Integrative analyses of convergent adaptation in sympatric extremophile fishes

Ryan Greenway[1], Rishi De-Kayne[2], Anthony P. Brown[3,4], Henry Camarillo[1,5], Cassandra Delich[1], Kerry L. McGowan[3], Joel Nelson[3], Lenin Arias-Rodriguez[6], Joanna L. Kelley[2‡], and Michael Tobler[7,8,9‡]

[1] Division of Biology, Kansas State University, Manhattan, KS, USA

[2] Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, USA

[3] School of Biological Sciences, Washington State University, Pullman, WA, USA

[4] Current address: California National Primate Research Center, University of California, Davis, Davis, CA, USA

[5] Current address: Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

[6] División Académica de Ciencias Biológicas, Universidad Juárez Autónoma de Tabasco, Villahermosa, Tabasco, Mexico

[7] Department of Biology, University of Missouri—St. Louis, St. Louis, MO 63121, USA

[8] Whitney R. Harris World Ecology Center, University of Missouri—St. Louis, St. Louis, MO 63121, USA

[9] WildCare Institute, St. Louis Zoo, St. Louis, MO 63110, USA

[‡] Corresponding authors: jokelley@ucsc.edu; tobler@umsl.edu

Summary

The evolution of independent lineages along replicated environmental gradients frequently results in convergent adaptation, yet the degree to which convergence is present across multiple levels of biological organization is often unclear. Additionally, inherent biases associated with shared ancestry and variation in selective regimes across geographic replicates often pose challenges for confidently identifying patterns of convergence. We investigated a system in which three species of poeciliid fishes sympatrically occur in a toxic spring rich in hydrogen sulfide (H_2S) and an adjacent nonsulfidic stream to examine patterns of adaptive evolution across levels of biological organization. We found convergence in morphological and physiological traits and genome-wide patterns of gene expression among all three species. In addition, there were shared signatures of selection on genes encoding H_2S toxicity targets in the mitochondrial genomes of each species. However, analyses of nuclear genomes revealed neither evidence for substantial genomic islands of divergence around genes involved in H_2S toxicity and detoxification nor substantial congruence of strongly differentiated regions across population pairs. These non-convergent, heterogenous patterns of genomic divergence may indicate that sulfide tolerance is highly polygenic, with shared allele frequency shifts present at many loci with small effects along the genome. Alternatively, H_2S tolerance may involve substantial genetic redundancy, with non-convergent lineage-specific variation at multiple loci along the genome underpinning similar changes in phenotypes and gene expression. Overall, we demonstrate variability in the extent of convergence across organizational levels and highlight the challenges of linking patterns of convergence across scales.

Keywords: Adaptive evolution | Convergent evolution | Ecological genomics | Hydrogen sulfide | Local adaptation | Mitochondria | Poeciliidae | Population genomics

Introduction

Evolutionarily independent lineages often evolve similar solutions when facing similar ecological conditions. Although such patterns of convergence may provide evidence for the repeatability and predictability of evolution by natural selection,[1–3] it is often challenging to determine the extent of convergence across biological levels of organization, and ultimately, whether convergent genomic shifts underpin convergence at other scales.[4] Convergent evolution is frequently studied in closely related, but geographically disjunct, population pairs exposed to similar sources of divergent selection, providing spatially replicated instances of local adaptation within a clade.[1] Investigations into convergent evolution in such systems have provided valuable insights across a wide variety of animal and plant systems by highlighting instances of evolutionary convergence at the level of single genes[5,6] and more polygenic signals of convergence.[7,8] However, studies that focus on closely related lineages likely biases our understanding of general patterns of convergence, because shared evolutionary responses may be the result of selection on standing genetic variation and constraints associated with shared genomic architecture.[5,9,10] Similarly, the focus on geographically disjunct populations exposed to seemingly similar selective regimes has the potential to ignore the effects of unquantified environmental variation, which often leads to idiosyncratic evolutionary responses.[11,12] Studying distantly related taxa that are experiencing a shared selective environment in sympatry eliminates these confounding influences, providing the opportunity to investigate convergence in adaptive strategies and determine the predictability and repeatability of evolution.[13–15]

We took advantage of a unique system to explore shared patterns of adaptive evolution among three sympatric lineages of livebearing fishes (Poeciliidae). In Mexico, populations of *Poecilia mexicana*, *Pseudoxiphophorus bimaculatus*, and *Xiphophorus hellerii* coexist both within the La Gloria sulfide spring complex, which is characterized by high concentrations of naturally occurring hydrogen sulfide (H_2S), and in adjacent nonsulfidic streams (Figure 1A, B). The presence and absence of H_2S creates strong divergent selection across habitat types because of its toxic properties. H_2S binds to cytochrome c oxidase (COX) of the electron transport chain and interrupts oxidative phosphorylation (OxPhos) in mitochondria, ultimately stopping aerobic ATP production necessary for the maintenance of cellular function.[16,17] At the molecular level, the clear biochemical and physiological consequences of H_2S predict adaptive modification of genes associated with OxPhos and enzymatic H_2S detoxification, which is mediated by the mitochondrial sulfide:quinone oxidoreductase (SQR) pathway.[16] Besides the presence and absence of H_2S, the two aquatic habitats

differ in other physical and chemical attributes, with the sulfide spring having lower dissolved oxygen concentrations and pH, but higher specific conductivity (Figure 1C).

The La Gloria system is particularly interesting because it allows for the juxtaposition of evolutionary responses to H_2S in distantly related species that occur in sympatry with evolutionary responses previously documented in geographically disparate populations of sulfide spring fishes within the family Poeciliidae.[18] Throughout the Americas, highly endemic poeciliids locally-adapted to sulfide springs have evolved from ancestral populations in adjacent freshwater habitats.[18–20] Evolutionary responses to the toxic conditions have primarily been studied in geographically replicated populations of *Poecilia mexicana*, which exhibit convergence in morphology, physiology, reproductive life histories, gene expression, and the establishment of reproductive isolation as a byproduct of adaptation via ecological speciation.[18] While some convergence appears to result from *de novo* modification of the same genes or biochemical pathways (*e.g.*, in OxPhos), there is also evidence for selection on standing genetic variation (*e.g.*, in the SQR pathway).[21–26] As in other systems, the focus on convergence across geographically replicated lineages within *P. mexicana* may have biased the interpretation of the repeatability and predictability of adaptive evolution in sulfide springs. Indeed, a recent broad-scale study of adaptation to sulfide springs across the Poeciliidae found evidence for convergence in gene expression profiles of ten evolutionarily independent sulfide-spring lineages, but only limited evidence for repeatable patterns of molecular evolution outside of a few mitochondrially-encoded OxPhos genes.[25] The degree to which this breakdown of molecular convergence at a broader phylogenetic scale is related to differences in the selective regimes among geographically disparate sulfide springs or attributable to lineage-specific evolutionary solutions to a shared source of selection remains unresolved, though evidence suggests that environmental similarity drives the degree of convergence at the transcriptomic level.[27]

In this study, we aimed to identify signatures of convergence associated with the adaptation of three independent lineages to the same H_2S-rich spring. This approach helped reduce the confounding effects of shared ancestral standing genetic variation and allowed us to stringently test for shared adaptive responses across biological levels of organization in different species. Our results also provided the opportunity to identify possible links between different aspects of convergence across scales, including whether convergence in morphology and gene expression are explicitly underpinned by convergent genomic variation. To address this overarching hypothesis, we (i)

112 identified signatures of convergent evolution in phenotypic traits that have previously been
113 associated with adaptation to sulfide spring environments (body shape, H_2S tolerance, and gene
114 expression), (ii) established evidence for local adaptation between sulfidic and nonsulfidic
115 populations of each species, and (iii) identified convergent patterns of genomic differentiation
116 among population pairs in sulfidic and nonsulfidic habitats. Overall, we found evidence for
117 convergence in adaptation across multiple levels of organization, but convergence at the genomic
118 level was limited and, outside of the mitochondrial genome, not associated with genomic regions
119 predicted to respond to selection from H_2S.
120
121 **Results**
122 ***Phenotypic variation among habitats***
123 Mechanisms of adaptation to H_2S-rich environments have been studied extensively in poeciliid
124 fishes, including the biochemical and physiological modifications that directly mitigate H_2S toxicity[25]
125 and modifications of other traits that are shaped by sources of selection that are correlated with the
126 presence of H_2S (*e.g.,* hypoxia, increased salinity, changes in trophic resource use, competition, and
127 predation.[18,28] We quantified variation of natural populations in three complex traits that typically
128 show adaptive divergence upon sulfide spring colonization: body shape, H_2S tolerance, and gene
129 expression. We found that sulfide spring populations of all three focal species exhibit convergent
130 patterns of trait divergence from adjacent nonsulfidic populations in response to colonization of
131 sulfide springs.
132
133 Geometric morphometric analyses of wild-caught, adult fish revealed that body shape significantly
134 differed between sulfidic and nonsulfidic populations of all three lineages (Procrustes ANOVA;
135 species × habitat: $P < 0.001$, $Z = 7.565$, Table S1). Furthermore, there was a significant signal of
136 morphological convergence across the sulfide spring populations of the three species (habitat: $P =$
137 0.018, $Z = 2.093$, Table S1). Investigation of thin-plate spline transformation grids visualizing shape
138 differences for both the "species × habitat" and "habitat" terms indicated that the shared aspects of
139 variation in body shape between sulfidic and nonsulfidic populations of each species involved
140 divergence in head size, with fish from the sulfide spring exhibiting larger heads than those from the
141 nonsulfidic habitat (Figure 2A).
142

143 Survival analyses indicated that wild-caught, adult fish from all three sulfide spring populations
144 exhibited an increased ability to tolerate acute H_2S exposure relative to nonsulfidic populations
145 (Figure 2B; Cox regression; habitat: $P = 0.034$, Figure S1). While H_2S tolerance was clearly affected
146 by habitat of origin, differences in tolerance between sulfidic and nonsulfidic populations also varied
147 among the three species, with sulfidic *P. mexicana* having a higher H_2S tolerance than both
148 nonsulfidic *Poecilia* and sulfidic populations of the other species (Figure S1). This observation
149 matched microhabitat uses by the three species within the sulfide spring complex; *P. mexicana* was
150 the only species inhabiting immediate spring outlets with the highest H_2S concentrations, and the
151 other two species were more common in areas with more oxygen-rich water (Figure S2).
152
153 Analysis of genome-wide gene expression in gill tissues from wild-caught sulfidic and nonsulfidic
154 fish revealed an overall phylogenetic signal in gene expression variation, with individuals clustering
155 by species and, secondarily, by habitat type (Figure 3A). There were many differentially expressed
156 genes (negative binomial regression, FDR < 0.05) between the sulfidic and nonsulfidic populations
157 of each species, representing 13.4% of all measured genes for *Pseudoxiphophorus*, 20.4% for *Poecilia*,
158 and 20.9% for *Xiphophorus* (Figure 3B). While most gene expression differences were species-specific
159 (Table S2), 245 genes (1.3% of all analyzed genes; 6.3–9.8% of differentially expressed genes in each
160 species) exhibited convergent expression changes across all species (153 upregulated, 92
161 downregulated; Figure 3B, Table S2), significantly more than predicted by chance (*SuperExactTest*[29]: P
162 < 0.0001). As predicted by the biochemical and physiological effects of H_2S, analysis of Gene
163 Ontology (GO) IDs associated with differentially expressed genes revealed that shared upregulated
164 genes were significantly enriched ($FDR < 0.05$) for associations with mitochondria and biological
165 processes involved in H_2S toxicity and detoxification, including aerobic respiration and the electron
166 transport chain (OxPhos), enzymatic H_2S detoxification, and the processing and transport of sulfur
167 compounds (Table S3). Notably, we found consistent upregulation of genes associated with the
168 primary target of H_2S toxicity (COX) and the primary H_2S detoxification pathway (SQR pathway;
169 Figure 3C–E). There was no evidence for significant enrichment of any GO terms in downregulated
170 genes (Table S3).
171
172 ***Population Structure***
173 As phenotypic divergence between sulfidic and nonsulfidic populations of each species was
174 measured using wild-caught fish, we cannot rule out the possibility that both adaptation and

175 phenotypic plasticity contribute in some part to phenotypic responses to the colonization of the
176 sulfidic habitat. However, since divergent selection on key ecological traits generally reduces gene
177 flow between populations, local adaptation and phenotypic plasticity are predicted to have different
178 population genomic signatures. When plasticity alone underpins phenotypic differentiation across
179 environmental contrasts, we expect to see little genome-wide population structure, as gene flow
180 between habitats prevents genomic differentiation from establishing. In contrast, where divergent
181 selection has resulted in the establishment of barriers to gene flow, we expect to find lower rates of
182 gene flow between populations and signals of population structure, even though the La Gloria
183 sulfide spring complex is spatially restricted (~200 m in length), flows directly into the adjacent
184 nonsulfidic stream (Figure 1A), and there are no physical barriers preventing fish movement
185 between habitat types. Additionally, previous investigations into the signatures of gene expression
186 associated with adaptation of *P. mexicana* to sulfide springs have provided evidence of heritable
187 variation in gene expression,[30,31] which highlights the potential for a component of adaption to be
188 underpinned by genomic variation rather than plasticity alone. To test for genetic signatures of local
189 adaptation, we sampled 20 individuals per species in each habitat type and conducted population
190 genomic analyses based on whole-genome resequencing (average sequence depth per individual: 3.8
191 ± 0.4; Table S4) to infer population structure between the sulfidic and nonsulfidic populations.
192 Maximum-likelihood-based estimates of admixture proportions with two inferred clusters ($k = 2$)
193 unambiguously clustered individuals by their habitat type of origin and only detected low levels of
194 admixture between sulfidic and nonsulfidic populations within each species (Figure 4A), indicating a
195 substantial reduction in gene flow between the two habitats. Principal component analysis (PCA)
196 corroborated these results, consistently separating populations by habitat type along the primary axis
197 of genetic variance in each species (17.5–53.5% of genomic variation explained by PC1; Figure S3).
198 Both structure analysis and PCA identified two *P. mexicana* individuals that appear to be early
199 generation hybrids (F1 and a possible F2 backcross), while all other individuals of all species exhibit
200 clear assignment to their habitat of origin with limited, asymmetric gene flow from sulfidic into
201 nonsulfidic populations.
202
203 We also conducted *in situ* reciprocal translocation experiments of wild-caught, adult fish, which
204 explicitly tested for local adaptation[32] and natural selection against migrants.[33] We found strong,
205 habitat-specific differences in the survival of sulfidic and nonsulfidic populations of each lineage
206 (Figure 4B; binomial GLMM; habitat of origin × testing habitat: $P < 0.001$; Table S5). Fish tested in

