from
160 this combined gvcf file as previously described (Schumer *et al.*, 2018). Briefly, we used hard-call
161 thresholds for variant quality scores recommended by GATK and previously validated in
162 swordtails using pedigree data (Schumer *et al.*, 2018). For both variant and invariant sites, we
163 masked sites within 5 bp of an INDEL or >2X or <0.5X the average genome-wide coverage. To
164 generate a pseudo-fasta file reflecting variant and masked sites, we used a custom script to
165 generate an insnp file (https://github.com/Schumerlab/Lab_shared_scripts). We used *seqtk*
166 (https://github.com/lh3/seqtk) to generate a new fasta file with variant sites and masked sites
167 updated to reflect the *X. variatus* individual being analyzed.
168   We next used this data to infer changes in historically effective population size through
169 time using the Pairwise Sequential Markovian Coalescence approach (PSMC, Li & Durbin,
170 2011). We used a custom script to convert the fasta file to a fastq file (the required input format
171 for PSMC) with uniform quality scores, (https://github.com/Schumerlab/Lab_shared_scripts) and
172 used *seqtk* to exclude scaffolds that did not belong to the 24 major *Xiphophorus* chromosomes.
173 We assumed a mutation rate of 3.5×10^{-9} per basepair per generation, a generation time of half a
174 year, and set the -r parameter to 2 (Schumer *et al.*, 2018). We compared these results for *X.*
175 *variatus* to those previously published for *X. birchmanni* and *X. malinche* (Schumer *et al.*, 2018).
176 For comparison to our results for *X. variatus*, we included only one sample per species.

178 *Low-coverage whole genome sequencing of individuals collected from Tlalica and nearby sites*
179   We extracted DNA from fin clips using the Agencourt DNAdvance bead-based kit as
180 specified by the manufacturer except that we used half-reactions. We prepared tagmentation-
181 based libraries for low-coverage whole genome sequencing as previously described (Payne *et al.*
182 2022). Briefly, DNA was enzymatically sheared using the Illumina Tagment DNA TDE1
183 Enzyme and Buffer Kit, amplified in a dual-indexed PCR reaction for 12 cycles, pooled, and
184 bead purified with 18% SPRI magnetic beads. Libraries were sent to Admera Health (South
185 Plainfield, NJ, USA) to be sequenced on a HiSeq 4000.

187 *Design and performance tests of three-way local ancestry inference*
188   To perform three-way local ancestry inference, we adapted our previously developed
189 pipeline, *ancestryinfer* (which only allowed for two source populations; Schumer *et al.*, 2020), to
190 accommodate three reference genomes and source populations. Briefly, we modified the
191 program to detect whether two or three reference genomes were provided in the configuration
192 file (see Appendix 1 – user manual). When three reference genomes are provided, *ancestryinfer*

193 maps reads to all three genomes and identifies and excludes any reads that do not map uniquely
194 to any of the three references. Using the coordinate space of reference genome 1, it tabulates
195 counts for each allele at each ancestry-informative site and runs AncestryHMM in the three
196 source population mode (Corbett-Detig & Nielsen, 2017). Users can optionally provide priors for
197 the number of generations since initial admixture for each source population and priors for
198 admixture proportions from each source population.
199   We searched for candidate ancestry-informative sites from high coverage whole genome
200 sequences (*X. variatus* n=2, *X. malinche* n=4, *X. birchmanni* n=26; samples from Schumer *et al.*,
201 2018; Powell *et al.*, 2020, 2021). Although we use a small number of high coverage samples in
202 this initial step based on available data, we filter sites using a large number of individuals of each
203 species (see below). We identified biallelic sites that differentiated any of the three focal
204 species. Initial analysis suggested that issues with accuracy arise from imbalance in the number
205 of ancestry-informative sites between pairs of species. We thinned to an approximately
206 equivalent number of informative sites between all pairs of species. To do so, we retained all
207 ancestry-informative sites that distinguished *X. birchmanni* and *X. malinche*, and every other site
208 that distinguished *X. variatus* from either of these two species.
209   We refined this candidate set of ancestry-informative sites using low-coverage population
210 data from each species (*X. malinche* n=28, *X. birchmanni* n=107 – Schumer *et al.*, 2018; *X.*
211 *variatus* n=145 – this study). Note that per basepair heterozygosity is much lower in *X. malinche,*
212 approximately ¼ of the levels observed in *X. birchmanni* or *X. variatus* (0.0003 per basepair
213 versus ~0.001 respectively). Low nucleotide diversity in *X. malinche* is attributable to low
214 historical effective population sizes in this species (Schumer *et al.*, 2018). Average coverage per
215 individual was ~1X. Because this is low coverage data, we did not perform explicit variant
216 calling but instead used bcftools mpileup to determine the observed counts for each allele at each
217 candidate ancestry-informative site in the three source populations. We then excluded ancestry-
218 informative sites that did not have equal to or greater than a 90% frequency difference between
219 at least one pair of species (e.g. *X. birchmanni* vs *X. malinche*, *X. birchmanni* vs *X. variatus*, *X.*
220 *malinche* vs *X. variatus*). This resulted in a final set of 997,366 ancestry-informative markers
221 throughout the 750 Mb genome.
222   Using this set of ancestry-informative sites and estimated parental allele frequencies
223 determined from the individuals described above, we ran *ancestryinfer* on a set of parental
224 individuals that were not used in the training datasets ($n_{variatus}$ = 30; $n_{birchmanni}$ = 12; $n_{malinche}$ = 10)
225 as a first performance check on empirical data. We found that *ancestryinfer* correctly inferred
226 that these individuals were unadmixed and derived from the correct parental population (Fig.
227 S2). We also performed a similar analysis on hybrids from *X. birchmanni* x *X. malinche* hybrid
228 populations that are allopatric with respect to *X. variatus* (Totonicapa, n=30, and Tlatemaco, n=
229 23; Fig. S3). See Supporting information 1 for additional performance testing and simulations.
230   We note that we do not have access to any populations in which *X. birchmanni* does not
231 co-occur with *X. variatus*. Thus, if there was admixture between *X. birchmanni* and *X. variatus*
232 in the *X. birchmanni* source populations that we have failed to detect, our approach could
233 underestimate the degree of contemporary gene flow between these species.
234
235 *Local ancestry inference and data processing of three-way hybrids*
236   We proceeded with local ancestry analysis of individuals collected from the Tlalica
237 population (n=64) and previously collected samples from upstream (n=553; Table S2) and
238 downstream (n=25) of this site on the Río Calnali. We also ran *ancestryinfer* on all sequenced

