Molecular evolution and expression of oxygen transport genes in livebearing fishes (Poeciliidae) from hydrogen sulfide rich springs

Journal:	<i>Genome</i>
Manuscript ID	gen-2017-0051.R1
Manuscript Type:	Article
Date Submitted by the Author:	21-Oct-2017
Complete List of Authors:	Barts, Nick; Kansas State University Greenway, Ryan; Kansas State University Passow, Courtney; University of Minnesota System ARIAS-RODRIGUEZ, LENIN; Universidad Juarez Autonoma de Tabasco Kelley, Joanna; Washington State University College of Arts and Sciences Tobler, Michael; Kansas State University,
Is the invited manuscript for consideration in a Special Issue? :	Ecological Genomics (closed)
Keyword:	adaptation, convergent evolution, hemoglobin, molecular evolution

SCHOLARONE™
Manuscripts

**Molecular evolution and expression of oxygen transport genes in
livebearing fishes (Poeciliidae) from hydrogen sulfide rich springs**

Nicholas Barts¹, Ryan Greenway¹, Courtney N. Passow², Lenin Arias-Rodriguez³, Joanna L. Kelley⁴,
and Michael Tobler^{1*}

¹Division of Biology, Kansas State University, 116 Ackert Hall, Manhattan, KS 66506, USA

²Ecology, Evolution and Behavior, University of Minnesota St. Paul 205 Cargill Building, St. Paul,
MN 55108, USA

³División Académica de Ciencias Biológicas, Universidad Juárez Autónoma de Tabasco (UJAT),
C.P. 86150, Villahermosa, Tabasco, México

⁴Department of Biological Sciences, Washington State University, 431 Heald Hall, Pullman, WA
99164, USA

*Corresponding author: tobler@ksu.edu

ABSTRACT

Hydrogen sulfide (H₂S) is a natural toxicant in some aquatic environments that has diverse molecular targets. It binds to oxygen transport proteins, rendering them non-functional by reducing oxygen-binding affinity. Hence, organisms permanently inhabiting H₂S-rich environments are predicted to exhibit adaptive modifications to compensate for the reduced capacity to transport oxygen. We investigated ten lineages of fish of the family Poeciliidae that have colonized freshwater springs rich in H₂S – along with related lineages from non-sulfidic environments – to test hypotheses about the expression and evolution of oxygen transport genes in a phylogenetic context. We predicted shifts in the expression of and signatures of positive selection on oxygen transport genes upon colonization of H₂S-rich habitats. Our analyses indicated significant shifts in gene expression for multiple hemoglobin genes in lineages that have colonized H₂S-rich environments, and three hemoglobin genes exhibited relaxed selection in sulfidic compared to non-sulfidic lineages. However, neither changes in gene expression nor signatures of selection were consistent among all lineages in H₂S-rich environments. Oxygen transport genes may consequently be predictable targets of selection during adaptation to sulfidic environments, but changes in gene expression and molecular evolution of oxygen transport genes in H₂S-rich environments are not necessarily repeatable across replicated lineages.

Keywords: adaptation; convergent evolution; hemoglobin; molecular evolution; myoglobin; sulfide springs