207 their own habitat had significantly higher survival rates (88–100%) than those tested in the opposite
208 habitat type, with the lowest survival observed for fish moved from the nonsulfidic into the sulfidic
209 habitat (0–21%). Fish moved from the sulfidic into the nonsulfidic habitat had comparatively high
210 survival (68–84%)—though significantly lower than resident nonsulfidic fish—ultimately supporting
211 the observed asymmetry of gene flow between habitat types. Overall, this experiment supported the
212 inference of local adaptation and that natural selection against migrants likely serves as a strong
213 mechanism generating population structure between sulfidic and nonsulfidic populations.
214
215 ***Landscape of Genomic Divergence***
216 The explicit biochemical and physiological effects of H_2S—along with the convergent changes at the
217 transcriptome level—predict clear targets of selection at the genomic level. We first explored
218 potential convergent changes in the mitochondrial genome, which includes genes for three subunits
219 of H_2S' primary toxicity target (COX). Mitochondrial haplotype networks revealed that sulfidic
220 populations harbor mitochondrial lineages that are highly divergent from those in the nonsulfidic
221 populations (Figure 5A). Consistent with the asymmetric patterns of gene flow, we found sulfidic
222 mitochondrial haplotypes in nonsulfidic habitats, especially in *P. mexicana* and *X. hellerii*, while
223 nonsulfidic mitochondrial haplotypes were not found in sulfidic populations (except for one
224 apparent first-generation hybrid in *P. mexicana* described above). Using analyses of molecular
225 evolution, we detected evidence for positive selection (elevated nonsynonymous to synonymous
226 substitution rates, ω; likelihood-ratio test, $P < 0.01$; Table S6) acting on five mitochondrial genes in
227 all three sulfidic populations, including two COX subunits (COX1 and COX3) and genes encoding
228 subunits of OxPhos complexes I (ND2 and ND3) and III (CYTB), corroborating previous analyses
229 of sulfide spring fishes.[21,25]
230
231 Unlike in the mitochondrial genome, recombination in the nuclear genome can mediate highly
232 heterogeneous divergence during local adaptation. When few genomic regions are directly involved
233 in adaptation, these regions often exhibit disproportionally high divergence (genomic islands of
234 divergence) against a background of regions of low divergence that is homogenized by gene flow.[34,35]
235 If the observed convergent morphological and transcriptomic differentiation had a convergent
236 genomic basis across lineages, we would predict overlapping genomic islands of divergence between
237 sulfidic and nonsulfidic populations across the three species, including regions with mitochondrially-
238 functioning genes that are involved in H_2S toxicity and detoxification. Considering the evidence for

239 positive selection on mitochondrially encoded genes, we also expected to detect evidence for mito-
240 nuclear coevolution,[36] with disproportionally high divergence in the nuclear genes encoding subunits
241 of OxPhos complexes I, III, and especially IV. To determine if selection on the same genomic
242 regions underlies convergence to sulfide spring environments, we used the whole-genome
243 resequencing data described above to characterize the landscape of genomic divergence in each
244 population pair using a combination of relative (F_{ST}) and absolute (d_{XY}) divergence metrics in non-
245 overlapping 5-kb windows.[37,38] Global estimates of relative divergence varied considerably among
246 species (mean F_{ST} = 0.23 for *Poecilia*, 0.06 for *Pseudoxiphophorus*, and 0.03 for *Xiphophorus*; Figure S4),
247 while absolute divergence across the genome was similar (mean d_{XY} = 0.003 for *Poecilia*, 0.004 for
248 *Pseudoxiphophorus*, and 0.003 for *Xiphophorus*; Figure S5).
249
250 Using a composite analysis that simultaneously considered relative and absolute divergence,[39] we
251 observed heterogeneous divergence across much of the genome within and among population pairs,
252 rather than locally pronounced and shared genomic islands of divergence (Figure 5B; Table S7). By
253 identifying the most differentiated regions between sulfidic and nonsulfidic individuals within each
254 species (top 1% most differentiated windows; 1,341–1,366 5-kb windows per species), we found that
255 regions of elevated divergence were significantly enriched for genes involved in oxygen transport
256 (*Pseudoxiphophorus*) and immune function (*Xiphophorus* and *Poecilia*), including hemoglobins and major
257 histocompatibility complex genes (GO enrichment analysis: *FDR* < 0.05; although in *Poecilia* this
258 enrichment was non-significant following FDR correction; Tables S8). To test for convergent
259 patterns of differentiation, we identified regions that were amongst the most differentiated windows
260 across all three species (top 1% most differentiated in each of the three comparisons to account for
261 variation in average genome-wide differentiation across the three comparisons; Figure 5B). Thirty-
262 four strongly differentiated regions—containing 28 genes—were shared across all three species
263 (significantly more than expected by chance; *SuperExactTest*, $P < 0.0001$), and 102–156 of these
264 regions were shared between pairs of the three species (Table S7). There was no significant
265 enrichment of GO terms in strongly differentiated regions shared among all lineages, and we found
266 no evidence for genes with known roles in H_2S detoxification or processing in regions of elevated
267 divergence within species (Tables S8). We did observe some overlap of shared differentiated regions
268 and differentially expressed genes within each species (150–198 genes), but not more than would be
269 expected by chance (Fisher's Exact Test, P = 0.24–1.0).
270

271 We also carried out site-frequency-spectrum based scans for selective sweeps[40] in each sulfidic
272 population and compared putative sweep regions across the three sulfidic and nonsulfidic contrasts.
273 We found no shared selective sweep outlier regions across all three species (Figure S6), and the small
274 subset of genes shared among pairs of species was not enriched for any biological functions (Table
275 S9). *Pseudoxiphophorus bimaculatus* exhibited enrichment of both ion transport and oxygen transport
276 genes, as well as evidence for a selective sweep in the region containing SQR, a key H_2S
277 detoxification enzyme (Table S10). However, neither *P. mexicana* nor *X. hellerii* exhibited any
278 evidence for selective sweeps in regions containing genes related to H_2S, and only *X. hellerii* exhibited
279 limited enrichment for GO terms in sweep regions (plasma membrane associated genes; Table S10).
280

281 **Discussion**

282 The three poeciliid fishes that have colonized the La Gloria sulfide spring complex provided a
283 unique opportunity to examine how distantly related lineages respond to the same sources of
284 divergent selection. This approach helped reveal which aspects of evolutionary responses are
285 repeatable and predictable and how responses and the predictability of these responses vary across
286 biological scales of organization. We found evidence for local adaptation in each of the lineages and
287 detected signatures of convergence in phenotypic traits that span levels of organismal organization.
288 In addition, there was convergent selection on mitochondrially-encoded genes associated with H_2S
289 toxicity. However, evidence for convergence in the nuclear genomes was scant and not associated
290 with genomic regions predicted to respond to selection from H_2S. Hence, landscapes of genomic
291 divergence can be highly idiosyncratic even when species experience the same selective regime in
292 sympatry, suggesting that evolutionary outcomes are contingent on the genomic substrates that
293 selection is acting on. These results also highlight the potential methodological challenges of
294 identifying convergent shifts associated with highly polygenic traits, phenotypic traits that are
295 associated with high genetic redundancy, or traits underpinned by convergent structural variation.
296

297 ***Evidence for Convergence in Traits Associated with Adaptation to Sulfide Springs***

298 Consistent with previous findings,[41,42] we found significant differences in body shape between
299 sulfidic and nonsulfidic populations of all three lineages at La Gloria, with sulfide spring fishes
300 exhibiting larger heads than those in nonsulfidic habitats. Previous studies have indicated that
301 increased head size in sulfide spring poeciliids reflects a response to selection on oxygen acquisition

302 in the hypoxic sulfide springs, as head size is correlated with increased gill surface area and higher
303 ventilation capacity in sulfide spring populations of *P. mexicana*[41,43] and other fishes.[44,45]
304
305 Convergent phenotypic differentiation in H_2S tolerance between sulfidic and nonsulfidic fishes was
306 also observed. Sulfide spring populations exhibited higher tolerance to acute H_2S exposure than
307 populations from adjacent nonsulfidic populations, and, accordingly, nonsulfidic fishes suffered high
308 mortality when placed into the natural conditions of the sulfide spring. Previous studies have
309 demonstrated that sulfide spring populations of *P. mexicana* consistently exhibit higher tolerance to
310 acute exposures to H_2S, whether it is quantified at a molecular or whole-organism level.[25,41] Increased
311 tolerance results from modification of the expression and protein coding sequences of H_2S toxicity
312 and detoxification genes in well characterized biological pathways.[21,25] Additionally, our findings
313 supported those of previous studies carried out in more divergent poeciliid lineages occupying
314 sulfide springs across the Neotropics, which indicated that convergent shifts in the expression of
315 H_2S toxicity and detoxification genes is correlated with H_2S tolerance.[25,46] We found that upregulated
316 genes in the gills of sulfidic populations from all three species were enriched for genes with
317 mitochondrial functions and H_2S processing, including genes (*COX* and *SQR*) demonstrated to
318 exhibit adaptations for processing H_2S or resisting its toxicity in sulfide spring *P. mexicana*.[21,25]
319 Ultimately, these shared expression patterns across the three sulfide spring populations provide
320 robust support for convergent regulation of genes with known roles in H_2S toxicity and
321 detoxification.
322
323 While these signals of convergence are substantial, we cannot eliminate the possibility that
324 differences in body shape, sulfide tolerance, and gene expression observed in our experiments with
325 wild-caught specimens were at least in part shaped by plastic responses to H_2S exposure, rather than
326 evolved differences between populations. Many physiological traits, such as sulfide tolerance and
327 gene expression, are known to exhibit short-term plastic responses to changes in environmental
328 conditions.[30] Common-garden rearing of these species will be needed to better understand how
329 genetic differentiation and plasticity interact to shape gene expression patterns observed in nature;
330 however, laboratory rearing of these sulfidic populations has proven difficult (Greenway and Tobler,
331 personal observation). Previous common-garden experiments of other sulfidic and nonsulfidic *P.*
332 *mexicana* populations, notably less divergent than the populations examined in this study, have
333 provided evidence for a heritable basis to divergence in body shape,[47] sulfide tolerance,[48] and

334 expression variation in H_2S-processing and OxPhos-related genes.[30,31] Furthermore, our population
335 genetic analyses demonstrated that fish collected from the sulfidic and nonsulfidic habitat exhibit
336 very little gene flow between them, despite the possibility of panmixia due to the connectedness of
337 the streams in which they are found, suggesting that the phenotypic differentiation we observed
338 likely does not solely reflect habitat-specific inducible effects but are at least in part the result of
339 evolved differences.
340
341 ***Genetic Basis of Adaptation***
342 Even with clear evidence for convergent phenotypic responses to shared environmental conditions,
343 we found that patterns of convergence were not consistent across scales of biological organization.
344 Specifically, we found limited evidence for convergent evolution at the genomic level, suggesting
345 that convergence in morphology, physiological tolerance, and gene expression do not simply stem
346 from genomic differentiation at one, or few, genomic loci. Our results show that nuclear genes
347 typically implicated in H_2S adaptation were not associated with shared or species-specific outlier
348 regions, including genes of the SQR pathway that are involved in enzymatic H_2S detoxification and
349 nuclear genes associated with COX and other OxPhos components. These results stand in stark
350 contrast to previous investigations of replicated *P. mexicana* populations adapted to sulfide springs
351 across southern Mexico, where a variety of approaches have demonstrated predictable evolution of
352 genes with known roles in H_2S detoxification and metabolism.[21,23,24] This high level of convergence
353 among closely related populations paired with the limited convergence observed among divergent
354 lineages in the present study suggests that even the strong physiological demands of H_2S toxicity are
355 not sufficient to constrain adaptation to one or a few solutions at the genomic level. Rather, access
356 to shared pools of genetic variation, as is the case for replicated *P. mexicana* populations,[23,26] appears
357 to facilitate convergence at the genomic level. These results align with previous studies investigating
358 genome-wide patterns of changes in protein coding sequences (based on transcriptomic data), which
359 found little evidence for convergent evolution of protein coding sequences outside of the
360 mitochondrial genome.[25] However, in line with previous studies, we did observe consistent
361 modification of genes encoded by the mitochondrial genome across all three sulfidic populations,
362 including the direct targets of H_2S toxicity (*COX* subunits). Additionally, nonsulfidic mitochondrial
363 haplotypes were not found in individuals sampled from the sulfidic habitat (except for one apparent
364 F1 hybrid), suggesting a critical role of mitochondrial DNA in mediating survival in the toxic
365 conditions. Despite evidence for selection on mitochondrial genes,[49] it remains unclear how genetic