239 pure *X. birchmanni* and *X. variatus* from Coacuilco, an allopatric site in a different drainage, to
240 confirm their ancestry (n=745; samples from Powell *et al.*, 2020 and this study; Table S3). For
241 hybrid individuals from the Río Calnali, we provided priors for admixture proportions from the
242 three source populations based on sppIDer results. *ancestryinfer* accepts priors for the time since
243 initial admixture for all source populations (see Corbett-Detig & Nielsen, 2017). Based on past
244 results for *X. birchmanni* x *X. malinche* (Schumer *et al.* 2014, 2017) and the results of an initial
245 run of *ancestryinfer* without specifying a prior for admixture time, we set the prior admixture
246 time between *X. malinche* and *X. birchmanni* to 50 and the prior admixture time between this
247 admixed population and *X. variatus* to 2. We excluded individuals with fewer than 500,000
248 reads, based on previous simulation results that indicated accuracy of local ancestry inference is
249 reduced in individuals with <0.2X coverage (Schumer *et al.*, 2020). This analysis resulted in
250 posterior probabilities for each of the six possible ancestry states (homozygous *X. birchmanni,*
251 homozygous *X. malinche,* homozygous *X. variatus* and each possible heterozygous combination)
252 at 900,343 ancestry-informative sites throughout the genome.
253   We used a posterior probability threshold of 0.9 to convert ancestry probabilities to hard-
254 calls. For ancestry-informative sites that did not have a probability of ≥0.9 for any ancestry state,
255 we converted the probabilities for those sites to NA. The average level of missing data in three-
256 way hybrid individuals after imposing this hard-call threshold was 0.03%.
257
258 *Water quality and chemistry at Tlalica*
259   Tlalica is ~1 km away from the municipal landfill of the city of Calnali and 2.7 km
260 downstream from the outfall of a sewage treatment plant. During the wet season, we observed a
261 small tributary running through the municipal landfill into the Río Calnali approximately 250
262 meters upstream of Tlalica. On one sampling occasion, we observed sewage effluent flowing
263 into the river from the treatment plant upstream of Tlalica. On another occasion a break in the
264 sewer line upstream of the treatment plant led to contamination of the river ~3 km upstream of
265 the sampling site. Accordingly, we expected water quality at the Tlalica site to be lower than
266 upstream sites (Fig. 1), and hypothesize that this could contribute to the hybridization observed
267 between distantly related *Xiphophorus* species.
268   We collected water samples in May and June of 2022 at a relatively undisturbed upstream
269 site (Plank), and at the three sites where we found genetic evidence of one or more three-way
270 hybrids (see Results; Plaza, Calnali Low, and Tlalica; Fig. 1). All focal sites contained both *X.*
271 *variatus* and *X. birchmanni* x *X. malinche* hybrids. We measured fluorescent dissolved organic
272 matter (DOM) and turbidity using an EXO2 multiparameter sonde (YSI, Yellow Springs, OH).
273 We used a 9300 colorimeter (YSI, Yellow Springs, OH) to quantify ammonia. We quantified
274 concentrations of dissolved copper (using a 0.45 μm polyethersulfone membrane, and
275 acidification to pH ~2.0 with trace metal grade nitric acid) in water using inductively coupled
276 plasma mass spectrometry (ICP-MS) by following the modified version of the
277 APHA3030B/6020A methods. See Supporting information 2 for additional water quality and
278 chemistry metrics collected.
279
280
281

Results

Morphological results and demographic survey of the Tlalica population

Three-way hybrids were morphologically distinct from pure parental individuals and from *X. birchmanni* x *X. malinche* hybrids found in nearby populations (Fig. 2, Table S4). They were most morphologically distinct from other groups analyzed along PC1 (74.9% of variation explained), and clustered with *X. malinche*, *X. variatus*, and Northern swordtail hybrid individuals along PC2 (21.6% of variation explained; Fig. 2).

Based on visual phenotypes, a large majority of individuals collected from the Tlalica site in May 2021, November 2021, February 2022, and May 2022 were classified as *X. variatus*. From visual phenotypic data alone, *X. variatus*-like individuals outnumbered *X. birchmanni* x *X. malinche* hybrids by ~33:1 (based on 571 individuals collected in May 2022). Genotype data from a subset of fish collected at Tlalica that were categorized as *X. variatus* indicate that we can accurately differentiate them based on morphology alone (Fig. S4).

History of divergence between X. variatus, X. birchmanni, and X. malinche

X. variatus and the Northern swordtail clade to which *X. birchmanni* and *X. malinche* belong are deeply divergent (Fig. 1; Schumer *et al.* 2014; 2016; 2018). Pairwise sequence divergence between *X. birchmanni* – *X. variatus* and *X. malinche* – *X. variatus* is 1.42% and 1.43% respectively. Because *X. malinche* has undergone a severe recent bottleneck (see below, Fig. 1), we focus on comparisons between *X. birchmanni* and *X. variatus* here. The per-site heterozygosity (θ_π) for *X. variatus* is 0.11%, similar to that observed in *X. birchmanni* (0.12%; Schumer *et al.*, 2018). Assuming that the ancestral θ is close to that of *X. birchmanni* and *X. variatus,* we estimate the divergence time between the two clades is approximately 7.5 in units of 4*Ne* generations (using the relationship $T_{div/2Ne} = D_{xy}/\theta - 1$).

Comparing PSMC results for *X. variatus* to those previously inferred for *X. birchmanni* and *X. malinche* highlights differences in the inferred effective population sizes of each species over time (Fig. 1). We estimated the long-term effective population size of *X. variatus* from one individual to be approximately 50,000 individuals, similar to our previous estimates for *X. birchmanni* (48,000–53,000; Powell *et al.*, 2021). However, the timing and extent of demographic fluctuations varies between the two species (Fig. 1). *X. malinche* differs more substantially in its inferred demographic history from both *X. birchmanni* and *X. variatus* given the strong bottleneck that has persisted through much of its recent history (Fig. 1; Schumer *et al.*, 2018).

Assuming a long-term effective population size of 50,000 individuals, and the divergence time in 4*Ne* generations calculated above, we estimate the divergence time between *X. variatus* and *X. birchmanni* (and *X. malinche*) to be approximately 1.5 million generations.

Ancestry analysis of Tlalica hybrids and nearby populations

Initial analysis of genomic data with sppIDer indicated that males visually categorized as three-way hybrids were likely hybrids between *X. variatus, X. birchmanni,* and *X. malinche* (Fig. S5). We found that for each Tlalica male 52% of the reads preferentially mapped to the *X. variatus* reference genome, 30-34% mapped to the *X. birchmanni* reference genome, 26-29% mapped to the *X. malinche* reference genome.