INTRODUCTION

Whether and under what circumstances evolution is repeatable and predictable remains a core question in biology. Phenotypic convergence is frequently invoked as evidence for the repeatability and predictability of evolution (Endler 1986; Schluter 2000; Vermeij 2006). Indeed, phenotypic convergence has been documented across broad phylogenetic scales and in response to a wide variety of sources of selection (e.g., Landry and Bernatchez 2010; Losos 2009; Wilkens and Strecker 2003; Winemiller et al. 1995). Nonetheless, convergent evolution is not ubiquitous, and disparate lineages can respond to shared evolutionary problems with unique adaptive solutions (Gould and Lewontin 1979; Kaeuffer et al. 2012; Langerhans and DeWitt 2004). Historical, ecological, genetic, and functional factors can all contribute to the degree of shared (convergent) *versus* unique (non-convergent) evolutionary changes that are observed when independent lineages are exposed to the same sources of selection (Blount et al. 2008; Kaeuffer et al. 2012; Stern 2013). But even when phenotypic convergence is common, convergent functional outcomes frequently arise from non-convergent modifications at lower hierarchical levels of biological organization (Elmer and Meyer 2011; Losos 2011), which suggests that multiple molecular changes can result in equivalent function and fitness (Natarajan et al. 2016). This begets questions about the circumstances under which we can actually anticipate convergent changes at a molecular level (Rosenblum et al. 2014). Evolutionary repeatability and predictability at a molecular level may be particularly high if constraints limit the number of solutions that maximize fitness under given environmental conditions. Such constraints may be evident when populations are exposed to sources of selection with clear-cut biochemical targets and physiological consequences (Tobler and Culumber 2016). Indeed, some of the best examples of repeated modifications of the same molecular pathways – and even genes – come from systems that are exposed to physiochemical stressors. For example, killifish have repeatedly evolved loss of function mutations in aryl hydrocarbon receptor signaling pathway to cope with

anthropogenic pollution in the form of aromatic hydrocarbons (Reid et al. 2016), exposure to secondary plant metabolites lead to the repeated modification of sodium-potassium pumps in insects (Dobler et al. 2012), snakes have repeatedly evolved tetrodotoxin-resistant sodium channels to cope with prey toxicity (Feldman et al. 2012), and poison frogs have repeatedly modified sodium channels to tolerate their own toxins (Tarvin et al. 2016). Consequently, natural systems in which selection is mediated by biochemically and physiologically explicit stressors provide excellent models to test for evolutionary repeatability and predictability at a molecular level.

Hydrogen sulfide (H_2S) is a naturally occurring toxicant with well-documented biochemical and physiological effects (Fukuto et al. 2012; Kabil and Banerjee 2010). The primary toxic effect of H_2S is the cessation of aerobic respiration. H_2S readily and reversibly binds to the heme group of cytochrome c oxidase (COX, complex IV of the mitochondrial respiratory chain), inhibiting aerobic ATP production (Cooper and Brown 2008; Khan 1990). Hence, exposure to H_2S is lethal in micromolar concentrations for most metazoans (Bagarinao 1992; Beauchamp et al. 1984). Nonetheless, animals can enzymatically oxidize and detoxify H_2S at low concentrations through the highly conserved sulfide:quinone oxidoreductase (SQR) pathway (Hildebrandt and Grieshaber 2008). Organisms living in naturally H_2S -rich environments – such as deep sea hydrothermal vents, cold seeps, anoxic sediments, and freshwater sulfide springs – are consequently expected to either exhibit modifications of direct toxicity targets or enzymes associated with detoxification (Tobler et al. 2016). Indeed, H_2S -resistant COX (Martin and Maricle 2015; Pfenninger et al. 2014) and modification of enzymes associated with the SQR pathway (Huang et al. 2013; Kelley et al. 2016; Liu et al. 2015; Ma et al. 2012; Pfenninger et al. 2015) have been documented in various taxa.

Besides COX and detoxification enzymes, H_2S has a variety of other putative molecular targets (Bagarinao 1992; Olson 2011). In particular, H_2S can bind to heme proteins involved in oxygen transport, including myoglobin and hemoglobin (Pietri et al. 2011). H_2S effectively reduces