366 variation in mitochondrial versus nuclear loci contribute to adaptive variation in OxPhos function.
367 The evolution of H_2S-resistance could be largely driven by modification of mitochondrial *COX*
368 subunits that form the reactive center of the protein and without coevolutionary changes in the
369 corresponding nuclear subunits. This finding challenges the general paradigm that adaptation of
370 mitochondrial function coincides with mito-nuclear coevolution.[49]
371
372 While evidence of selection on nuclear genes associated with H_2S detoxification and OxPhos was
373 sparse, functional annotation of shared and unique outlier regions did find limited evidence for
374 enrichment of genes involved in oxygen transport and immune function (Tables S8). Evidence for
375 selection on oxygen transport proteins, such as hemoglobin is not surprising, as H_2S is capable of
376 binding to and impairing hemeprotein function.[50,51] The presence of hemoglobin genes in outlier
377 regions also confirms findings of positive selection on hemoglobin genes in other sulfide spring
378 poeciliids.[52] In contrast, the large number of genes associated with immune function was not
379 expected, although immune genes are known hotspots for selection in vertebrates.[53] Selection on
380 immune genes could be driven by differences in parasite and pathogen communities between
381 sulfidic and nonsulfidic habitats, which have been documented in another sulfide spring inhabited by
382 a poeciliid fish.[54] Alternatively, there have been long-standing, but poorly investigated, hypotheses
383 about the potential role of H_2S-oxidizing microsymbionts in adaptation to sulfide spring
384 environments,[16,55] and changes in host-microbe interactions upon sulfide-spring colonization may
385 also exert selection on immune genes.[56]
386
387 Even though H_2S has clear-cut biochemical and physiological effects, tolerance to this toxicant
388 involves numerous genes. OxPhos complexes alone contain over 80 nuclear-encoded genes,[57] and
389 dozens of genes contribute to the physiological processing of H_2S and other sulfur compounds.[58,59]
390 For this reason, the lack of convergence we observed at the genome level could be the result of the
391 polygenic nature of H_2S tolerance, which is known to pose substantial challenges for detecting
392 patterns of genomic convergence, since associated loci are unlikely to form large islands of
393 divergence but rather be scattered throughout the genome. It is possible that subtle allele frequency
394 shifts across multiple polymorphic OxPhos or H_2S processing genes could result in adaptive
395 modifications to these pathways without leaving a molecular signature detectable by scans for
396 genomic differentiation between divergent species pairs.[60,61] Such subtle allele frequency shifts are
397 now detectable when closely related populations with shared allelic variants are exposed to similar

398 sources of selection[62] but not in highly divergent population pairs that were investigated here. It is
399 also important to emphasize that environmental gradients between sulfidic and adjacent nonsulfidic
400 habitats are complex (Figure 1C). The multifarious sources of selection have caused adaptive
401 modifications of whole suites of complex traits beyond H_2S tolerance, and most of these traits likely
402 have different genomic bases.[18] Finally, another important factor influencing our ability to detect
403 genomic convergence is whether H_2S tolerance involves substantial genetic redundancy, where many
404 different combinations of alleles throughout the genome can underpin the same phenotype. If
405 genetic redundancy is high, lineage-specific responses to selection that are in part shaped by
406 historical contingencies[63–65] may underpin the convergence signatures of phenotypic traits and gene
407 expression without genomic convergence across species.[66]
408
409 It is also possible that adaptation to H_2S may primarily be driven by changes in mitochondrial
410 protein function and the expression of detoxification genes, with lineage-specific mechanisms
411 producing the convergent gene expression patterns uncovered by transcriptome analyses. Because
412 differentially expressed genes were not overrepresented in the outlier genome regions as might be
413 expected by selection acting on *cis*-regulatory elements,[24] outlier regions may harbor *trans*-regulatory
414 elements mediating adaptive shifts in gene expression.[67] Divergent selection acting on regulatory
415 elements has been implicated in adaptive divergence for several well-known models in evolutionary
416 ecology[68–70]. The problem of detecting the genetic basis of trait variation by any of the above means
417 can be exacerbated when adaptive evolution gives rise to reproductive isolation. Selection on small-
418 effect loci spread across the genome can facilitate strong and stable reductions in gene flow over
419 short timescales,[71,72] similar to the reduced gene flow observed for all three lineages in our study.
420 High levels of background divergence resulting from post-speciation divergence can further obscure
421 the signatures of selection on the loci mediating adaptation during earlier stages of divergence.[73–75]
422
423 **Conclusions**
424 Our results showed that while convergent evolution in response to shared selective pressure is
425 present across diverse lineages, patterns of convergence to extreme environments are not consistent
426 across levels of biological organization. We also demonstrated that mechanistically linking strong
427 signals of convergence in morphology, physiology, and gene expression to genomic variation is
428 challenging due to the magnitude of ways in which different aspects of genomic variation interplay
429 to affect phenotypic variation. Although we observed signatures of independent local adaptation and

430 reproductive isolation between sulfidic and nonsulfidic populations in three highly divergent
431 lineages, variation in the complexity and genomic architecture of traits and the degree of genetic
432 redundancy may reduce our ability to confidently identify signals of convergence at the genome
433 level. While substantial progress has been made in understanding the molecular basis of trait
434 evolution, methodological and statistical limitations are prevalent when traits are complex, leading to
435 an underrepresentation of studies covering the adaptation of non-model organisms to complex
436 ecological stressors.
437
438 **STAR Methods**
439 ***Study sites and sample collection***
440 In the foothills of the Sierra Madre de Chiapas, several freshwater springs rich in naturally occurring
441 H_2S form the La Gloria spring complex.[41,76] The spring complex is located near the city of Teapa,
442 Mexico, and is part of the Río Pichucalco drainage. Habitats with high H_2S concentrations are
443 spatially restricted (~200 m in length), and the spring run flows directly into a nearby nonsulfidic
444 stream (Arroyo Caracol). There are no physical barriers that prevent fish movement between the
445 sulfidic spring run and the nonsulfidic stream, but there are stark differences in water chemistry
446 (Figure 1). Populations of the poeciliid fishes *Poecilia mexicana, Pseudoxiphophorus bimaculatus*, and
447 *Xiphophorus hellerii* have colonized H_2S-rich habitats at La Gloria, and ancestral populations also
448 occur in adjacent nonsulfidic habitats.[77] Note that the population of *P. mexicana* at La Gloria likely
449 belongs to an endemic species described from sulfide springs, *P. sulphuraria.*[77,78] All samples used in
450 this study were collected by seine from the La Gloria sulfide spring complex and nonsulfidic habitats
451 in the adjacent Arroyo Caracol. The sole exception were *Pseudoxiphophorus* gill tissues for the
452 nonsulfidic population, which were collected from another nearby stream (Table S11).
453
454 Physical and chemical water parameters were analyzed for several sites within the La Gloria spring
455 complex and Arroyo Caracol. Temperature, specific conductivity, pH, and oxygen content were
456 measured using a Hydrolab Multisonde 4A (Hach Environmental). Environmental H_2S
457 concentrations were measured with a methylene blue assay using a Hach DR1900 Portable
458 Spectrophotometer (Hach Company, Loveland, CO, USA). Measurements and calibration of probes
459 were conducted according to the manufacturer's recommendations. At least three Hydrolab readings
460 and two H_2S samples were taken at each site. Measurements were averaged for each site prior to use
461 in a principal component analysis.

462

463 ***Evidence for convergence in traits associated with adaptation to sulfide springs***

464 To identify potential convergent evolution across the three focal species, we investigated three
465 distinct complex traits that have previously been implicated in adaptation to sulfide spring
466 environments: (1) body shape; (2) tolerance to acute H_2S exposure; and (3) genome-wide patterns of
467 gene expression.

468

469 *Analysis of body shape*

470 To test for divergence in body shape among populations in sulfidic and nonsulfidic habitats, adult
471 *Poecilia*, *Pseudoxiphophorus*, and *Xiphophorus* were collected from the two sites using seines. Lateral
472 photographs were taken for all specimens using a Canon EOS 400D digital camera (Canon USA
473 Inc., Lake Success, NY, USA) mounted on a copy stand. We digitized 14 morphological landmarks
474 (Figure S7) on each photograph using the software program tpsDig.[79] We extracted coordinates of
475 these digitized landmarks to conduct geometric morphometric analyses[80] using the GEOMORPH
476 package in R.[81] We first aligned landmark coordinates via least-squares superimposition implemented
477 in the gpagen function to remove effects of translation, rotation, and scale, yielding aligned
478 Procrustes coordinates that describe body shape variation and centroid size as a measure of size.
479 Procrustes coordinates were then analyzed using a Procrustes ANOVA (procD.lm) using residual
480 randomization over 9999 iterations and type III sums of squares. Predictor variables included
481 species (*Poecilia*, *Pseudoxiphophorus*, or *Xiphophorus*), habitat (sulfidic vs. non-sulfidic), and sex as
482 factors, centroid size as a covariate, as well as their interaction terms. *P*-values were based on a
483 Cohen's f-squared sampling distribution. For data visualization, we calculated divergence vector
484 scores[82] for each individual based on the first principle component of the among-group covariance
485 matrix for the habitat term from a preparatory MANCOVA using the same model as for the
486 Procrustes ANOVA.[83] This approach allowed us to visualize aspects of body shape variation
487 associated with habitat type among species (*i.e.*, convergent aspects of body shape), while accounting
488 for all other terms in the model.

489

490 *Tolerance to acute H_2S exposure*

491 To test for differences in H_2S tolerance between individuals collected from the sulfidic and
492 nonsulfidic populations, we conducted acute H_2S exposure trials, subjecting wild-caught fish to
493 logarithmically increasing concentrations of H_2S.[41] We collected adult *Poecilia*, *Pseudoxiphophorus*, and

494 *Xiphophorus* from the two sites using seines, transferred them to insulated coolers filled with water
495 from the collection site, and transported them to a nearby field station. Fish were kept in population
496 and species-specific holding tanks with aeration for at least 24 hours prior to testing. To standardize
497 experimental conditions, water from the collection sites was slowly replaced with H_2S-free well water
498 over the first 8 hours in the holding tanks. Water was continuously aerated and filtered during this
499 time, and the fish received no food.

500 For the experiment, we prepared stock solutions of 10 mM aqueous H_2S solution by
501 dissolving 2.4 g sodium sulfide hydrate ($Na_2S{\cdot}6H_2O$) into 1 L of well water deoxygenated through
502 purging with nitrogen.[84] Individual fish were placed into clear plastic containers with 150 mL water
503 from the holding tanks and allowed to acclimate for 5 minutes. Following acclimation, 10 mL of H_2S
504 solution were added to the experimental container at 2-minute intervals using a syringe placed under
505 the water surface. Each fish was observed as H_2S concentration increased in the experimental
506 container. We measured the time until the fish lost equilibrium, at which point the fish was removed
507 from the container, sexed, weighed, and placed into a heavily aerated recovery tank. Experiments
508 were ended after 32 minutes (15 sulfide additions) if a fish did not lose equilibrium.

509 We analyzed variation in H_2S tolerance using a survival analysis (Cox regression) with the
510 coxph function implemented in the R package SURVIVAL.[85] We used time to loss of equilibrium as
511 the response variable, and individuals that did not lose equilibrium during the experiment were
512 censored. Body mass (log_{10}-transformed), species, sex, and habitat of origin (sulfidic vs. nonsulfidic)
513 were used as independent variables. Adjusted survival curves were generated for the habitat term of
514 the Cox regression to visualize expected survival curves, using the ggadjustedcurves function from
515 the SURVMINER package.[86]

516

517 *Transcriptome sequencing and alignment*

518 We used an RNA-sequencing approach to identify differentially expressed genes between the
519 sulfidic and nonsulfidic populations of each species. We collecteded adult female individuals of
520 *Poecilia*, *Pseudoxiphophorus*, and *Xiphophorus* in the La Gloria sulfide spring and from the adjacent
521 nonsulfidic stream. Following capture with a seine, fish were instantly euthanized, and gill tissues
522 were extracted from both sides of the head and immediately preserved in 2 mL of RNAlater
523 (Ambion, Inc.). We selected gill tissue because the gills mediate physiological processes necessary for
524 the maintenance of homeostasis,[87] are in direct contact with the H_2S-rich water,[16] and show strong,
525 heritable transcriptional responses upon H_2S exposure.[30,46] RNA was isolated from gill tissues, and

526 cDNA libraries were constructed for each sample as described for samples in previous studies.[25,46]
527 Briefly, we extracted total RNA from pulverized gill tissues using the NucleoSpin RNA kit
528 (Macherey-Nagel GmbH & Co. KG, Düren, Germany) following the manufacturer's protocol. We
529 conducted mRNA isolation and cDNA library preparation with the NEBNext Poly(A) mRNA
530 Magnetic Isolation Module (New England Biolabs, Inc., Ipswich, MA, USA) and NEBNext Ultra
531 Directional RNA Library Prep Kit for Illumina (New England Biolabs, Inc., Ipswich, MA, USA),
532 following the manufacturers' protocol with minor modifications. Each cDNA library was assigned a
533 unique barcode, quantified, and pooled in sets of 11–12 samples that were sequenced together in a
534 way that different species and habitat types were sequenced together. cDNA libraries were
535 sequenced on an Illumina HiSeq 2500 with paired-end, 100-bp reads at the Washington State
536 University Spokane Genomics Core.
537 Raw RNA-sequencing reads were sorted by barcode and trimmed twice (*--quality* 0 to remove
538 adapters, followed by *--quality* 24) with Trimgalore![88] Trimmed reads for all 36 individuals were
539 mapped to the *X. maculatus* reference genome used for genomic analyses[89] using BWA-MEM.[90] On
540 average, ~97% of reads mapped across all individuals (Table S12). We functionally annotated genes
541 from the *X. maculatus* reference genome by extracting the longest transcript per gene (using the perl
542 script gff2fasta.pl[91]) and then searching against the human SWISSPROT database (critical E-value
543 0.001; access date 11/15/2018) using BLASTx.[92] Each *X. maculatus* gene was annotated with the top
544 BLAST hit based on the top high-scoring segment pair. Annotations were used for analysis of
545 differentially expressed genes as well as for analyses of outlier regions from genomic data (see
546 below).
547 We used STRINGTIE v.1.3.3b[93,94] to quantify the number of RNA-seq reads mapped to
548 each gene in the *X. maculatus* reference genome for each individual and used a Python script
549 provided with STRINGTIE (prepDE.py) to generate a read counts matrix.[94] We then removed
550 genes that did not have at least two counts per million in 3 or more individuals across all species,
551 resulting in a set of 18,598 genes for analysis of gene expression patterns. We first conducted
552 hierarchical cluster analysis on the full set of retained genes. The hierarchical cluster analysis was
553 performed on $\log_2$-transformed counts-per-million and visualized with the heatmap.2 function in the
554 GPLOTS package in R.[95]
555 To identify differentially expressed genes, we used generalized linear models (GLMs) as
556 implemented in the Bioconductor package EDGER.[96] We used glmFit to fit a negative binomial GLM
557 to the normalized read counts of each gene based on tagwise dispersion estimates and a design