This finding led us to develop local ancestry inference for three-way admixture for these species (see Methods). The results of our local ancestry inference analysis indicated that

328 seventeen individuals sequenced from the Tlalica population were early generation hybrids
329 between *X. variatus, X. birchmanni,* and *X. malinche*. Among individuals at Tlalica with
330 Northern swordtail ancestry, we estimate the frequency of three-way hybrids to be ~10% of
331 individuals; this estimate is based on individuals collected before May 2022 since in later
332 samples we selectively collected suspected three-way hybrids (Fig. 3B; Table S5). Three-way
333 hybrid individuals derived ~50-75% of their genomes from *X. variatus*.
334   Samples collected from Tlalica that did not show evidence of ancestry derived from all
335 three species fell into two categories: hybrids between *X. birchmanni* and *X. malinche* and pure
336 *X. variatus* (Fig. 3). Hybrids between *X. birchmanni* and *X. malinche* derived approximately 25-
337 75% of their genomes from either of these parent species (Fig. 3). This is consistent with
338 admixture proportions observed in hybrids between *X. birchmanni* and *X. malinche* at sites up
339 and downstream of Tlalica (Fig. 1; Fig. 2; Schumer *et al.*, 2017). This suggests that three-way
340 hybrids with *X. variatus* originated from admixture with already extant *X. birchmanni* x *X.
341 malinche* hybrids. Indeed, *X. birchmanni* and *X. malinche* ancestry tract lengths in three-way
342 hybrids are similar to those observed in *X. birchmanni* x *X. malinche* hybrids at nearby sites
343 (average minor parent ancestry tract length ~150-200 kb; Schumer *et al.*, 2017). Samples
344 preliminarily categorized as *X. variatus* based on morphology from the Tlalica population show
345 no evidence of introgression from *X. birchmanni* or *X. malinche* (Fig. S4).
346   Given the proximity of Tlalica to previously sampled sites on the Río Calnali (~3 km),
347 and the fact that *X. variatus* is sympatric with several *X. birchmanni* x *X. malinche* hybrid
348 populations along the river (Fig. 1, Fig. 3), we asked if there is evidence of three-way
349 hybridization at other sites. We performed three-way local ancestry inference on 578 historically
350 and newly collected samples from other sites on the Río Calnali from 2003 to 2022 (previously
351 assumed based on morphology to represent *X. birchmanni* x *X. malinche* hybrids; Schumer *et al.*,
352 2017; Table S2). We identified only two three-way hybrids from these sites — a female from the
353 Plaza site who derived ~25% of her genome from *X. variatus* and a male from Calnali Low who
354 derived 50% of his genome from *X. variatus* (Fig. 3; 0.4% of sequenced specimens from other
355 sites).
356
357 *Inference about the generation of admixture using ancestry tract lengths*
358   Observed admixture proportions for three-way hybrids (25-75% *X. variatus* ancestry
359 across all samples) suggests that these individuals might be early generation hybrids between *X.
360 variatus* and Northern swordtail hybrids. Some of these samples are clearly first generation
361 hybrids between *X. variatus* and *X. birchmanni* x *X. malinche* hybrids based on local ancestry
362 patterns (N=17; Fig. 3). These individuals derived 50% of their genomes from *X. variatus* and
363 were heterozygous for *X. variatus* ancestry at nearly every ancestry-informative site across the
364 genome (>99.5% across individuals; Fig. 3). The few sites inferred to be homozygous *X.
365 variatus* are consistent with our expected error rate (see Methods).
366   Two individuals with substantial *X. variatus* ancestry did not have ancestry patterns
367 consistent with those expected for first generation hybrids. Both their observed admixture
368 proportions (25% and 75% *X. variatus,* respectively) and the lengths of ancestry tracts
369 heterozygous or homozygous for *X. variatus* ancestry indicate that these individuals are likely
370 backcrosses between a three-way hybrid with a pure *X. variatus* individual (Fig. 3). The
371 identification of two second-generation three-way hybrids indicates that hybrids between *X.
372 birchmanni*, *X. malinche*, and *X. variatus* are at least partially fertile.

373   Since the mitochondrial genome is maternally inherited, it allows us to infer the likely
374 maternal ancestry for the three-way hybrids identified. All three-way hybrids (n = 19) sequenced
375 had mitochondrial ancestry derived from either *X. birchmanni* or *X. malinche* (Table S5),
376 suggesting that the mothers of all three-way hybrid individuals sampled to date were *X.*
377 *birchmanni* x *X. malinche* hybrids. Skews in maternal ancestry may be the result of population
378 demography, differences in the strength of mate discrimination across groups, or impacts of
379 cross direction on the viability of hybrids.
380
381 *Water quality and chemistry analysis*
382   Dissolved organic matter, ammonia, dissolved copper, and turbidity were all elevated at
383 the Tlalica sampling site in comparison with the relatively undisturbed upstream Plank site when
384 we collected water samples in the spring of 2022 (Fig. 2B). Two sites between Plank and Tlalica
385 where three-way hybrids were detected at low frequencies, Plaza and Calnali Low, had
386 intermediate values of dissolved organic matter, ammonia, and dissolved copper (Fig. 2B). With
387 one season of data collection, we focus on qualitative patterns in the results. However, our results
388 show a pattern of elevated pollution and turbidity at sites with higher frequencies of
389 hybridization between *X. variatus* and *X. birchmanni* x *X. malinche* hybrids. Additional water
390 quality and chemistry metrics are shown in Fig. S6.