the porphyrin ring within the globin structure, forming sulfmyoglobin and sulfhemoglobin (Grieshaber and Völkel 1998). Sulfmyoglobin and sulfhemoglobin are characterized by drastically reduced oxygen binding affinity, rendering them almost non-functional for oxygen transport (Evans et al. 1994; Johnson 1970). Consequently, theory predicts that myoglobin and hemoglobin proteins are likely targets of adaptation to chronic H₂S exposure. Several marine invertebrates that inhabit H₂S-rich environments (e.g., the bivalve *Lucina pectinata* and the giant tubeworm *Riftia pachyptila*) have evolved specialized heme proteins that function in H₂S binding and are critical for supplying symbiotic sulfur-oxidizing microbes with resources (Flores et al. 2005; Kraus and Wittenberg 1990; Zal et al. 1998). In addition, a polychaete worm (*Methanoaricia dendrobranchiata*) that inhabits environments with low oxygen and high H₂S concentrations along hydrocarbon seeps exhibits hemoglobin with exceptionally high oxygen binding affinity and a pronounced Bohr effect (Hourdez et al. 2002). However, it largely remains untested whether hemoglobin and myoglobin are repeatable and predictable targets of selection across a diversity of lineages that have colonized H₂S-rich environments.

We investigated multiple lineages of fish of the family Poeciliidae that have independently colonized freshwater springs rich in H₂S in parts of North and Latin America, as well as the Caribbean (Greenway et al. 2014). These sulfide springs exhibit continuously high concentrations of H₂S and low concentrations of oxygen (Tobler and Plath 2011). Fish populations in sulfide springs are locally adapted and characterized by divergence in complex phenotypic traits from ancestral populations in non-sulfidic environments (Riesch et al. 2015). Trait evolution in independent sulfide spring lineages is governed by patterns of convergence in molecular (Kelley et al. 2016; Pfenninger et al. 2015), physiological (Passow et al. in press; Pfenninger et al. 2014), morphological (Tobler and Hastings 2011; Tobler et al. 2011), and life history traits (Riesch et al. 2010; Riesch et al. 2014). Furthermore, sulfide spring populations are often reproductively isolated from adjacent, non-sulfidic

populations, even when physical barriers that prevent fish movement are absent (Plath et al. 2013; Riesch et al. 2016).

Here, we collected samples of sulfide spring poeciliids from a broad geographic and phylogenetic range (Fig. 1) in order to test whether evolution in H₂S-rich environments has led to the repeatable and predictable modification of oxygen transport genes. We sought to address two broad questions: (1) Is there evidence for potential shifts in the expression of myoglobin and hemoglobin genes associated with the colonization of H₂S-rich habitats? Both the presence of H₂S and the rampant hypoxia in sulfide springs could cause differential transcription of oxygen transport genes either plastically or through evolved changes in gene regulation. If H₂S reduces the functionality of hemoglobin and myoglobin in oxygen transport and storage (Evans et al. 1994; Johnson 1970), organisms in sulfidic habitats may compensate through up-regulation of associated genes. Similarly, up-regulation of hemoglobins and myoglobin could also be related to an increased capacity for H₂S sequestration and oxidation that might mitigate toxic effects (Vivitsky et al. 2015). Finally, variation in ambient oxygen concentrations have been documented to induce complex transcriptional changes in oxygen transport genes (Roesner et al. 2006), and hypoxia can be associated with increased gene expression of hemoglobins and myoglobin (Fraser et al. 2006; Richards 2009; van der Meer et al. 2005; Wawrowski et al. 2011). (2) Is there evidence for positive selection on oxygen transport genes, and are there specific codons that have been the target of selection across different lineages of sulfide spring fishes? Selection on oxygen transport genes is expected to shape the proteins' oxygen binding affinity in a way that simultaneously optimizes oxygen uptake from the environment and its delivery to mitochondria in tissues where oxygen is consumed by aerobic metabolism. Selection may also act to shape the proteins' propensity to interact with allosteric effectors – like H₂S – that affect their function (Storz et al. 2010; Weber 2007; Weber and Fago 2004). Theoretical considerations suggest that adaptive modifications of oxygen

transport proteins should be restricted to specific amino acids that are involved in heme binding, inter-subunit contacts, or interaction with allosteric effectors, which collectively shape oxygen-binding affinity (Natarajan et al. 2016; Perutz 1983). Hence, if there is positive selection on oxygen transport genes in H₂S-rich environments, we predict convergent modifications of specific codons in replicated lineages inhabiting sulfide spring environments. We addressed these questions by analyzing the expression levels and molecular evolution of hemoglobin and myoglobin genes in an explicit phylogenetic context, including ten lineages from sulfide springs and ten reference lineages from non-sulfidic habitats.