558 matrix describing the comparison between the sulfidic and nonsulfidic population of each species.
559 We assessed statistical significance with the GLM likelihood-ratio test with a false-discovery rate
560 (FDR) of < 0.05 calculated with the Benjamini-Hochberg correction.[97] We then intersected the
561 significantly upregulated and downregulated genes from all three species-specific comparisons
562 separately to identify genes with consistent evidence for differential expression among all three
563 species. After identifying the set of genes with differential expression between sulfidic and
564 nonsulfidic populations of each species, and genes with evidence for shared differential expression
565 across all three species-specific analyses (Table S2), we used a Gene Ontology (GO) enrichment
566 analysis to explore the putative biological functions of these candidate sets of genes (Tables S3). We
567 first annotated all genes that had a match in the human SWISSPROT database with GO IDs.[98] In
568 total, 10,817 unique SWISSPROT annotations were associated with a term in the GO database. We
569 then tested for the enrichment of specific GO IDs in these sets of genes with differential expression
570 relative to the full dataset of 18,598 analyzed genes using GORILLA (*P*-value threshold: 0.0001,
571 accessed 04/11/2019).[99]
572
573 ***Evidence for local adaptation***
574 To test for local adaptation in each of the three sulfidic/nonsulfidic population pairs, we used
575 whole-genome resequencing to assess the contemporary levels of population genetic structure. We
576 also conducted a reciprocal translocation experiment to compare the survival of individuals across
577 the sulfidic/nonsulfidic environmental gradient.
578
579 *Genomic DNA library preparation and whole genome sequencing*
580 We resequenced the genomes of 120 individuals (40 per species, 20 per population). DNA was
581 extracted from fin clip tissues that were stored in ethanol using the NucleoSpin Tissue kit
582 (Macherey-Nagel GmbH & Co. KG, Düren, Germany) following the protocol in the user manual
583 (version January 2017/Rev. 17). For the final elution step, extracted DNA was eluted in 100 μL of
584 elution buffer.
585
586 Library preparation and sequencing were conducted at the Washington State University Spokane
587 Genomics Core. Genomic DNA (gDNA) was quantified using the Qubit 2.0 fluorometer with the
588 dsDNA HS assay kit (Thermo Fisher Scientific, Waltham, MA). One hundred nanograms of gDNA
589 were used as input for library preparation using the TruSeq Nano DNA Library Prep Kit (Illumina,

590 San Diego, CA). gDNA was sheared using the M220 Focused-Ultrasonicater (Covaris, Woburn,
591 MA), followed by end repair, size selection (350-bp insert size), dA-tailing, adaptor ligation, and
592 library amplification by eight rounds of PCR. The quality of the DNA libraries was assessed by
593 Fragment Analyzer with the High Sensitivity NGS Fragment Analysis Kit (Advanced Analytical
594 Technologies, Ankeny, IA). Library concentrations were measured by the StepOnePlus Real-Time
595 PCR System (ThermoFisher Scientific, Waltham, MA) with the KAPA Library Quantification Kit
596 (Kapabiosystems, Wilmington, MA). Each library was labeled with an index sequence during the
597 adaptor-ligation step using Illumina's TruSeq Dual Indexing system (Illumina, San Diego, CA),
598 which contains 96 unique indices. The 120 libraries were split and pooled into two groups at equal
599 molar concentrations, with each library having a unique index within its group. The library pools
600 were diluted to 2 nM with Reticulocyte Standard Buffer (RSB; 10 mM Tris-HCl, pH 8.5) and
601 denatured with 0.1 M NaOH. Eighteen pM libraries were clustered in a high-output flow cell using
602 the HiSeq Cluster Kit v4 on a cBot (Illumina, San Diego, CA). After cluster generation, the flow cell
603 was loaded onto HiSeq 2500 for sequencing using the HiSeq SBS kit v4 (Illumina, San Diego, CA).
604 DNA was sequenced with paired-end, 100-bp chemistry among four lanes. Raw bcl files were
605 converted to fastq files and demultiplexed using the program bcl2fastq2.17.1.14
606 (https://github.com/thepler/bcl2fastq2).
607
608 *Mapping and genotyping*
609 Raw reads were trimmed to remove adapters with Trimgalore! *(--quality 0)*.[88] Trimmed reads for all
610 120 individuals were mapped to the *Xiphophorus maculatus* reference genome (v.5.0, RefSeq accession
611 number: GCF_002775205.1)[89] using BWA-MEM.[90] We used this reference genome because *X.*
612 *maculatus* was the only species in the family Poeciliidae with a chromosome-scale genome assembly.
613 On average, ~94% of trimmed reads mapped across all individuals (Table S4). After mapping, based
614 on percent of mapped reads, it became clear that one sample (sample MX16-205) labeled as
615 *Pseudoxiphophorus* from a sulfidic habitat was likely mislabeled, and this sample was excluded from all
616 further analyses. Note that we conducted the same analyses mapping to the *P. mexicana* reference
617 genome, and all results were consistent.
618
619 The resequenced genomes generated for this study had an average sequence depth of 3.8 (± 0.4;
620 Table S4). We used the Genome Analysis Toolkit (GATK, v.3.5)[100] and Picard Tools[101] to mark
621 duplicate reads (using the MarkDuplicates tool) and realign mapped reads around indels (using the

622 RealignerTargetCreator and IndelRealigner tools) following GATK best practices.[102,103] The resulting
623 realigned bam-file for each individual was used in subsequent analyses. GATK was also used to
624 identify single nucleotide polymorphisms (SNPs) per population. HaplotypeCaller was used to
625 identify genotypes per individual with the *--ERC* GVCF flag.[100]
626
627 Samples were genotyped jointly on a per-species basis using GATK's GenotypeGVCFs tool after
628 combining GVCF files from the sulfidic and nonsulfidic populations using the GATK
629 CombineGVCFs tool. Genotypes were identified using the *--include-non-variant-sites* flag to retain all
630 sites in the initial VCF. We recoded genotypes inferred as phased missing (.|.) as unphased missing
631 (./.) due to a formatting issue with VCFtools. We used the GATK SelectVariants to extract biallelic
632 SNPs from the original genotyped GVCF files, except for the d_{XY} analysis (see below). For all
633 analyses relying on biallelic SNPs, combined VCF files with both the sulfidic and nonsulfidic
634 populations of each species were filtered to exclude sites with >25% missing data, a minor allele
635 frequency below 5%, and alleles that were fixed for the non-reference allele were removed using
636 VCFtools v.0.1.17 with the following flags: *--max-missing* 0.75, *--remove-filtered-all*, *--maf 0.05*, and *--max-*
637 *non-ref-af-any* 0.999. After filtering, 1,994,206 biallelic SNPs were retained for *Poecilia*, 5,284,684 for
638 *Pseudoxiphophorus*, and 4,624,801 for *Xiphophorus*. For calculating d_{XY}, only indels and sites with >25%
639 missing data were removed from the merged VCF using VCFtools with the *-remove-indels* and *-max-*
640 *missing 0.75* flags, respectively (note that this filtering retained monomorphic sites for accurate
641 estimation of d_{XY}).[104] This resulted in the retention of 439,564,228 monomorphic and biallelic sites
642 for *Poecilia*, 528,674,992 for *Pseudoxiphophorus*, and 611,098,475 for *Xiphophorus*.
643
644 *Analyzing population structure*
645 We assessed structure between populations in sulfidic and nonsulfidic habitats of each species. To
646 do so, we removed sites with an r^2 above 0.7 within 5 kb within VCFtools *(--geno-r2 --min-r2* 0.7 *--ld-*
647 *window* 5000), producing a set of unlinked SNPs for each species (62,738 in *Poecilia*, 64,405 in
648 *Pseudoxiphophorus*, and 65,515 in *Xiphophorus*). We investigated the extent of gene flow between
649 populations using FastStructure,[105] a maximum-likelihood-based assignment method that estimates
650 admixture proportions for each individual. FastStructure was run with the number of populations
651 (*K*) set as 2 to test if individuals cluster by habitat of origin (*i.e.*, sulfidic vs. nonsulfidic). Population
652 structure was further examined by principal component analysis (PCA) based on a variance-
653 standardized genetic relationship matrix computed with *PLINK* v.2.00.[106]

654

655 From the structure analyses, we identified an admixed individual of *P. mexicana* that was likely an F1
656 hybrid between the sulfidic and non-sulfidic populations (MX16-023). We also identified another
657 admixed individual in *P. mexicana* that was possibly an F1 or F2 backcross (MX16-121). We removed
658 these admixed individuals from the *Poecilia* SNP sets for downstream analyses.

659

660 *Reciprocal translocation experiments*

661 To experimentally validate local adaptation[32] and test for natural selection against migrants between
662 sulfidic and nonsulfidic populations,[33] we conducted reciprocal translocation experiments using
663 previously established approaches.[19,107] Large (20-liter) plastic buckets were placed into the two
664 habitat types as experimental mesocosms. Two holes (18 × 32 cm) were cut into opposite sides of
665 each bucket and sealed with 1.5 mm plastic mesh to allow the free exchange of water with the
666 environment. Approximately 50 small holes (<1 mm) were drilled into the bucket lids to facilitate air
667 exchange. Mesocosms were placed into shallow areas of the sulfidic and nonsulfidic streams, seeded
668 with a 3–4 cm layer of natural substrate, and allowed to settle overnight before the start of the
669 experiment. We established ten mesocosms in each habitat. Water conditions in these mesocosms
670 have been shown to closely match conditions in the surrounding stream.[107]

671

672 We collected adult *Poecilia*, *Pseudoxiphophorus*, and *Xiphophorus* by seine and placed them in insulated
673 coolers with aerated water for transport to the mesocosm locations. Five haphazardly chosen
674 individuals of the same species and habitat type were introduced into a mesocosm. Half of the
675 mesocosms in each habitat were used as controls to test the survival of resident fish, and the other
676 half to test survival of fish from the opposite habitat type. Transportation and handling times were
677 minimal (<1 hour) and were balanced for resident and translocated individuals. Fish were measured
678 for standard length prior to introduction. Experiments ran for ~20 hours before termination.
679 Following the experimental period, we quantified survival and returned surviving individuals to their
680 original collection site.

681

682 To analyze variation in survival, we used a generalized linear mixed model (GLMM) with a binomial
683 error distribution and a logit-link function, as implemented in the R package LME4.[108] Survival
684 (binary data: 0 = died; 1 = survived) was used as the response variable. We included species, habitat
685 of origin (sulfidic vs. nonsulfidic), and testing habitat (sulfidic vs. non-sulfidic) as fixed effects

686 (including their interactions terms), and standard length (log_{10}-transformed) as a covariate.
687 Additionally, mesocosm ID was included as a random effect.
688
689 ***Evidence for convergence at a genomic level***
690 The final objective of this study was to identify genomic signatures of convergent evolution. To do
691 so, we leveraged the genome resequencing data described above. We analyzed patterns of molecular
692 evolution in the mitochondrial genomes of focal populations and explored patterns of divergence
693 and signatures of selection in the nuclear genomes as well.
694
695 *Mitochondrial genome assembly and analysis*
696 Trimmed reads from the whole-genome re-sequencing were used as input for NOVOPlasty,[109]
697 which allows for *de novo* assembly and variant-calling in short circular genomes. The *X. maculatus*
698 reference mitochondrial genome (NCBI Accession NC_011379) was used as the seed input and
699 reference sequence. Sequences for the 13 protein-coding genes were identified using
700 MitoAnnotator.[110]
701
702 To investigate the evolutionary history of mitochondrial genomes among populations, we performed
703 phylogenetic analyses using 12 of the 13 protein-coding mitochondrial genes (*atp8* was excluded
704 from analyses due to poor assembly for this short gene in many individuals). For each gene, we
705 aligned nucleotide sequences of all individuals using the default options of the MAFFT progressive
706 alignment method,[111] guided by a translated amino acid sequence with TranslatorX.[112] We used
707 PopART v.1.7[113] to construct a TCS haplotype network[114] for the concatenated mitochondrial
708 sequences of each population pair separately.
709
710 Since mitochondria and OxPhos are the direct target of H_2S toxicity, we analyzed patterns of
711 molecular evolution and the effect of positive selection in lineages from sulfidic habitats for 12 of
712 the 13 mitochondrially encoded protein-coding genes, following a previously developed approach.[25]
713 In brief, because of the high coverage of the RNAseq data, we used GATK[100] to realign BWA-MEM
714 mapped RNAseq reads around indels and used the UnifiedGenotyper to detect variants following
715 GATK best practices.[102,103] We genotyped individuals on a per-population basis using the
716 UnifiedGenotyper EMIT_ALL_SITES output mode to create a VCF-file for each population. We
717 then used a custom pipeline to fill positions lacking sequence data in each VCF with Ns, rather than