Discussion

391
392
393   Here we characterize wild caught *Xiphophorus* individuals with ancestry from three
394   parental species — *X. variatus*, *X. birchmanni*, and *X. malinche* — using multiple approaches.
395   Though *X. birchmanni* x *X. malinche* hybrids were first reported nearly two decades ago
396   (Rosenthal *et al.* 2003), hybridization with *X. variatus* has not been previously reported. These
397   results are remarkable given that *X. variatus* have ~1.5% sequence divergence from the Northern
398   swordtail clade. This is similar to divergence between chimpanzees (*Pan troglodytes*) and
399   gorillas (*Gorilla spp.*) (Chen & Li, 2001), highlighting the unusually deep nature of this
400   hybridization event.
401   To facilitate this work, we developed a user-friendly pipeline for running local ancestry
402   inference in hybrids which derive their genomes from three source populations, as well as a
403   collection of simulation scripts to test expected performance (see Appendix 1). Although several
404   methods have been developed that accommodate local ancestry inference with three source
405   populations (reviewed in Wu *et al.* 2021), there are few pipelines available that allow researchers
406   to move from raw reads to probabilities of ancestry across the genome. By expanding our
407   previously developed local ancestry inference pipeline and simulation scripts (Schumer *et al.*,
408   2020), we are able to provide a toolkit that can be used by researchers to study complex
409   hybridization events in diverse species groups.
410   The existence of natural hybrids between *X. variatus* and *X. birchmanni* x *X. malinche*
411   hybrids is surprising. Despite extensive collections over the past two decades and substantial
412   range overlap between *X. variatus, X. birchmanni,* and *X. birchmanni* x *X. malinche* hybrids,
413   there are no reports of contemporary hybridization involving *X. variatus*. However, past work
414   has identified a small genomic contribution from the lineage leading to *X. variatus* to the
415   ancestors of *X. birchmanni* and *X. malinche* (Schumer *et al.* 2018), which indicates that gene
416   flow has occurred historically. This event contributed ~2-4% of the genome to present-day *X.*
417   *birchmanni* and *X. malinche* (Schumer *et al.* 2018). Together with our present data, this suggests
418   that hybridization between these groups is possible, albeit rare.
419   *X. variatus* are sympatric with natural *X. malinche* x *X. birchmanni* hybrid populations at
420   several upstream sites along the Río Calnali where we have found no evidence of three-way
421   hybridization. In our analyses of 642 Northern swordtail individuals from sites along the river,
422   we have identified fifteen individuals with three-way hybrid ancestry at Tlalica, one individual at
423   Plaza, and one individual at Calnali Low. Moreover, *X. birchmanni* and *X. variatus* are sympatric
424   over much of *X. birchmanni*'s range, but there has been no evidence of hybridization outside of
425   the three-way hybrids reported here (Kallman & Kazianis 2006).
426   The lack of evidence for contemporary hybridization involving *X. variatus*, *X.*
427   *birchmanni* or *X. birchmanni* x *X. malinche* hybrids outside of the Tlalica site and nearby sites
428   suggests that something is unique about this locality. We predict that the demography of the
429   Tlalica community and its disrupted water quality and chemistry play an important role in this
430   unusual hybridization event. Anthropogenic disturbance via wastewater effluent and landfill
431   leachate could have facilitated hybridization via two mechanisms: 1) by decreasing the
432   abundance of *X. birchmanni* x *X. malinche* hybrids with respect to *X. variatus* and 2) by
433   disrupting sensory cues used in mate choice.
434   *X. birchmanni*, *X. malinche,* and their hybrids are more sensitive to poor water quality
435   than *X. variatus* (Mercado-Silva *et al.*, 2006; personal observation), and *X. variatus* individuals
436   vastly outnumber *X. birchmanni* x *X. malinche* hybrids at the Tlalica site. Notably, nowhere else

437   on the Rio Calnali have we observed such a strong demographic skew toward *X. variatus*. Past
438   research has suggested that female mate preferences can weaken when the density of
439   conspecifics is low and encounter rate with heterospecifics is high, or the search cost of finding a
440   conspecific mate is very high (Cotton *et al.* 2006; Lehmann 2007; Verzijden *et al.* 2011; Stoffer
441   & Uetz 2015; Delclos *et al.* 2020), including in *X. birchmanni*, *X. malinche,* and *X. variatus*
442   (Fisher & Rosenthal, 2010). This raises the possibility that female *X. birchmanni* x *X. malinche*
443   hybrids mated with *X. variatus* males because there were so few Northern swordtail mates
444   available. This hypothesis is further supported by the lack of *X. variatus* mitochondrial ancestry
445   in three-way hybrids (Table S5). In *Xiphophorus*, females tend to have much stronger
446   conspecific mate preferences than males (Rosenthal & Garcia de Leon, 2011) and *X. variatus*
447   females in the Tlalica population have access to many conspecific males.
448   In addition to demography, our results are consistent with a role of anthropogenic shifts
449   in water quality and chemistry in this rare hybridization event. In many *Xiphophorus* species,
450   female mate choice is driven in large part by species-specific olfactory signals (Crapon De
451   Caprona and Ryan 1990; McLennan and Ryan, 1999; Wong *et al.* 2005; Fisher & Rosenthal
452   2006; Rosenthal *et al.* 2011; Verzijden *et al.* 2011). Research has shown that organic and
453   inorganic substances can alter the ability of female *X. birchmanni* to distinguish conspecific from
454   heterospecific males (Fisher, Wong, & Rosenthal, 2006, Powell *et al.* 2022). Levels of several
455   chemicals observed at Tlalica may be sufficient to disrupt olfactory communication and drive
456   hybridization between *X. variatus* and *X. birchmanni* x *X. malinche* hybrids. Elevated dissolved
457   organic matter has been shown to impair chemical and/or visual communication in some fish
458   species at concentrations of ~1 mg/L of humic acid (Hubbard *et al.*, 2002; Mobley *et al.*, 2020),
459   similar to concentrations found in Tlalica and Calnali (Table S6). Ammonia can impair
460   generation of electric impulses in neurons and lead to health effects at low concentrations (1.3-
461   3.5 mg N/L; Ip, Chew, and Randall 2001). Because ammonia was found at concentrations above
462   toxic thresholds at Tlalica (Fig. 2), future research should test hypotheses regarding whether
463   ammonia affects mating behavior. Finally, the copper concentrations detected at Tlalica were
464   similar to the concentrations reported to disturb olfactory perception and olfactory-mediated
465   behaviors in other fish species (~2 μg/L; Sandahl *et al.*, 2007; Morris *et al.*, 2019).
466   Although *Xiphophorus* respond most strongly to olfactory sexual signals, visual cues are
467   also important in mating decisions (Crapon de Caprona and Ryan, 1990; McLennan and Ryan,
468   1999; Fisher *et al.*, 2006b; Verzijden and Rosenthal 2011; Delclos *et al.* 2020). Thus, the
469   increased turbidity of water at Tlalica and nearby sites may also play a role in the breakdown of
470   reproductive barriers (Seehausen et al. 1997). Turbid waters could indirectly facilitate
471   hybridization among *X. variatus* and *X. birchmanni* x *X. malinche* by disturbing the transmission
472   of visual cues involved in species/mate recognition. Testing hypotheses about the impacts of
473   water chemistry on mate choice using both chemical treatments and mate choice trials is an
474   exciting avenue for future investigation (e.g. as in Fisher *et al.*, 2006).
475   What are the consequences of the complex hybridization events that researchers are
476   beginning to uncover? One possible outcome when three species hybridize is the potential for
477   gene flow between two species that would otherwise not contact each other. Such "conduit
478   introgression" has been described in several systems (Heliconius Genome Consortium 2012;
479   Toews *et al.* 2018; Langdon *et al.* 2019; Grant and Grant 2020, Natola *et al.* 2022). While the
480   scenario uncovered at Tlalica is more complex since hybridization is occurring between pure *X.*
481   *variatus* and *X. birchmanni* x *X. malinche* hybrids, the effects on dynamics of gene flow could be
482   similar. Specifically, because *X. malinche* does not overlap with *X. variatus*, admixture with *X.*