MATERIALS AND METHODS

Sample Collection

To analyze variation in oxygen transport gene expression and evolution, we collected specimens from multiple lineages of poeciliid fishes that have independently colonized H₂S-rich springs in the United States, Mexico, and the island of Hispaniola, as well as from geographically and phylogenetically proximate lineages in non-sulfidic habitats (Fig. 1; Greenway et al. 2014). This approach led to sets of closely related lineages found in sulfidic and non-sulfidic habitats, including populations of *Poecilia latipinna* and *Gambusia holbrooki* in Florida; populations of the *Poecilia mexicana* species complex (including *Poecilia mexicana*, *Poecilia sulphuraria*, and *Poecilia limantouri*) (Palacios et al. 2016), the *Gambusia sexradiata* species complex (including *G. sexradiata* and *Gambusia eurystoma*) (Jourdan et al. 2016), *Pseudoxiphophorus bimaculatus*, and *Xiphophorus hellerii* in Mexico; as well as populations of the *Limia perugiae* species complex (*L. perugiae* and *Limia sulphurophila*) (Weaver et al. 2016) in the Dominican Republic (see Table 1 for localities and habitat associations of specific lineages). Note that data for some lineages of the *P. mexicana* species complex were reanalyzed from a previous study (GenBank BioProject accession number: PRJNA290391; Kelley et al. 2016). In

addition, we used published genomic data from *Xiphophorus maculatus* (for mapping of RNA-seq reads; GenBank accession number: AGAJ000000000.1; Scharl et al. 2013) and *Fundulus heteroclitus* (as an outgroup in phylogenetic analyses; GenBank accession number: JXMV000000000.1; Waits et al. 2016).

All fish were caught using a seine (2 × 5 meters; 3 mm mesh size). Immediately upon capture, adult females ($N=5-6$ per site; Table 1) were euthanized, and the gills were extracted from both sides of the body using sterilized scissors and forceps. Tissues were then preserved in RNAlater (Ambion Inc.) for subsequent analysis in the laboratory. We focused on gill tissues, because they are in direct contact with the toxic environment (Tobler et al. 2016), mediate a variety of physiological processes involved in the maintenance of homeostasis (Evans and Claiborne 2006; Evans et al. 2011), and exhibit strong transcriptional responses upon exposure to H_2S (Kelley et al. 2016; Tobler et al. 2014). Procedures for these experiments were approved by the Institutional Animal Care and Use Committee at Kansas State University (Protocol #3418).

RNA extraction, cDNA library preparation, and sequencing

The general protocols for RNA extraction, library preparation and sequencing of all new samples closely followed procedures previously employed for *Poecilia* (Kelley et al. 2016; Kelley et al. 2012). In brief, 10-30 mg of gill tissues from each individual were sealed in a Covaris TT1 TissueTube (Covaris, Inc., Woburn, MA, USA), frozen in liquid nitrogen, and pulverized. Total RNA was extracted from pulverized tissue using the NucleoSpin RNA kit (Machery-Nagel, Düren, Germany) following the manufacturer's protocol. mRNA isolation and cDNA library preparation were conducted using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, Inc., Ipswich, MA, USA) and NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs, Inc., Ipswich, MA, USA) following the manufacturers' protocol with minor