718 the GATK default (which fills in missing data with the reference allele).[115] We used BCFtools v.1.3
719 to designate the genotype at all variable sites as the major allele in each population's VCF-file with
720 the setGT tool.[116] Using the coordinates for each mitochondrial gene from the *X. maculatus* reference
721 genome gene annotation file (GTF), we created population-specific consensus sequences for each
722 mitochondrial gene across all individuals in a population using the consensus tool in BCFtools. We
723 aligned nucleotide sequences for all populations using the MAFFT progressive alignment method[111]
724 guided by the translated amino acid sequence as implemented in TranslatorX,[112] producing an in-
725 frame multispecies alignment of each gene. We constructed gene-specific phylogenies (gene trees)
726 for each sequence using the DNAML program in PHYLIP v.3.696.[117]
727
728 To identify mitochondrial genes that have experienced positive selection in multiple sulfide spring
729 populations, as evidenced by an excess rate of non-synonymous to synonymous substitutions (ω) for
730 codons in each mitochondrial gene across the sulfidic branches of our phylogenies, we used branch
731 models implemented by codeml in PAML v.4.9a.[118] For each mitochondrial gene, codeml was run
732 using the in-frame multispecies alignment and gene tree with the three sulfidic populations
733 designated as foreground branches and nonsulfidic populations as background branches, allowing us
734 to test whether ω differed between foreground and background branches in a two-ratio branch
735 model (M2; model = 2, NSsites = 0) compared against a null model with a single ω value for all
736 branches (M0; model = 0, NSsites = 0). A likelihood-ratio test was used to assess whether models
737 significantly differed for each gene, with the test statistic calculated as two times the difference
738 between the ln-likelihood of the branch model and the ln-likelihood of the null model, and *P*-values
739 calculated from a c^2 approximation.[119]
740
741 *Characterizing the genomic landscape of divergence*
742 To characterize the landscape of genomic divergence, we calculated genomic differentiation (F_{ST},
743 relative divergence) and divergence (d_{XY}, absolute divergence) between the sulfidic and nonsulfidic
744 population of each species across non-overlapping 5-kb windows spanning the genome from our
745 filtered dataset (excluding the two putative, early-stage hybrids in *Poecilia*, as described above) using
746 the popgenWindows.py script.[120–122] Divergent selection between habitats should lead to elevated
747 levels of both F_{ST} and d_{XY} at and around loci under selection (those contributing to the maintenance
748 of adaptive divergence), while regions of the genome with reduced diversity in one or both
749 populations (likely a consequence of background selection or low recombination rates) are expected

750 to exhibit elevated F_{ST} but not d_{XY}.[34,37,123] To identify probable targets of divergent selection between
751 sulfidic and nonsulfidic populations, we used a composite outlier analysis of both statistics as
752 implemented in the R package MINOTAUR.[38] First, we converted F_{ST} and d_{XY} to rank-based *P*-values
753 using a right-tailed test, then $-\log_{10}$-transformed these *P*-values, and calculated the Mahalanobis
754 distances among windows (the "Md-rank-P" method recommended for summary statistics).[39] We
755 considered windows in the top 1% of the resulting Mahalanobis distance distribution as genomic
756 regions likely experiencing divergent natural selection between habitat types in each species. We also
757 identified regions that were classified as outlier windows across all three species and between each
758 pairwise combination of species.
759
760 We investigated the link between divergent selection and the putative biological functions of genes
761 found in these outlier windows for each species using a GO-enrichment analysis (as described
762 above). We first identified genes in outlier regions from the *X. maculatus* genome annotation file
763 (GTF) using the *intersect* tool in BEDTools v.2.25.0.[124] We then annotated all genes that had a match
764 in the human SWISSPROT database (21,891 genes) with GO IDs and tested for enrichment of
765 specific GO IDs in outlier regions relative to all human annotated genes in the *X. maculatus* genome
766 using GORILLA (*P*-value threshold: 0.0001, accessed 09/29/2020).[99] In total, 12,247 unique
767 SWISSPROT annotations were associated with a GO term.
768
769 *Scans for positive selection*
770 To further test for signals of natural selection driving genomic divergence between sulfidic and
771 nonsulfidic populations, we implemented the cross-population composite likelihood ratio (XP-
772 CLR)[40] test to identify regions of the genome with evidence for selective sweeps in the sulfide spring
773 population of each species based on differentiation in multilocus allele frequencies compared to the
774 nonsulfidic population. We further filtered our biallelic SNP data to include only sites with no more
775 than 10% missing data (957,379 in *Poecilia*, 2,885,156 in *Pseudoxiphophorus*, and 3,419,489 in
776 *Xiphophorus*), and limited our analyzes to chromosomes in the assembly, excluding all short scaffolds.
777 For XP-CLR, the local two-dimensional site-frequency-spectrum (SFS) was compared to the
778 genome-wide two-dimensional SFS to identify genomic regions putatively under selection.[40] For
779 each species, we designated the nonsulfidic population as the reference population, and the sulfidic
780 population was the target population for each of the comparisons. Prior to running XP-CLR, we
781 partitioned the genome into discrete 25-kb windows. If a window contained more than 200 SNPs,

782 only the first 200 SNPs were used to calculate the composite likelihood ratio. The XP-CLR test
783 statistic was not calculated if a window contained fewer than ten SNPs. We considered windows
784 within the top 1% as putative selective sweeps.
785
786
787 **Data Availability**
788 All sequence data are available at the National Center for Biotechnology Information (NCBI)
789 (https://www.ncbi.nlm.nih.gov under BioProject accession no. PRJNA738871).
790
791 **Acknowledgements**
792 We are indebted to the community of Teapa, especially the owners of Rancho La Gloria, for
793 providing access to study sites, and to the Centro de Investigación e Innovación para la Enseñanza y
794 Aprendizaje (CIIEA) for their hospitality and support during the nearly decade-long course of this
795 project. We thank N. Barts, C. Carson, Z. Culumber, T. Doumas, G. W. Hopper, C. N. Passow, and
796 E. A. Renner for assistance in the field, as well as O. Cornejo, K. B. Gido, A. G. Hope, D. A.
797 Marques, and T. J. Morgan for their comments and discussions. This work was supported by grants
798 from the NSF (IOS-1557860, IOS-1931657, and IOS-2311366) and the US Army Research Office
799 (W911NF-15-1-0175, W911NF-16-1-0225) to JLK and MT. RG was supported by the American
800 Museum of Natural History (Theodore Roosevelt Memorial Fund grant), Friends of Sunset Zoo
801 (Conservation Scholar Program), American Livebearer Association (Vern Parish Fund award),
802 Society for the Study of Evolution (Rosemary Grant Advanced Graduate Research Excellence
803 Grant), American Society of Naturalists (Student Research Award), and a NSF Graduate Research
804 Fellowship (DGE-1746899). MT was supported by the Des Lee Collaborative Vision in Zoological
805 Studies.
806
807 **References**
808 1. Schluter, D., and Nagel, L.M. (1995). Parallel speciation by natural selection. Am. Nat. *146*,
809 292–301. 10.1086/285799.

810 2. Kaeuffer, R., Peichel, C.L., Bolnick, D.I., and Hendry, A.P. (2012). Parallel and nonparallel
811 aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of
812 lake and stream stickleback. Evolution *66*, 402–418. 10.1111/j.1558-5646.2011.01440.x.

813 3. Blount, Z.D., Lenski, R.E., and Losos, J.B. (2018). Contingency and determinism in evolution:
814 Replaying life's tape. Science *362*, eaam5979. 10.1126/science.aam5979.

4. Rosenblum, E.B., Parent, C.E., and Brandt, E.E. (2014). The molecular basis of convergent evolution. Annu. Rev. Ecol. Evol. Syst. *45*, 203–226.

5. Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D., and Kingsley, D.M. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science *307*, 1928–1933.

6. Rosenblum, E.B., Römpler, H., Schöneberg, T., and Hoekstra, H.E. (2010). Molecular and functional basis of phenotypic convergence in white lizards at White Sands. Proc. Natl. Acad. Sci. U. S. A. *107*, 2113–2117.

7. Chaturvedi, S., Gompert, Z., Feder, J.L., Osborne, O.G., Muschick, M., Riesch, R., Soria-Carrasco, V., and Nosil, P. (2022). Climatic similarity and genomic background shape the extent of parallel adaptation in Timema stick insects. Nat Ecol Evol *6*, 1952–1964. 10.1038/s41559-022-01909-6.

8. Yeaman, S., Hodgins, K.A., Lotterhos, K.E., Suren, H., Nadeau, S., Degner, J.C., Nurkowski, K.A., Smets, P., Wang, T., Gray, L.K., et al. (2016). Convergent local adaptation to climate in distantly related conifers. Science *353*, 1431–1433. 10.1126/science.aaf7812.

9. Conte, G.L., Arnegard, M.E., Peichel, C.L., and Schluter, D. (2012). The probability of genetic parallelism and convergence in natural populations. Proceedings of the Royal Society B *279*, 5039–5047. 10.1098/rspb.2012.2146.

10. Soria-Carrasco, V., Gompert, Z., Comeault, A.A., Farkas, T.E., Parchman, T.L., Johnston, J.S., Buerkle, C.A., Feder, J.L., Bast, J., Schwander, T., et al. (2014). Stick insect genomes reveal natural selection's role in parallel speciation. Science *344*, 738–742. 10.1126/science.1252136.

11. Stuart, Y.E., Veen, T., Weber, J.N., Hanson, D., Ravinet, M., Lohman, B.K., Thompson, C.J., Tasneem, T., Doggett, A., Izen, R., et al. (2017). Contrasting effects of envrionment and genetics generate a continuum of parallel evolution. Nature Ecology & Evolution *1*, 0158.

12. Morales, H.E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A.M., and Butlin, R.K. (2019). Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast. Sci. Adv. *5*, eaav9963. 10.1126/sciadv.aav9963.

13. Rosenblum, E.B., and Harmon, L.J. (2011). "Same same but different": replicated ecological speciation at White Sands. Evolution *65*, 946–960. 10.1111/j.1558-5646.2010.01190.x.

14. Raeymaekers, J.A.M., Chaturvedi, A., Hablützel, P.I., Verdonck, I., Hellemans, B., Maes, G.E., De Meester, L., and Volckaert, F.A.M. (2017). Adaptive and non-adaptive divergence in a common landscape. Nat. Commun. *8*. 10.1038/s41467-017-00256-6.

15. Weber, A.A.-T., Rajkov, J., Smailus, K., Egger, B., and Salzburger, W. (2021). Speciation dynamics and extent of parallel evolution along a lake-stream environmental contrast in African cichlid fishes. Sci Adv *7*, eabg5391. 10.1126/sciadv.abg5391.

850 16. Tobler, M., Passow, C.N., Greenway, R., Kelley, J.L., and Shaw, J.H. (2016). The evolutionary
851 ecology of animals inhabiting hydrogen sulfide-rich environments. Annu. Rev. Ecol. Evol. Syst.
852 *47*, 239–262. 10.1146/annurev-ecolsys-121415-032418.

853 17. Cooper, C.E., and Brown, G.C. (2008). The inhibition of mitochondrial cytochrome oxidase by
854 the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical
855 mechanism and physiological significance. Journal of Bioenergy and Biomembranes *40*, 533–
856 539.

857 18. Tobler, M., Kelley, J.L., Plath, M., and Riesch, R. (2018). Extreme environments and the origins
858 of biodiversity: adaptation and speciation in sulphide spring fishes. Mol. Ecol. *27*, 843–859.
859 10.1111/mec.14497.

860 19. Plath, M., Pfenninger, M., Lerp, H., Riesch, R., Eschenbrenner, C., Slattery, P.A., Bierbach, D.,
861 Herrmann, N., Schulte, M., Arias-Rodriguez, L., et al. (2013). Genetic differentiation and
862 selection against migrants in evolutionarily replicated extreme environments. Evolution *67*,
863 2647–2661. 10.1111/evo.12133.

864 20. Riesch, R., Tobler, M., Lerp, H., Jourdan, J., Doumas, T., Nosil, P., Langerhans, R.B., and
865 Plath, M. (2016). Extremophile Poeciliidae: multivariate insights into the complexity of
866 speciation along replicated ecological gradients. BMC Evol. Biol. *16*, 136. 10.1186/s12862-016-
867 0705-1.

868 21. Pfenninger, M., Lerp, H., Tobler, M., Passow, C., Kelley, J.L., Funke, E., Greshake, B., Erkoc,
869 U.K., Berberich, T., and Plath, M. (2014). Parallel evolution of cox genes in H2S-tolerant fish
870 as key adaptation to a toxic envrionment. Nat. Commun. *5*, 3873.

871 22. Pfenninger, M., Patel, S., Arias-Rodriguez, L., Feldmeyer, B., Riesch, R., and Plath, M. (2015).
872 Unique evolutionary trajectories in repeated adaptation to hydrogen sulphide-toxic habitats of a
873 neotropical fish (*Poecilia mexicana*). Mol. Ecol. *24*, 5446–5459. 10.1111/mec.13397.

874 23. Brown, A.P., McGowan, K.L., Schwarzkopf, E.J., Greenway, R., Rodriguez, L.A., Tobler, M.,
875 and Kelley, J.L. (2019). Local ancestry analysis reveals genomic convergence in extremophile
876 fishes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *374*, 20180240. 10.1098/rstb.2018.0240.

877 24. Brown, A.P., Arias-Rodriguez, L., Yee, M.-C., Tobler, M., and Kelley, J.L. (2018). Concordant
878 changes in gene expression and nucleotides underlie independent adaptation to hydrogen-
879 sulfide-rich environments. Genome Biol. Evol. *10*, 2867–2881. 10.1093/gbe/evy198.

880 25. Greenway, R., Barts, N., Henpita, C., Brown, A.P., Arias Rodriguez, L., Rodríguez Peña, C.M.,
881 Arndt, S., Lau, G.Y., Murphy, M.P., Wu, L., et al. (2020). Convergent evolution of conserved
882 mitochondrial pathways underlies repeated adaptation to extreme environments. Proc. Natl.
883 Acad. Sci. U. S. A. *117*, 16424–16430. 10.1073/pnas.2004223117.

884 26. Ryan, K., Greenway, R., Landers, J., Arias-Rodriguez, L., Tobler, M., and Kelley, J.L. (2023).
885 Selection on standing genetic variation mediates convergent evolution in extremophile fish.
886 Mol. Ecol. *32*, 5042–5054. 10.1111/mec.17081.

887 27. Tobler, M., Greenway, R., and Kelley, J.L. (2021). Ecology drives the degree of convergence in
888 the gene expression of extremophile fishes. bioRxiv, 2021.12.13.472416.
889 10.1101/2021.12.13.472416.

890 28. Tobler, M., and Plath, M. (2011). Living in extreme habitats. In Ecology and evolution of
891 poeciliid fishes, J. Evans, A. Pilastro, and I. Schlupp, eds. (University of Chicago Press), pp.
892 120–127.

893 29. Wang, M., Zhao, Y., and Zhang, B. (2015). Efficient Test and Visualization of Multi-Set
894 Intersections. Sci. Rep. *5*, 16923. 10.1038/srep16923.