483   *birchmanni* x *X. malinche* hybrids could be a route for contemporary gene flow between *X.*
484   *variatus* and *X. malinche*.
485
486   *Conclusions*
487   We present evidence of a deep hybridization event between the platyfish and Northern
488   swordtail clade, involving genetic material from three source species. Given the abundance of
489   first-generation three-way hybrids over multiple sampling seasons and the rarity of second
490   generation or later hybrids in our sample, we predict that there are significant costs in terms of
491   viability or fertility of this distant cross. While contemporary hybrids between these species have
492   not been previously reported, ancient hybridization between them has been inferred (Schumer *et*
493   *al.* 2018). Our data suggest that this unusual hybridization event may be linked to anthropogenic
494   disturbance in the local environment. Disentangling the mechanisms through which
495   anthropogenic disturbance contributes to hybridization in this system in an exciting direction for
496   future work.
497

498   **Figures**

499
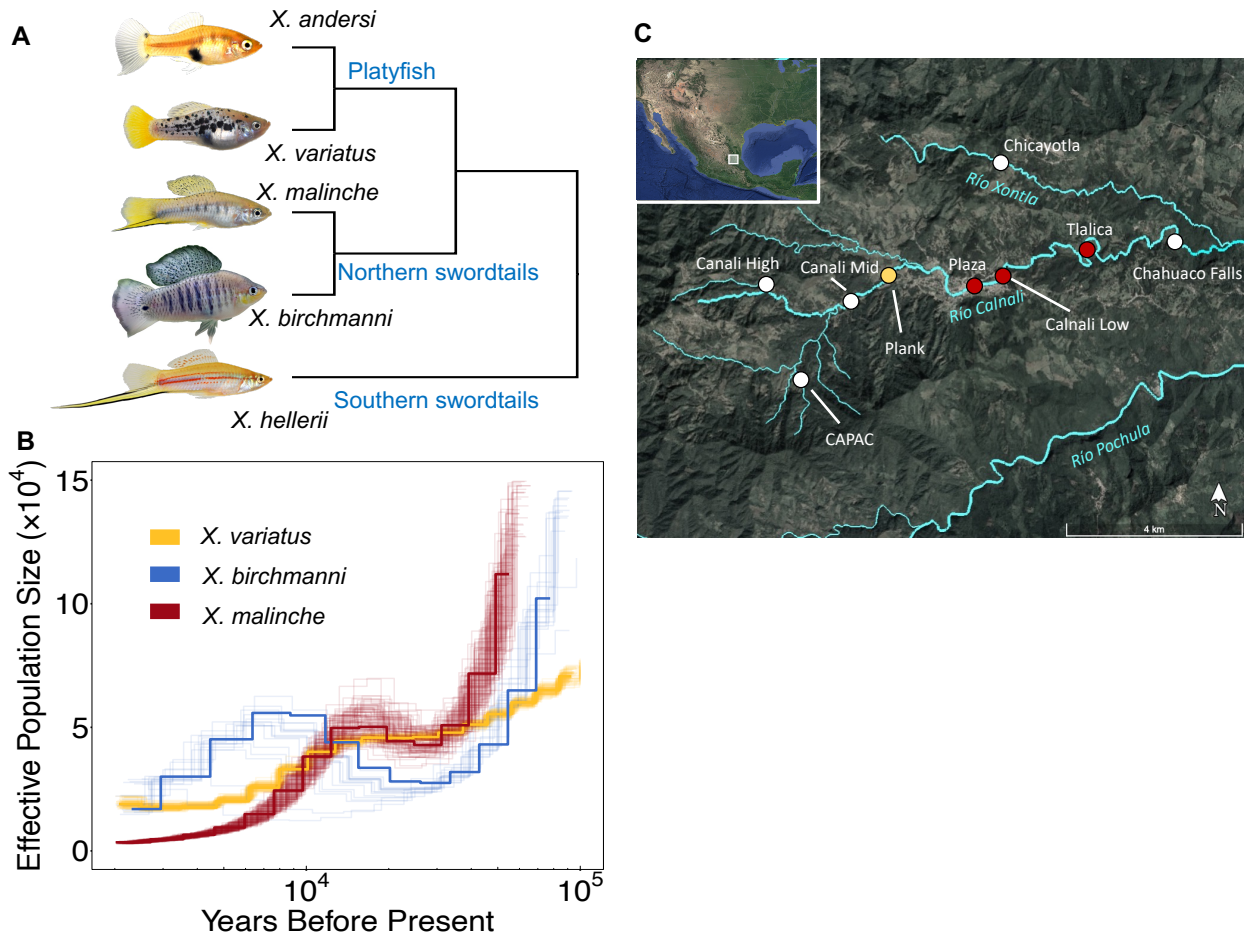
500   **Figure 1**. **A**) Simplified phylogeny of the genus *Xiphophorus*. *X. birchmanni* and *X. malinche*
501   are sister species in the Northern swordtail clade and *X. variatus* belongs to the distantly related
502   platyfish clade. **B**) PSMC results analyzing population history from a single whole-genome
503   sample of *X. variatus* and previously collected data from *X. birchmanni* and *X. malinche* (from
504   Schumer *et al.* 2018). Analysis was conducted with a ρ/θ ratio of 2, generation time of two
505   generations per year, and mutation rate of 3.5×10^{-9}. Faint lines reflect bootstrap resampling of
506   the data. **C**) Map of collection sites along the Río Calnali with the focal sites where three-way
507   hybrids have been collected highlighted in red, and upstream site used for comparison in water
508   quality and chemistry samples highlighted in yellow. Inset shows location of *X. birchmanni* x *X.*
509   *malinche* hybrid populations in Hidalgo, Mexico relative to a map of North and Central America.
510   Images adapted from Google Earth.