modifications. All cDNA libraries were individually barcoded (so that reads could be linked back to a specific individual), quantified using Qubit and an Agilent 2100 Bioanalyzer High Sensitivity DNA chip, and pooled in sets of 11-12 samples based on nanomolar concentrations. Samples were split into pools such that samples from different species and habitat types were sequenced together. There was no evidence for significant lane effects. Libraries were sequenced on an Illumina HiSeq 2500 with paired-end 101 base pair (bp) reads at the Washington State University Spokane Genomics Core. Due to low coverage, some samples were re-sequenced, and reads from separate runs were combined for each sample for data analysis. After sequencing, all raw RNA-seq reads were sorted by barcode and trimmed twice (quality 0 to remove adapters and then quality 24) using the program Trimalore! (v. 0.4.0; Krueger 2014). Trimmed reads were then mapped to the *Xiphophorus maculatus* reference genome (Schartl et al. 2013) using BWA-MEM (v. 0.7.12; Li and Durbin 2009). Depending on the lineage, between 81.0 and 94.5 % of reads mapped to the reference genome (mean $\pm$ standard deviation: 87.9 ± 3.8 %). Note that mapping to an alternative reference genome (*Poecilia mexicana*) (as in Passow et al. 2017) yielded in qualitatively similar results, indicating that mapping biases among lineages did not play a major role in shaping analytical outcomes (results not shown).

Identifying focal and background genes

To identify focal genes, we extracted 11 hemoglobin and 1 myoglobin genes (Table 2) using the *X. maculatus* genome annotation file (GTF) (Schartl et al. 2013). For a subset of individuals in each population, the sequence for each gene was compared to the nr/nt database using BLASTn to confirm that the *X. maculatus* gene was within the set of top hits. For a subset of *P. latipinna* and *P. mexicana* individuals, each gene's sequence was compared to the respective available reference genomes (*P. latipinna* or *P. mexicana*) using BLASTn (Critical E-value: 0.001). The top hit in the

respective genome was then compared with BLASTn back against the *X. maculatus* genome to confirm the best hit for each gene. To generate a background set for reconstructing phylogenetic relationships and conducting analyses of gene expression patterns, we also extracted 171 genes representing functionally important, highly expressed loci for comparison with the focal oxygen transport genes (Table S1; Amsterdam et al. 2004; Calvo et al. 2015; Pagliarini et al. 2008; Shen et al. 2010; Warrington et al. 2000; Zhang and Broughton 2013).

Variant detection and consensus sequence generation

To detect variants and generate a consensus, we utilized the Genome Analysis Toolkit (GATK, v. 3.5; McKenna et al. 2010) on the mapped reads for indel realignment and variant detection (via the UnifiedGenotyper tool) across the coding regions of all samples simultaneously, following GATK best practices (Auwera et al. 2013; DePristo et al. 2011). Individuals were genotyped on a per-population basis using the EMIT_ALL_SITES output mode in the UnifiedGenotyper, creating a vcf file for each population. A custom script was used to fill positions in each vcf file that were missing sequence data with Ns. For each vcf, we used the setGT plugin to set all genotypes to the major allele and then created a consensus sequence across all individuals from each population with the consensus tool, as implemented in BCFtools (v.1.3; Li et al. 2009). The consensus sequences for the coding regions of each gene were combined, resulting in a single sequence file for each gene and population to be used in subsequent analyses. It is important to note that polymorphic sites identified in this manner are not necessarily fixed in a population, and differences between closely related populations may be due to differences in major allele frequencies (see results for details).