895 30. Passow, C.N., Henpita, C., Shaw, J.H., Quackenbush, C.R., Warren, W.C., Schartl, M., Arias-
896 Rodriguez, L., Kelley, J.L., and Tobler, M. (2017). The roles of plasticity and evolutionary
897 change in shaping gene expression variation in natural populations of extremophile fish. Mol.
898 Ecol. *26*, 6384–6399. 10.1111/mec.14360.

899 31. Tobler, M., Henpita, C., Bassett, B., Kelley, J.L., and Shaw, J.H. (2014). H_2S exposure elicits
900 differential expression of candidate genes in fish adapted to sulfidic and non-sulfidic
901 environments. Comp. Biochem. Physiol. A Mol. Integr. Physiol. *175*, 7–14.
902 10.1016/j.cbpa.2014.04.012.

903 32. Kawecki, T.J., and Ebert, D. (2004). Conceptual issues in local adaptation. Ecol. Lett. *7*, 1225–
904 1241.

905 33. Nosil, P., Vines, T.H., and Funk, D.J. (2005). Perspective: Reproductive isolation caused by
906 natural selection against immigrants from divergent habitats. Evolution *59*, 705–719.

907 34. Ravinet, M., Faria, R., Butlin, R.K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M.A.F.,
908 Mehlig, B., and Westram, A.M. (2017). Interpreting the genomic landscape of speciation: a road
909 map for finding barriers to gene flow. J. Evol. Biol. *30*, 1450–1477. 10.1111/jeb.13047.

910 35. Nosil, P., and Feder, J.L. (2012). Genomic divergence during speciation: causes and
911 consequences. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 332–342.

912 36. Rand, D.M., Haney, R.A., and Fry, A.J. (2004). Cytonuclear coevolution: the genomics of
913 cooperation. Trends Ecol. Evol. *19*, 645–653.

914 37. Cruickshank, T.E., and Hahn, M.W. (2014). Reanalysis suggests that genomic islands of
915 speciation are due to reduced diversity, not reduced gene flow. Mol. Ecol. *23*, 3133–3157.
916 10.1111/mec.12796.

917 38. Verity, R., Collins, C., Card, D.C., Schaal, S.M., Wang, L., and Lotterhos, K.E. (2017).
918 minotaur: A platform for the analysis and visualization of multivariate results from genome
919 scans with R Shiny. Mol. Ecol. Resour. *17*, 33–43. 10.1111/1755-0998.12579.

920 39. Lotterhos, K.E., Card, D.C., Schaal, S.M., Wang, L., Collins, C., and Verity, B. (2017).
921 Composite measures of selection can improve the signal-to-noise ratio in genome scans.
922 Methods Ecol. Evol. *8*, 717–727. 10.1111/2041-210x.12774.

923 40. Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective
924 sweeps. Genome Res. *20*, 393–402. 10.1101/gr.100545.109.

925 41. Tobler, M., Palacios, M., Chapman, L.J., Mitrofanov, I., Bierbach, D., Plath, M., Arias-
926 Rodriguez, L., de León, F.J.G., and Mateos, M. (2011). Evolution in extreme environments:
927 replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. Evolution
928 *65*, 2213–2228. 10.1111/j.1558-5646.2011.01298.x.

929 42. Tobler, M., and Hastings, L. (2011). Convergent patterns of body shape differentiation in four
930 different clades of poeciliid fishes inhabiting sulfide springs. Evol. Biol. *38*, 412–421.
931 10.1007/s11692-011-9129-4.

932 43. Camarillo, H., Arias Rodriguez, L., and Tobler, M. (2020). Functional consequences of
933 phenotypic variation between locally adapted populations: swimming performance and
934 ventilation in extremophile fish. J. Evol. Biol. *33*, 512–523. 10.1111/jeb.13586.

935 44. Langerhans, R.B., Chapman, L.J., and DeWitt, T.J. (2007). Complex phenotype-environment
936 associations revealed in an East African cyprinid. J. Evol. Biol. *20*, 1171–1181.

937 45. Chapman, L.J., Galis, F., and Shinn, J. (2000). Phenotypic plasticity and the possible role of
938 genetic assimilation: Hypoxia-induced trade-offs in the morphological traits of an African
939 cichlid. Ecol. Lett. *3*, 387–393.

940 46. Kelley, J.L., Arias-Rodriguez, L., Patacsil Martin, D., Yee, M.-C., Bustamante, C.D., and Tobler,
941 M. (2016). Mechanisms underlying adaptation to life in hydrogen sulfide-rich environments.
942 Mol. Biol. Evol. *33*, 1419–1434. 10.1093/molbev/msw020.

943 47. Tobler, M., Dewitt, T.J., Schlupp, I., García de León, F.J., Herrmann, R., Feulner, P.G.D.,
944 Tiedemann, R., and Plath, M. (2008). Toxic hydrogen sulfide and dark caves: phenotypic and
945 genetic divergence across two abiotic environmental gradients in *Poecilia mexicana*. Evolution *62*,
946 2643–2659. 10.1111/j.1558-5646.2008.00466.x.

947 48. Barts, N. (2020). Comparative physiology and the predictability of evolution in extreme
948 environments.

949 49. Hill, G.E. (2019). Mitonuclear Ecology. Mol. Biol. Evol. *32*, 1917–1927.
950 10.1093/molbev/msv104.

951 50. Pietri, R., Roman-Morales, E., and Lopez-Garriga, J. (2011). Hydrogen sulfide and heme
952 proteins: knowledge and mysteries. Antioxid. Redox Signal. *15*, 393–404.

953 51. Evans, S.V., Sishta, B.P., Mauk, A.G., and Brayer, G.D. (1994). Three-dimensional structure of
954 cyanomet-sulfmyoglobin C. Proc. Natl. Acad. Sci. U. S. A. *91*, 4723–4726.

955 52. Barts, N., Greenway, R., Passow, C.N., Arias-Rodriguez, L., Kelley, J.L., and Tobler, M. (2018).
956 Molecular evolution and expression of oxygen transport genes in livebearing fishes (Poeciliidae)
957 from hydrogen sulfide rich springs. Genome *61*, 273–286. 10.1139/gen-2017-0051.

958 53. Shultz, A.J., and Sackton, T.B. (2019). Immune genes are hotspots of shared positive selection
959 across birds and mammals. Elife *8*. 10.7554/eLife.41815.

960 54. Tobler, M., Plath, M., Riesch, R., Schlupp, I., Grasse, A., Munimanda, G.K., Setzer, C., Penn,
961 D.J., and Moodley, Y. (2014). Selection from parasites favours immunogenetic diversity but not
962 divergence among locally adapted host populations. J. Evol. Biol. *27*, 960–974.
963 10.1111/jeb.12370.

964 55. Hotaling, S., Quackenbush, C.R., Bennett-Ponsford, J., New, D.D., Arias-Rodriguez, L.,
965 Tobler, M., and Kelley, J.L. (2019). Bacterial Diversity in Replicated Hydrogen Sulfide-Rich
966 Streams. Microb. Ecol. *77*, 559–573. 10.1007/s00248-018-1237-6.

967 56. Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the
968 microbiota and the immune system. Science *336*, 1268–1273. 10.1126/science.1223490.

969 57. Zhang, F., and Broughton, R.E. (2013). Mitochondrial-nuclear interactions: compensatory
970 evolution or variable functional constraint among vertebrate oxidative phosphorylation genes?
971 Genome Biol. Evol. *5*, 1781–1791.

972 58. Olson, K.R. (2018). H_2S and polysulfide metabolism: Conventional and unconventional
973 pathways. Biochem. Pharmacol. *149*, 77–90. 10.1016/j.bcp.2017.12.010.

974 59. Stipanuk, M.H., and Ueki, I. (2011). Dealing with methionine/homocysteine sulfur: cysteine
975 metabolism to taurine and inorganic sulfur. J. Inherit. Metab. Dis. *34*, 17–32. 10.1007/s10545-
976 009-9006-9.

977 60. Yeaman, S. (2015). Local Adaptation by Alleles of Small Effect. Am. Nat. *186 Suppl 1*, S74-89.
978 10.1086/682405.

979 61. Simons, Y.B., Bullaughey, K., Hudson, R.R., and Sella, G. (2018). A population genetic
980 interpretation of GWAS findings for human quantitative traits. PLoS Biol. *16*, e2002985.
981 10.1371/journal.pbio.2002985.

982 62. Whiting, J.R., Paris, J.R., van der Zee, M.J., and Fraser, B.A. (2022). AF-vapeR : A multivariate
983 genome scan for detecting parallel evolution using allele frequency change vectors. Methods
984 Ecol. Evol. 10.1111/2041-210x.13952.

985 63. Barghi, N., Tobler, R., Nolte, V., Jakšić, A.M., Mallard, F., Otte, K.A., Dolezal, M., Taus, T.,
986 Kofler, R., and Schlötterer, C. (2019). Genetic redundancy fuels polygenic adaptation in
987 Drosophila. PLoS Biol. *17*, e3000128. 10.1371/journal.pbio.3000128.

988 64. Barghi, N., Hermisson, J., and Schlötterer, C. (2020). Polygenic adaptation: a unifying
989 framework to understand positive selection. Nat. Rev. Genet. *21*, 769–781. 10.1038/s41576-
990 020-0250-z.

991 65. Láruson, Á.J., Yeaman, S., and Lotterhos, K.E. (2020). The importance of genetic redundancy
992 in evolution. Trends Ecol. Evol. *35*, 809–822. 10.1016/j.tree.2020.04.009.

993 66. Steiner, C.C., Römpler, H., Boettger, L.M., Schöneberg, T., and Hoekstra, H.E. (2009). The
994 genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different
995 genes. Mol. Biol. Evol. *26*, 35–45. 10.1093/molbev/msn218.

996 67. Mack, K.L., and Nachman, M.W. (2017). Gene Regulation and Speciation. Trends Genet. *33*,
997 68–80. 10.1016/j.tig.2016.11.003.

998 68. Hebert, F.O., Renaut, S., and Bernatchez, L. (2013). Targeted sequence capture and
999 resequencing implies a predominant role of regulatory regions in the divergence of a sympatric
1000 lake whitefish species pair (Coregonus clupeaformis). Mol. Ecol. *22*, 4896–4914.
1001 10.1111/mec.12447.

1002 69. Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A.,
1003 Peichel, C.L., Saetre, G.P., Bank, C., Brannstrom, A., et al. (2014). Genomics and the origin of
1004 species. Nat. Rev. Genet. *15*, 176–192. 10.1038/nrg3644.

1005 70. Ishikawa, A., Kusakabe, M., Yoshida, K., Ravinet, M., Makino, T., Toyoda, A., Fujiyama, A.,
1006 and Kitano, J. (2017). Different contributions of local- and distant-regulatory changes to
1007 transcriptome divergence between stickleback ecotypes. Evolution *71*, 565–581.
1008 10.1111/evo.13175.

1009 71. Kautt, A.F., Kratochwil, C.F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F.,
1010 Torres-Dowdall, J., Härer, A., Hulsey, C.D., Franchini, P., et al. (2020). Contrasting signatures
1011 of genomic divergence during sympatric speciation. Nature *588*, 106–111. 10.1038/s41586-
1012 020-2845-0.

1013 72. Flaxman, S.M., Wacholder, A.C., Feder, J.L., and Nosil, P. (2014). Theoretical models of the
1014 influence of genomic architecture on the dynamics of speciation. Mol. Ecol. *23*, 4074–4088.
1015 10.1111/mec.12750.

1016 73. Renaut, S., Grassa, C.J., Yeaman, S., Moyers, B.T., Lai, Z., Kane, N.C., Bowers, J.E., Burke,
1017 J.M., and Rieseberg, L.H. (2013). Genomic islands of divergence are not affected by geography
1018 of speciation in sunflowers. Nat. Commun. *4*, 1827.

1019 74. Han, F., Lamichhaney, S., Grant, B.R., Grant, P.R., Andersson, L., and Webster, M.T. (2017).
1020 Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of
1021 divergence among Darwin's finches. Genome Res. *27*, 1004–1015. 10.1101/gr.212522.116.

1022 75. Feder, J.L., Nosil, P., Wacholder, A.C., Egan, S.P., Berlocher, S.H., and Flaxman, S.M. (2014).
1023 Genome-wide congealing and rapid transitions across the speciation continuum during
1024 speciation with gene flow. J. Hered. *105 Suppl 1*, 810–820. 10.1093/jhered/esu038.

1025 76. Greenway, R., Arias-Rodriguez, L., Diaz, P., and Tobler, M. (2014). Patterns of
1026 macroinvertebrate and fish diversity in freshwater sulphide springs. Diversity *6*, 597–632.
1027 10.3390/d6030597.

1028 77. Palacios, M., Arias-Rodriguez, L., Plath, M., Eifert, C., Lerp, H., Lamboj, A., Voelker, G., and
1029 Tobler, M. (2013). The rediscovery of a long described species reveals additional complexity in

1030 speciation patterns of poeciliid fishes in sulfide springs. PLoS One *8*, e71069.
1031 10.1371/journal.pone.0071069.

1032 78. Álvarez del Villar, J. (1948). Descripción de una nueva especie de *Mollienisia* capturada en Baños
1033 del Azufre, Tabasco (Pisces, Poeciliidae). Anales de la Escuela Nacional de Ciencias Biológicas
1034 *5*, 275–281.

1035 79. Rohlf, F.J. (2004). tpsDig. Available from http://life.bio.sunysb.edu/morph/.

1036 80. Zelditch, M.L., Swiderski, D.L., and David Sheets, H. (2012). Geometric Morphometrics for
1037 Biologists: A Primer (Academic Press).

1038 81. Adams, D. C., M. L. Collyer, A. Kaliontzopoulou, and E.K. Balken (2021). Geomorph:
1039 Software for geometric morphometric analyses.

1040 82. Langerhans, R.B. (2009). Trade-off between steady and unsteady swimming underlies predator-
1041 driven divergence in Gambusia affinis. J. Evol. Biol. *22*, 1057–1075.

1042 83. Klingenberg, C.P., and Spence, J.R. (1993). Heterochrony and allometry: lessons from the
1043 water strider genus Limnoporus. Evolution *47*, 1834–1853.