511

512
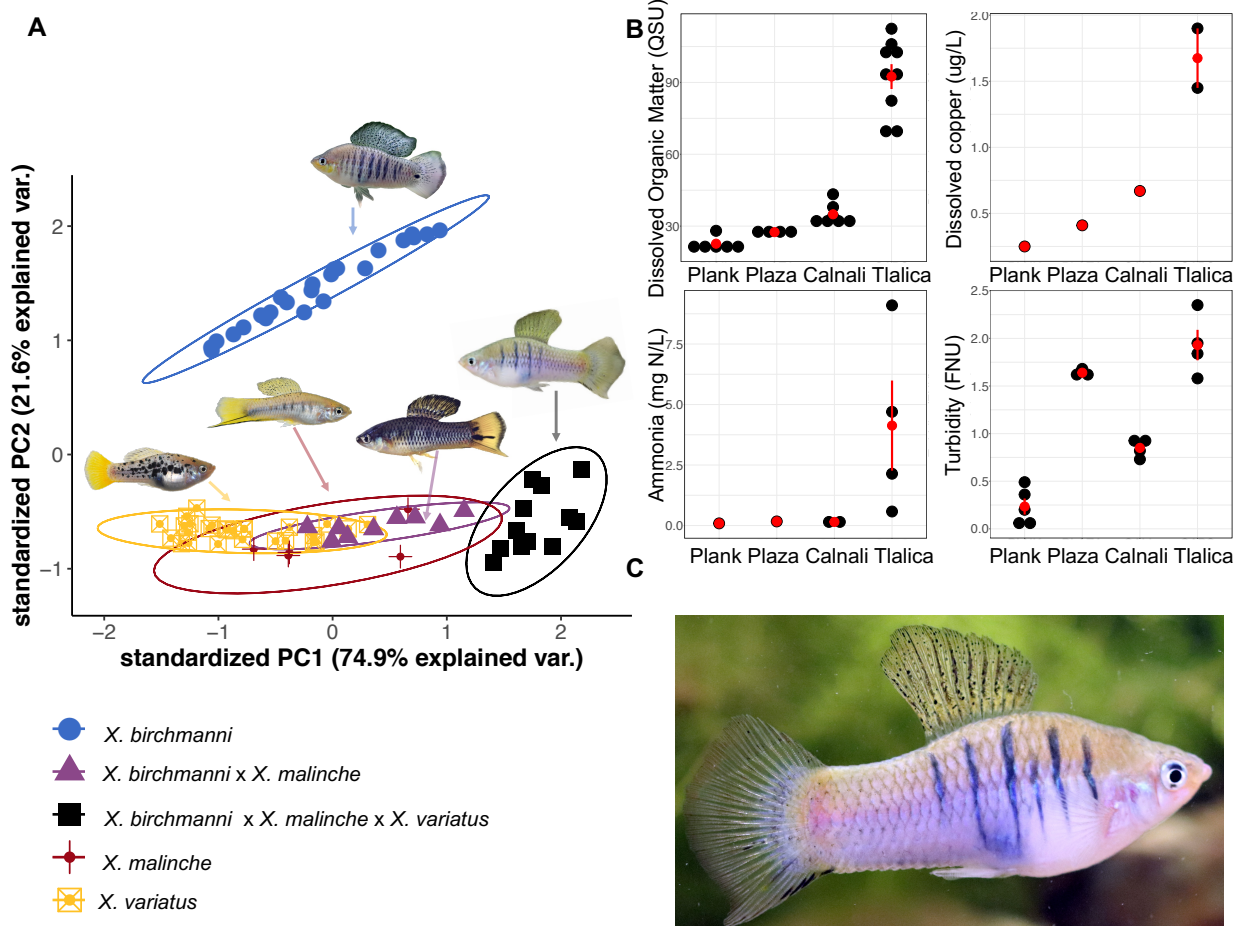
Figure 2. **A**) Principal component analysis of morphology of male *X. malinche, X. birchmanni, X. variatus*, *X. birchmanni* x *X. malinche* hybrids and confirmed three-way hybrids. **B**) Dissolved organic matter (quinine sulfite units - QSU), ammonia (mg N/L), dissolved copper (ug/L), and turbidity (formazin nephelometric units - FNU) levels measured at Plank, Plaza, Calnali Low (Calnali), and Tlalica in May and June of 2022. Black dots represent independent measurements, red dots represent means, and red bars show one standard error of repeated measurements. **C**) Example of a first generation hybrid individual with 50% *X. variatus* ancestry, 23.8% *X. birchmanni* ancestry, and 26.2% *X. malinche* ancestry. This individual has a short sword, a trait which is unique to *X. malinche*, and large dorsal fin characteristic of *X. birchmanni*, and an overall body shape and vertical barring characteristics of *X. variatus*.
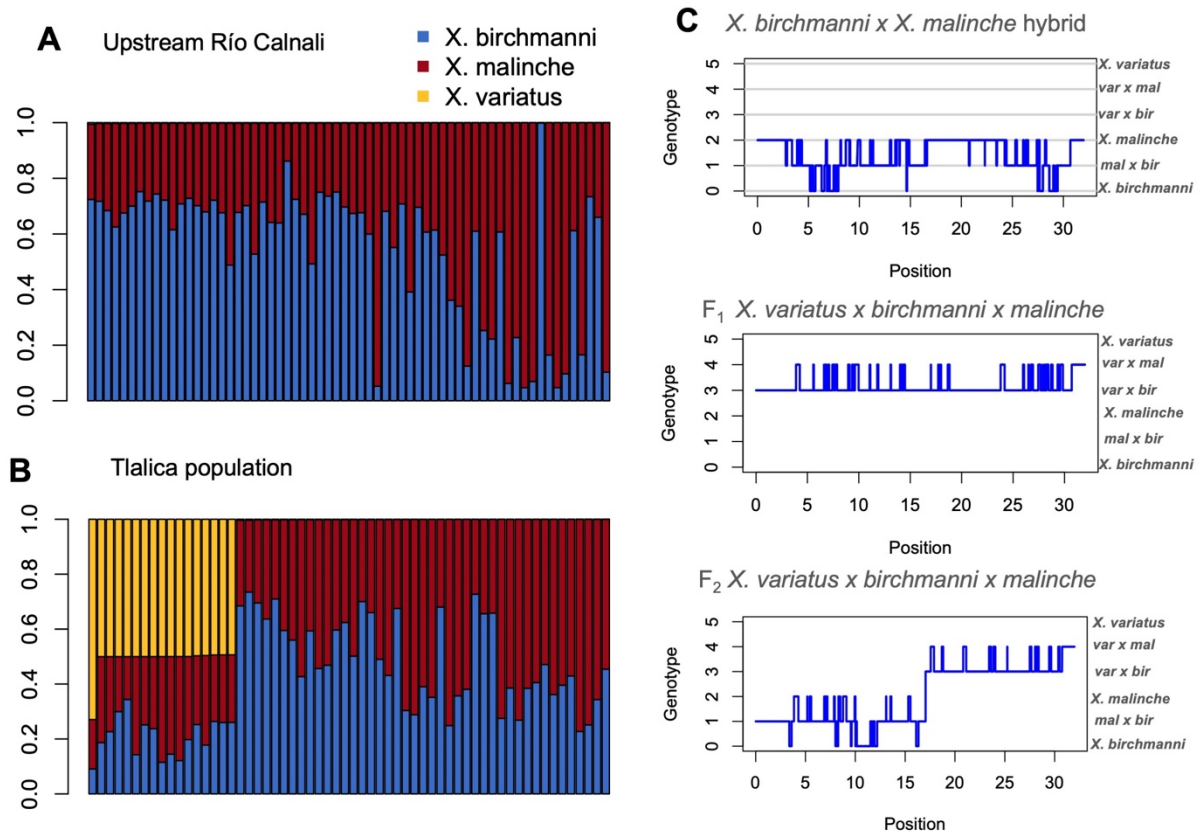
Figure 3. A) Distribution of genome-wide ancestry for individuals collected upstream of the Tlalica site on the Río Calnali (N=553, downsampled for visualization to 64). Stacked plots show the estimated proportion of each individual's genome derived from *X. variatus* (yellow), *X. birchmanni* (blue), and *X. malinche* (red) based on local ancestry inference with *ancestryinfer*. Individuals sampled from upstream sites have near-zero introgression from *X. variatus*. **B**) Distribution of genome-wide ancestry for individuals with suspected swordtail ancestry collected at the Tlalica site on the Río Calnali (N=64). Stacked plots show the estimated proportion of each individual's genome derived from *X. variatus* (yellow), *X. birchmanni* (blue), and *X. malinche* (red). While some *X. birchmanni* x *X. malinche* hybrid individuals sampled from Tlalica lack *X. variatus* ancestry, a substantial proportion derive some of their genome from *X. variatus*. **C**) Local ancestry inferred along chromosome 1 for individuals of different hybrid types. The genotype on the y-axis corresponds to the ancestry class at that marker: 0 – homozygous *X. birchmanni*, 1 – heterozygous *X. birchmanni* x *X. malinche*, 2 – homozygous *X. malinche*, 3 – heterozygous *X. birchmanni* x *X. variatus*, 4 – heterozygous *X. malinche* x *X. variatus*, 5 – homozygous *X. birchmanni*. In the top plot, a typical *X. birchmanni* x *X. malinche* hybrid from the Río Calnali is shown. In the middle plot, a first generation hybrid between an *X. birchmanni* x *X. malinche* mother and *X. variatus* father is shown (parental source populations inferred based on mitochondrial ancestry, Table S5). Note that across the entire chromosome this individual is either heterozygous *X. birchmanni* x *X. variatus* or heterozygous *X. malinche* x *X. variatus* (genotype classes 3 or 4). In the bottom plot, ancestry for a "backcrossed" three-way