Phylogenetic analysis

To establish a phylogenetic framework for subsequent analyses, we conducted phylogenetic analyses using the consensus sequences from 167 background genes (Table S1; note that four genes were excluded due to alignment issues). Individual genes were aligned using ClustalW (v.2.1; Larkin et al. 2007). We used jModelTest 2 (v. 2.1.10; Darriba et al. 2012) to determine the best partition scheme based on the most likely model of DNA substitution among 88 candidate models, which was established by evaluating likelihood scores on a fixed BioNJ-JC tree under the Akaike Information Criterion (AIC) (Table S1). Genes were then concatenated and partitioned according to the substitution models indicated by jModelTest and then used in Maximum Likelihood analyses using RAxML (v. 8.2.9; Stamatakis 2014). We ran RAxML analyses under the Rapid Bootstrap Algorithm with 1,000 bootstrap replicates and used a thorough maximum likelihood (ML) search with each partition assigned its own GTR + Γ + I model. The best scoring ML tree was utilized for subsequent analyses of gene expression variation and molecular evolution. These analyses recovered a robust hypothesis for evolutionary relationships among the focal taxa (bootstrap support generally $\geq 99\%$, Fig. 1) that was consistent with previous phylogenetic studies of the family Poeciliidae (Hrbek et al. 2007; Palacios et al. 2016; Pollux et al. 2014). The only groups with moderately supported relationships included the two populations of *G. sexradiata* and *G. eurystoma* (bootstrap support 60%, Fig. 1), which have previously been found to represent a complex of closely related and recently diverged species (Jourdan et al. 2016).

Evolution of gene expression

To describe variation in gene expression among lineages, we quantified the relative number of RNA-seq reads that mapped to our focal and background genes (measured as fragments per kilobase of transcript per million mapped reads, FPKM) using cufflinks (v.2.2.1; Trapnell et al. 2010) run on each sample. FPKM values for each gene were extracted using a custom script, leading to a

measurement of gene expression for each individual ($N=5-6$ per lineage; see Table 1). Individual-level expression of focal and background genes were then used as dependent variables in Expression Variance and Evolution (EVE) models (Rohlf et al. 2013; Rohlf and Nielsen 2015), which implement Ornstein-Uhlenbeck models to conduct phylogenetic analysis of variance (ANOVA) to test for branch-specific shifts in gene expression associated with the colonization of H_2S -rich environments (Rohlf and Nielsen 2015). EVE models explicitly test for branch-specific shifts in gene expression by comparing likelihoods associated with an expression shift parameter (θ_i) for a gene being shared between two sets of lineages *vs.* θ_i being significantly different between them (Rohlf and Nielsen 2015). We tested for branch-specific expression shifts in two ways: (1) We tested whether colonization of H_2S -rich habitats was consistently associated with gene expression shifts by contrasting θ_i between all sulfidic and all non-sulfidic lineages (i.e., $\theta_i^{sulfidic} \neq \theta_i^{non-sulfidic}$). (2) We tested whether colonization of H_2S -rich habitats was associated with gene expression shifts in specific sulfidic lineages by contrasting θ_i between a foreground lineage (each sulfidic lineage as designated in Table 1) and all other lineages (i.e., $\theta_i^{foreground\ lineage} \neq \theta_i^{other\ lineages}$).

For statistical analyses, we calculated a likelihood ratio test (LRT_0) contrasting the null hypothesis (θ_i equal between groups) to the alternative hypotheses (θ_i unequal between groups) (Rohlf and Nielsen 2015). To minimize issues of multiple testing, we avoided significance testing using χ^2 -tests for each gene as suggested by Rohlf and Nielsen (2015). We instead generated empirical LRT_0 distributions including both oxygen-transport and background genes, deeming shifts in the focal oxygen transport genes significant if their LRT_0 values were in the top 5% of all values.

Quantifying positive selection

To test for positive selection on the focal myoglobin and hemoglobin genes associated with the colonization of H₂S-rich environments, we used branch models as implemented in codeml of the PAML package (v.4.9; Yang 2007), which evaluate variation in the rates of non-synonymous to synonymous substitutions (ω) across the phylogeny. We created multispecies codon alignments for each of our focal genes with the “codon” option in PRANK (v.150803; Löytynoja and Goldman 2005). For analysis in codeml, lineages from H₂S-rich habitats were set as foreground branches, and lineages from non-sulfidic habitats were set as background (Yang et al. 2000). For each gene, we compared a branch model ($\omega_{\text{non-sulfidic}} \neq \omega_{\text{sulfidic}}$) to a null model ($\omega_{\text{non-sulfidic}} = \omega_{\text{sulfidic}}$). *P*-values were generated based on a maximum likelihood ratio test with a χ^2 approximation (Zhang et al. 2005). In addition, we also conducted branch-site tests to evaluate evidence for positive selection on specific branches and codons (Yang and Nielsen 2002; Zhang et al. 2005).