1044 84. Butler, I.B., Schoonen, M.A.A., and Rickard, D.T. (1994). Removal of dissolved oxygen from
1045 water: a comparison of four common techniques. Talanta *41*, 211–215.

1046 85. Therneau, T. (2021). A package for survival analysis in R.

1047 86. Kassambara, A., Kosinski, M., and Biecek, P. (2021). survminer: drawing survival curves using
1048 "ggplot2."

1049 87. Evans, D.H., and Claiborne, J.B. (2006). The physiology of fishes. In (Taylor & Francis).

1050 88. Krueger, F. (2014). Trim Galore!
1051 http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

1052 89. Schartl, M., Walter, R.B., Shen, Y., Garcia, T., Catchen, J., Amores, A., Braasch, I., Chalopin,
1053 D., Volff, J.-N., Lesch, K.-P., et al. (2013). The genome of the platyfish, Xiphophorus
1054 maculatus, provides insights into evolutionary adaptation and several complex traits. Nat.
1055 Genet. *45*, 567–572. 10.1038/ng.2604.

1056 90. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
1057 transform. Bioinformatics *25*, 1754–1760. 10.1093/bioinformatics/btp324.

1058 91. Genomics, I.S.U. GFF2FASTA.

1059 92. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden,
1060 T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.
1061 10.1186/1471-2105-10-421.

1062 93. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T., and Salzberg, S.L.
1063 (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads.
1064 Nat. Biotechnol. *33*, 290–295. 10.1038/nbt.3122.

1065 94. Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level
1066 expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat.
1067 Protoc. *11*, 1650–1667. 10.1038/nprot.2016.095.

1068 95. Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T.,
1069 Maechler, M., Magnusson, A., Moeller, S., et al. (2009). gplots: various R programming tools
1070 for plotting data.

1071 96. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for
1072 differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.
1073 10.1093/bioinformatics/btp616.

1074 97. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and
1075 powerful approach to multiple testing. J. R. Stat. Soc. *57*, 289–300. 10.1111/j.2517-
1076 6161.1995.tb02031.x.

1077 98. Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis,
1078 S., Marshall, B., Mungall, C., et al. (2004). The Gene Ontology (GO) database and informatics
1079 resource. Nucleic Acids Res. *32*, D258-61. 10.1093/nar/gkh036.

1080 99. Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for
1081 discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*,
1082 48. 10.1186/1471-2105-10-48.

1083 100. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella,
1084 K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a
1085 MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*,
1086 1297–1303.

1087 101. Broad Institute (2018). Picard Toolkit.

1088 102. DePristo, M., Banks, E., Poplin, R., Garimella, K., Maguire, J., Hartl., C., Philippakis, A., del
1089 Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and
1090 genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.

1091 103. Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine,
1092 A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high
1093 confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Protoc.
1094 Bioinformatics *43*, 11.10.1-11.10.33. 10.1002/0471250953.bi1110s43.

1095 104. Korunes, K.L., and Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and
1096 divergence in the presence of missing data. Mol. Ecol. Resour. *21*, 1359–1368. 10.1111/1755-
1097 0998.13326.

1098 105. Raj, A., Stephens, M., and Pritchard, J.K. (2014). fastSTRUCTURE: variational inference of
1099 population structure in large SNP data sets. Genetics *197*, 573–589.
1100 10.1534/genetics.114.164350.

1101 106. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-
1102 generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.
1103 10.1186/s13742-015-0047-8.

1104 107. Tobler, M., Riesch, R., Tobler, C.M., Schulz-Mirbach, T., and Plath, M. (2009). Natural and
1105 sexual selection against immigrants maintains differentiation among micro-allopatric
1106 populations. J. Evol. Biol. *22*, 2298–2304. 10.1111/j.1420-9101.2009.01844.x.

1107 108. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models
1108 Using lme4. Journal of Statistical Software, Articles *67*, 1–48. 10.18637/jss.v067.i01.

1109 109. Dierckxsens, N., Mardulyn, P., and Smits, G. (2017). NOVOPlasty: de novo assembly of
1110 organelle genomes from whole genome data. Nucleic Acids Res. *45*, e18. 10.1093/nar/gkw955.

1111 110. Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T.P., Sado, T.,
1112 Mabuchi, K., Takeshima, H., Miya, M., et al. (2013). MitoFish and MitoAnnotator: a
1113 mitochondrial genome database of fish with an accurate and automatic annotation pipeline.
1114 Mol. Biol. Evol. *30*, 2531–2540. 10.1093/molbev/mst141.

1115 111. Katoh, K., Kuma, K.-I., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in
1116 accuracy of multiple sequence alignment. Nucleic Acids Res. *33*, 511–518. 10.1093/nar/gki198.

1117 112. Abascal, F., Zardoya, R., and Telford, M.J. (2010). TranslatorX: multiple alignment of
1118 nucleotide sequences guided by amino acid translations. Nucleic Acids Res. *38*, W7-13.
1119 10.1093/nar/gkq291.

1120 113. Leigh, J.W., and Bryant, D. (2015). Popart : Full-feature software for haplotype network
1121 construction. Methods Ecol. Evol. *6*, 1110–1116. 10.1111/2041-210x.12410.

1122 114. Clement, M., Snell, Q., Walke, P., Posada, D., and Crandall, K. (2002). TCS: estimating gene
1123 genealogies. In Proceedings 16th International Parallel and Distributed Processing Symposium
1124 (IEEE). 10.1109/ipdps.2002.1016585.

1125 115. Mapping to EVE (Github).

1126 116. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,
1127 Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence
1128 Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
1129 10.1093/bioinformatics/btp352.

1130 117. Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics *5*, 164–
1131 166.

1132 118. Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*,
1133 1586–1591.

1134 119. Zhang, J., Nielsen, R., and Yang, Z. (2005). Evaluation of an improved branch-site likelihood
1135 method for detecting positive selection at the molecular level. Mol. Biol. Evol. *22*, 2472–2479.

1136 120. Martin, S.H., Davey, J.W., and Jiggins, C.D. (2015). Evaluating the use of ABBA-BABA
1137 statistics to locate introgressed loci. Mol. Biol. Evol. *32*, 244–257. 10.1093/molbev/msu269.

1138 121. Martin, S.H., Singh, K.S., Gordon, I.J., Omufwoko, K.S., Collins, S., Warren, I.A., Munby, H.,
1139 Brattström, O., Traut, W., Martins, D.J., et al. (2020). Whole-chromosome hitchhiking driven
1140 by a male-killing endosymbiont. PLoS Biol. *18*, e3000610. 10.1371/journal.pbio.3000610.

1141 122. Martin, S. (2020). General tools for genomic analyses.

1142 123. Delmore, K.E., Lugo Ramos, J.S., Van Doren, B.M., Lundberg, M., Bensch, S., Irwin, D.E.,
1143 and Liedvogel, M. (2018). Comparative analysis examining patterns of genomic differentiation
1144 across multiple episodes of population divergence in birds. Evol Lett *2*, 76–87.
1145 10.1002/evl3.46.

1146 124. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing
1147 genomic features. Bioinformatics *26*, 841–842. 10.1093/bioinformatics/btq033.

1148

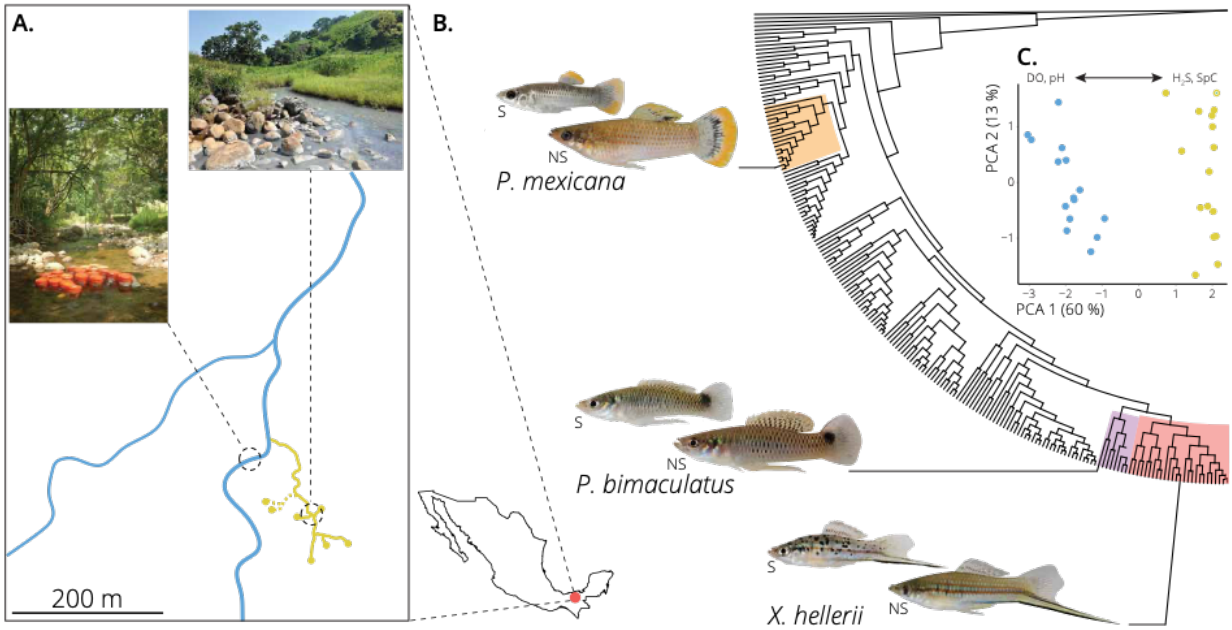

Figure 1. Sympatric populations of three highly divergent poeciliid lineages co-occur in a sulfide spring and adjacent nonsulfidic stream in southern Mexico. A. Map of the study site, where yellow points indicate the H_2S spring heads of the La Gloria spring system, while yellow lines indicate H_2S-rich stream segments. Blue lines indicate the adjacent nonsulfidic stream system, Arroyo Caracol. Inset photos depict representative habitats within the two streams (top: La Gloria, sulfidic; left: Arroyo Caracol, nonsulfidic). B. Phylogeny of the family Poeciliidae highlighting the three focal genera: *Poecilia* (orange), *Pseudoxiphophorus* (purple), and *Xiphophorus* (red). Each species' position within its genus is denoted by the extended branch leading to representative photographs of a sulfidic (S; top) and nonsulfidic (NS; bottom) male of each species. C. Nonsulfidic and sulfidic habitats differed significantly in physical and chemical water parameters. Principal component analysis indicated that sites within the La Gloria spring complex (yellow) and Arroyo Caracol (blue) separated along the primary axis of variation, with nonsulfidic sites exhibiting higher dissolved oxygen concentrations (DO) and pH and sulfidic sites exhibiting higher H_2S concentrations and specific conductivities (SpC).

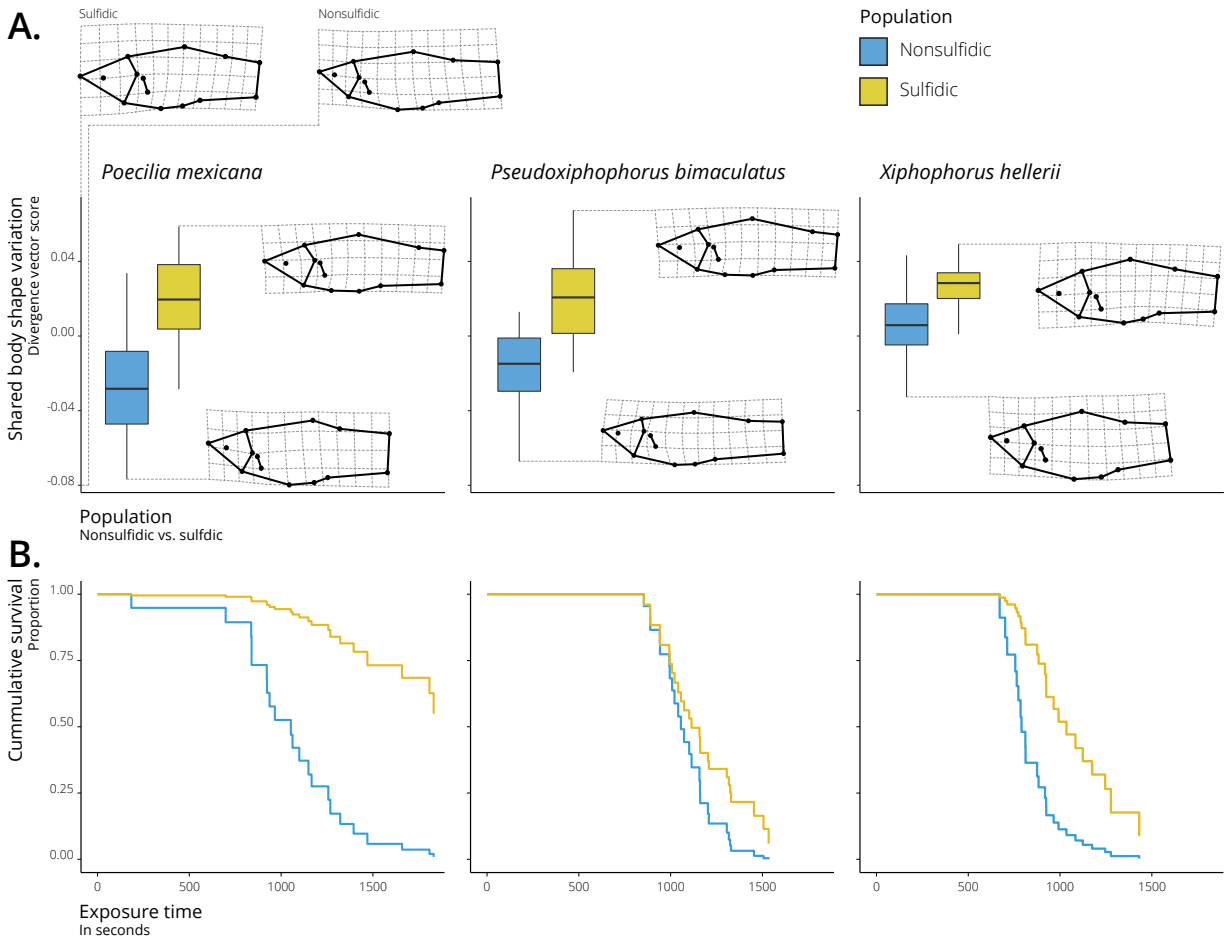

Figure 2. Phenotypic convergence in morphological and physiological traits of sympatric sulfidic populations. A. Convergent morphological evolution among species between habitat types. Body shape differentiation along a shared axis of divergence (habitat type) is depicted by the transformation grids on the top. Significant differences in body shape between populations of each species is depicted in insets (Table S1). B. Convergence in H_2S tolerance among all populations from the sulfidic habitat. Cumulative survival in acute H_2S toxicity trials revealed significant differences between populations, which were most pronounced in *P. mexicana* (Figure S1).