547   hybrid individual is shown. This individual is inferred to be the offspring of a first generation

548   three-way hybrid and a *X. birchmanni* x *X. malinche* mother. Note that this individual is

549   heterozygous for *X. variatus* ancestry over only approximately half of its chromosome.

Acknowledgements

Author contributions

S.M. Banerjee, D.L. Powell, and M. Schumer conceived of this project. S.M. Banerjee, D.L. Powell, B.M. Moran, T. Gunn, and W.F. Ramírez-Duarte collected data, S.M. Banerjee, D.L. Powell, B.M. Moran, Q. Langdon, W.F. Ramírez-Duarte, and M. Schumer analyzed data. M. Schumer adapted *ancestryinfer* for three-way local ancestry inference. M. Schumer and C. Rochman oversaw the project. All authors wrote the manuscript.

Data availability

All raw sequencing data for this project will be deposited on NCBI's SRA. All water quality and chemistry data and ancestry calls will be deposited on Dryad. Computational pipelines and analysis scripts are available at https://github.com/Schumerlab.

References

Armstrong, J., Hickey, G., et al. (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature 587, 246–251.

Borowsky, R. (1981). Tailspots of *Xiphophorus* and the evolution of conspicuous polymorphism. Evolution 35(2), 345–358.

Chen, F. & Li, W. (2001). Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am J Hum Genet 68:2, 444-456.

Chunco, A. (2014). Hybridization in a warmer world. Ecol Evol. 2014 May;4(10):2019-31.

Corbett-Detig, R. & Nielsen, R. (2017). A Hidden Markov Model approach for simultaneously estimating local ancestry and admixture time using next generation sequence data in samples of arbitrary ploidy. PLOS Genet, 13, e1006529.

Cotton, S., Small, J., Pomiankowski, A. (2006). Selection and condition-dependent mate preferences. Curr Biol 16:PR755-R765.

Crapon De Caprona, M. & Ryan, M. J. (1990). Conspecific mate recognition in swordtails, *Xiphophorus nigrensis* and *X. pygmaeus* (Poeciliidae): olfactory and visual cues. Anim Behav 39, 290-296.

Cui, R., Schumer, M., Kruesi, K., Walter, R., Andolfatto, P. & Rosenthal, G. (2013). Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. Evolution, 67, 2166–2179.

Culumber, Z. & Rosenthal, G. (2013). Mating preferences do not maintain the tailspot polymorphism in the platyfish, *Xiphophorus variatus*, Behav Ecol, 24: 1286–1291.

Culumber, Z. (2014). Pigmentation in *Xiphophorus*: An emerging system in ecological and evolutionary genetics. Zebrafish 11, 57–70.

Danecek P, Bonfield JK, et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience 10(2):giab008

Delclos, P., Forero, S., Rosenthal, G. (2020). Divergent neurogenomic responses shape social learning of both personality and mate preference. J Exp Biol 223(6): jeb220707.

Fisher, H. & Rosenthal, G. (2006) Female swordtail fish use chemical cues to select well-fed mates. Anim Behav 72:721-725.

Fisher, H.S., Wong, B.B.M. & Rosenthal, G.G. (2006). Alteration of the chemical environment disrupts communication in a freshwater fish. Proc R Soc London Ser B, 273, 1187–1193.

Fisher, H.S. & Rosenthal, G.G. (2010). Relative abundance of *Xiphophorus* fishes and its effect on sexual communication. Ethology, 116: 32-38.

Grant, P. & Grant, R. (2020). Triad hybridization via a conduit species. Proc Nat Acad Sci 117(14): 7888-7896.

The Heliconius Genome Consortium. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487, 94–98.

Hubbard, P.C., Barata, E.N., Canario, A.V.M. (2002). Possible disruption of pheromonal communication by humic acid in the goldfish, *Carassius auratus*. Aquat Toxicol 60, 169–183.

Ip, Y.K., Chew, S.F., Randall, D.J (2001). Ammonia toxicity, tolerance, and excretion. Fish Physiol 20: 109–148.

Kallman, K. & Kazianis, S. (2006). The genus *Xiphophorus* in Mexico and Central America. Zebrafish. 3:271-285.

642 Kelly, B., Whiteley, A., Tallmon, D. (2010). The Arctic melting pot. Nature 468, 891.

643 Langdon, Q.K., Peris, D., Kyle, B., Hittinger, C.T. (2018). sppIDer: a species identification tool
644 to investigate hybrid genomes with high-throughput sequencing. Mol Biol Evol, 35, 2835–2849.

645 Langdon, Q.K., Peris, D., et al. (2019). Fermentation innovation through complex hybridization
646 of wild and domesticated yeasts. Nat Ecol Evol 3, 1576–1586.