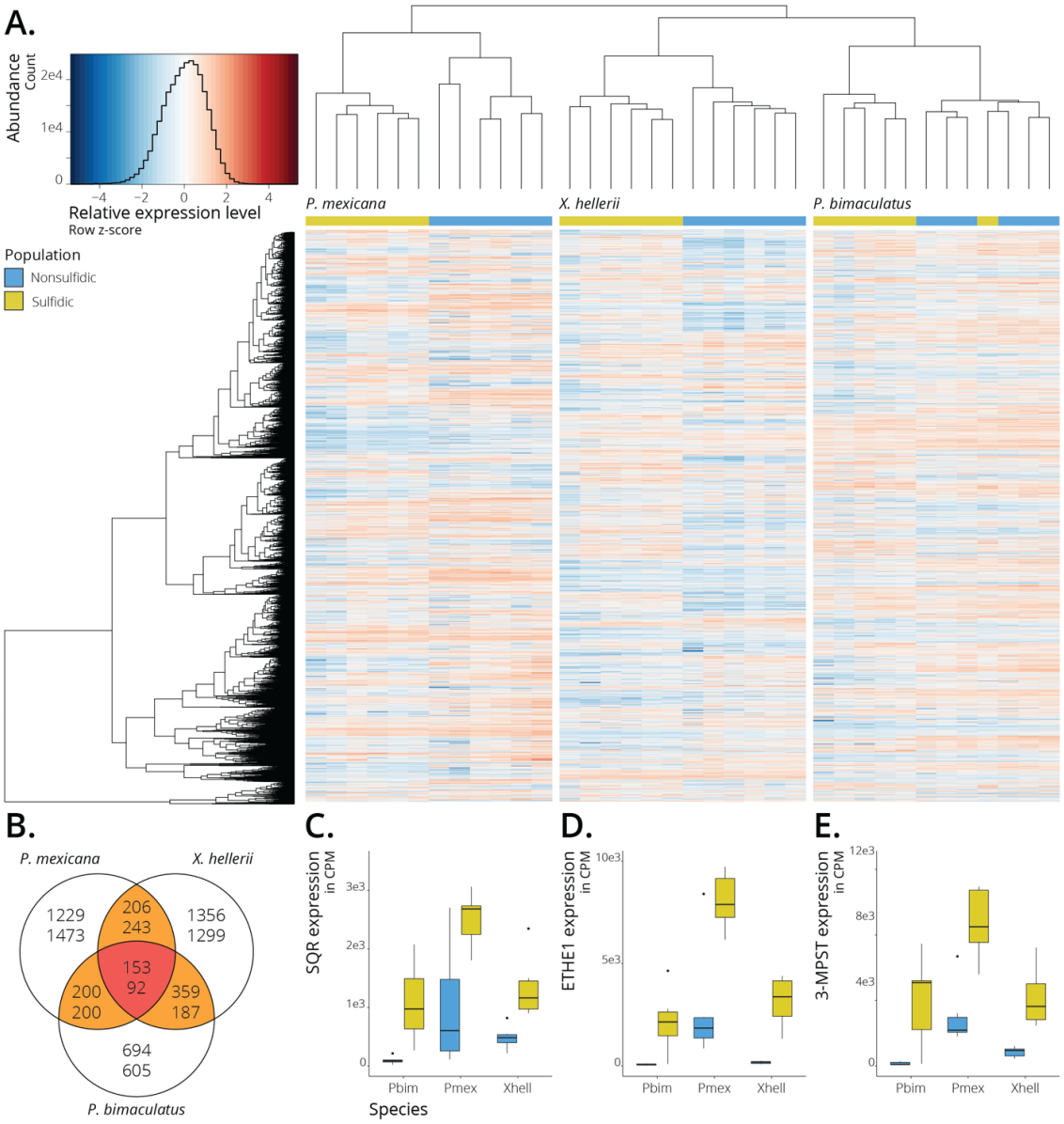

Figure 3. Convergent shifts in gene expression underlie local adaptation. A. Hierarchical clustering analysis indicates that gene expression profiles cluster by habitat type within each species. Each row represents expression data for one gene, and each column corresponds to an individual. Genes with higher relative expression in individuals from sulfidic habitats are indicated in red, genes with lower relative expression in blue. B. Venn diagram showing the number of differentially expressed genes between the sulfidic and nonsulfidic populations of each species and those shared among species. Top numbers are genes upregulated in the sulfidic habitat; bottom numbers are downregulated genes. C-E. Sulfidic populations consistently exhibit higher expression of key H_2S detoxification genes (SQR: sulfide:quinone oxidoreductase; ETHE1: persulfide dioxygenase; 3-MPST: 3-

1183 mercaptopyruvate sulfurtransferase) than nonsulfidic populations of the same species. Box plots of
1184 counts per million (CPM) by population. All boxplots have the following elements: center line,
1185 median; box limits, upper and lower quartiles; whiskers, ×1.5 interquartile range; points, outliers
1186 closed circles.
1187

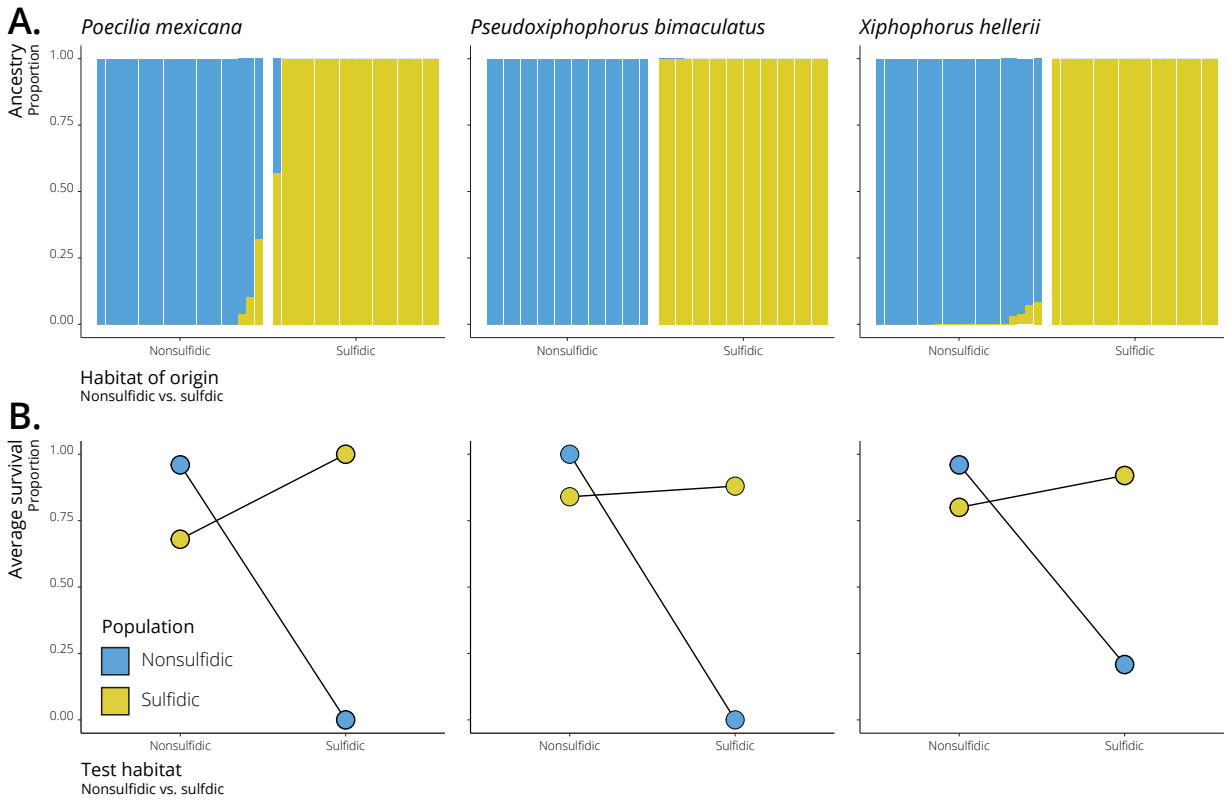

Figure 4. Nonsulfidic and sulfidic populations of three sympatric species exhibit limited gene flow between habitats mediated by strong natural selection against migrants. A. Admixture analyses revealed substantial population genetic differentiation in each population pair. Individuals are grouped by sampling site, corresponding with the detected genetic population clusters (*K*=2) shown in blue (nonsulfidic) and yellow (sulfidic). B. Translocation experiments revealed strong selection against migrants, especially those from nonsulfidic to sulfidic habitats (Table S5). Fish were transferred between sites (test habitat) to estimate survival of migrants between habitats, with individuals collected from the nonsulfidic population in blue and from the sulfidic population in yellow.

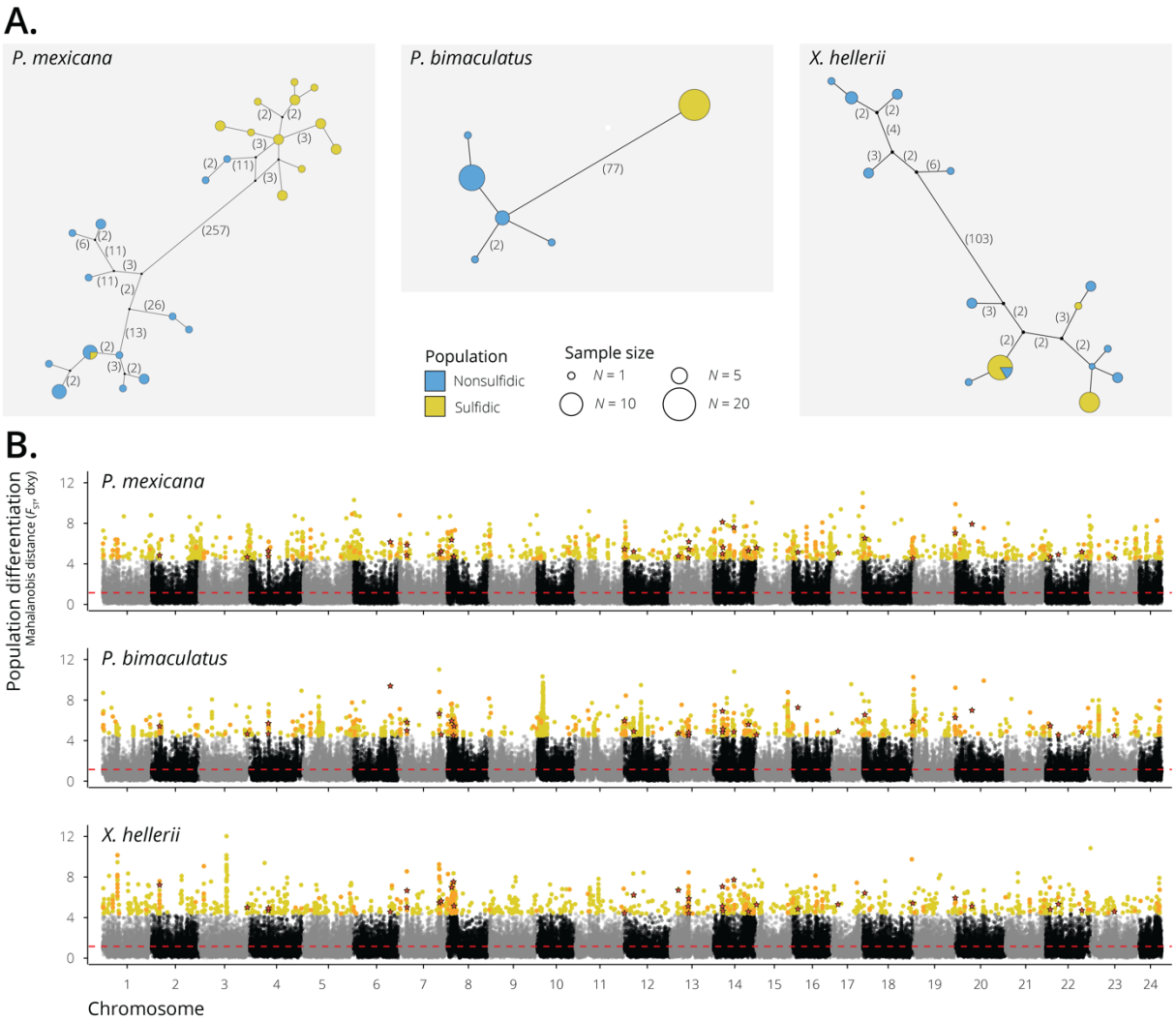

Figure 5. Convergent patterns of genomic adaptation are limited to the mitochondrial genome and few nuclear loci. A. Haplotype networks based on concatenated mitochondrial genes indicate high levels of mitochondrial divergence between nonsulfidic (blue) and sulfidic (yellow) populations, with asymmetrical introgression of sulfidic mitochondria into nonsulfidic individuals. The relative size of each circle indicates the number of individuals with a particular mitochondrial haplotype, while numbers in parentheses represent the total number of nucleotide differences between haplotypes. B. Mahalanobis distance ($-\log_{10}$ *P*-values of F_{ST} and d_{XY}) of non-overlapping 5-kb windows across the genome as a measurement of genomic divergence driven by natural selection. Outlier windows (top 1%) unique to each population pair are shown in yellow, while those shared between two population pairs are in orange. The outlier windows shared among all three population pairs are denoted with red stars.