647 Lehmann (2006). Density-dependent plasticity of sequential mate choice in a bushcricket
648 (Orthoptera : Tettigoniidae). Austral J Zool 55(2): 123-130.

649 Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
650 Transform. Bioinformatics, 25:1754-60.

651 Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-
652 genome sequences. Nature 475, 493–496.

653 Mclennan, D. A. & Ryan, M. J. (1999). Interspecific recognition and discrimination based upon
654 olfactory cues in northern swordtails. Evolution 53, 880-888.

655 Mercado-Silva, N., Lyons, J., et al. (2006). Long-term changes in the fish assemblage of the Laja
656 River, Guanajuato, central Mexico. Aquatic Conserv: Mar. Freshw. Ecosyst., 16: 533-546.

657 Mobley, R.B., Weigel, E.G., Boughman, J.W. (2020). Does humic acid alter visually and
658 chemically guided foraging in stickleback fish? Anim Cogn 23, 101–108.

659 Moran, B.M., Payne, C.Y., et al. (2021). A lethal genetic incompatibility between naturally
660 hybridizing species in mitochondrial complex I. bioRxiv. doi:
661 https://doi.org/10.1101/2021.07.13.452279

662 Moran, B.M., Payne, C., Langdon, Q., Powell, D.L., Brandvain, Y., Schumer, M. (2021). The
663 genomic consequences of hybridization. eLife, 10, e69016.

664 Morris, J.M., Brinkman, S.F., Takeshita, R., McFadden, A.K., Carney, M.W., Lipton, J. (2019).
665 Copper toxicity in Bristol Bay headwaters: part 2—olfactory inhibition in low-hardness water.
666 Environ Toxicol Chem 38, 198–209.

667 Natola, L., Seneviratne, S. S., Irwin, D. (2022). Population genomics of an emergent tri-species
668 hybrid zone. Mol Ecol 00, 1– 12.

669 Ottenburghs, J. (2020). Ghost introgression: spooky gene flow in the distant past. BioEssays, 42,
670 2000012.

671 Oziolor, E.M., Reid, N.M., et al. (2019). Adaptive introgression enables evolutionary rescue
672 from extreme environmental pollution. Science, 364, 455–457.

673 Pampoulie, C., Gíslason, D., et al. (2021). Evidence of unidirectional hybridization and second-
674 generation adult hybrid between the two largest animals on Earth, the fin and blue whales. Evol
675 Appl 14: 314– 321.

676 Payne, C., Bovio, R., et al. (2022). Genomic insights into variation in thermotolerance between
677 hybridizing swordtail fishes. Mol Ecol.

678 Powell, D.L., García-Olazábal, et al. (2020). Natural hybridization reveals incompatible alleles
679 that cause melanoma in swordtail fish. Science, 368, 731–736.

680 Powell, D.L., Moran, B.M., et al. (2021). Two new hybrid populations expand the swordtail
681 hybridization model system. Evolution, 75, 2524–2539.

682 Rauchenberger, M., Kallman, K.D., Morizot, D.C. (1990). Monophyly and geography of the Río
683 Pánuco Basin swordtails (genus *Xiphophorus*) with descriptions of four new species. Am Mus
684 Nov no. 2975.

685 Rosenthal, G.G., Rosa Reyna, X.F., et al. (2003). Dissolution of sexual signal complexes in a
686 hybrid zone between the swordtails *Xiphophorus birchmanni* and *Xiphophorus malinche*
687 (Poeciliidae). Copeia 299–307.

688 Rosenthal, G.G. & Garcia de Leon, F.J. (2011). Speciation and hybridization. In: (*Ecology and*
689 *evolution of poeciliid fishes),* 109–119.
690 Rosenthal, G., Fitzsimmons, J.N., Woods, K.U., Gerlach, G., Fisher, H.S. (2011). Tactical
691 release of a sexually-selected pheromone in a swordtail fish. PLoS ONE 6(2): e16994.
692 Sandahl, J.F., Baldwin, D.H., Jenkins, J.J., Scholz, N.L. (2007). A sensory system at the interface
693 between urban stormwater runoff and salmon survival. Environl Sci Technol 41,
694 2998–3004.
695 Schneider, C. A., Rasband, W. S., Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of
696 Image Analysis. Nat Methods, 9(7), 671–675.
697 Schumer, M., Cui, R., Powell, D.L., Dresner, R., Rosenthal, G.G., Andolfatto, P. (2014). High-
698 resolution mapping reveals hundreds of genetic incompatibilities in hybridizing fish species.
699 eLife, 3, e02535.
700 Schumer, M., Powell, D.L., et al. (2017). Assortative mating and persistent reproductive
701 isolation in hybrids. Proc Nat Acad Sci, 114, 10936.
702 Schumer, M., Xu, C., et al. (2018). Natural selection interacts with recombination to shape the
703 evolution of hybrid genomes. Science, 360, 656.
704 Schumer, M., Powell, D.L., Corbett-Detig, R. (2020). Versatile simulations of admixture and
705 accurate local ancestry inference with mixnmatch and ancestryinfer. Mol Ecol Res, 20, 1141–
706 1151.
707 Seehausen, O., Alphen, J., Witte, F. (1997). Cichlid fish diversity threatened by eutrophication
708 that curbs sexual selection. Science, 277, 1808–1811.
709 Stoffer, B., Uetz, G.W. (2015). The effects of social experience with varying male availability on
710 female mate preferences in a wolf spider. Behav Ecol Sociobiol 69, 927–937.
711 Toews, D., Streby, H., Burket, L., Taylor, S. (2018). A wood-warbler produced through both
712 interspecific and intergeneric hybridization. Biol Lett.
713 Vallejo-Marín, M. Hiscock, S.J. (2016). Hybridization and hybrid speciation under global
714 change. New Phytol, 211: 1170-1187.
715 Van der Auwera, G.A. & O'Connor, B.D. (2020). Genomics in the cloud: Using Docker, GATK,
716 and WDL in Terra (1st Edition). O'Reilly Media.
717 Verzijden, M. & Rosenthal, G. (2011). Effects of sensory modality on learned mate preferences
718 in female swordtails. Anim Behav 82(3): 557-562.
719 Willis, P.M., Ryan, M.J., Rosenthal, G.G. (2011). Encounter rates with conspecific males
720 influence female mate choice in a naturally hybridizing fish. Behav Ecol, 22, 1234–1240.
721 Wong, B., Fisher, H., Rosenthal, G. (2005). Species recognition by male swordtails via chemical
722 cues. Behav Ecol 16(4): 818–822
723 Wu, J., Liu, Y., Zhao, Y. (2021). Systematic review on local ancestor inference from a
724 mathematical and algorithmic perspective. Front Genet.