We used Protein Variation Effect Analyzer (PROVEAN, v. 1.1.3; Choi and Chan 2015; Choi et al. 2012) to test whether non-synonymous mutations in codons with evidence of positive selection are functionally important. For PROVEAN analyses, we tested whether an amino acid substitution found in a sulfidic lineage occurs in a highly conserved region of the protein using the protein sequence of the most closely related non-sulfidic reference lineage as our query sequence and the amino acid substitution found in the sulfidic lineage as our variant, applying the default delta alignment score threshold of -2.5 for determination of potential functional effects.

RESULTS

Focal oxygen transport genes

Our analyses recovered a single gene encoding for myoglobin and eleven genes encoding for hemoglobin subunits (Table 2), including six α - and five β -type subunits. Two of the α - (HEMOA1 and HEMOAL2) and two of the β -type (HEMOB1 and HEMOBAL2) subunits of hemoglobin

exhibited very high levels of gene expression across all taxa, which were orders of magnitude higher than the expression of the other hemoglobin genes (Fig. 2; Table S2). These findings are consistent with previous studies on hemoglobin gene evolution documenting widespread duplication of genes that give rise to functionally distinct heme proteins and whose expression varies depending on developmental stage and environmental conditions (Hardison 2012; Opazo et al. 2013).

Expression variation

Analysis of evolutionary patterns of gene expression with EVE revealed no evidence for significant shifts in gene expression between all sulfidic vs. all non-sulfidic lineages (Table 2, Fig. 3). However, the expression of oxygen-transport genes shifted in a subset of sulfidic lineages. In three sulfidic lineages, there was evidence for a significant expression shift in a single gene: *G. eurystoma* shifted in the expression of HEMOAL, *L. sulphurophila* in HEMOB1L, and *P. mexicana* (Puyacatengo) in HEMOESA2 (Table 2, Fig. 3). In *P. sulphuraria*, we detected significant expression shifts in five oxygen-transport genes, including myoglobin, the two top-expressed heme genes (HEMOBAL and HEMOAA), as well as two additional heme genes (HEMOB1L2 and HEMOADL). In all documented cases, significant shifts were associated with increased levels of gene expression relative to other lineages (Fig. 4).

Evidence for positive selection

Analysis of molecular evolution based on branch models revealed evidence for different rates of non-synonymous to synonymous amino acid substitutions in sulfidic lineages ($\omega_{\text{non-sulfidic}} \neq \omega_{\text{sulfidic}}$) for two highly expressed hemoglobin genes (HEMOAA and HEMOBAL2; Table 3). In addition, branch-site models provided evidence for positive selection on specific codons in HEMOAA and HEMOBAL (Table 4).

HEMOAA exhibited evidence for positive selection at two codons (positions 73 and 143) with high support (probability >0.95 based on Bayes Empirical Bayes analysis; Fig. 5). Evidence for positive selection at position 73 was driven by a shift from glycine to alanine in the sulfidic lineage of *P. bimaculatus* as well as all sulfidic lineages in the *P. mexicana* species complex (*P. sulphuraria* and *P. mexicana* from the Puyacatengo and Tacotalpa river drainages). Analyses using PROVEAN suggested that these amino acid substitutions likely have neutral effects on protein function (PROVEAN scores between -1.913 and -1.640). The effect at position 143 was driven by sulfidic populations of *P. bimaculatus*, *G. eurystoma*, and *P. mexicana* (Puyacatengo) exhibiting a histidine instead of an arginine in all other lineages. PROVEAN results indicated that changes from arginine to histidine at position 143 have significant functional consequences (PROVEAN scores between -4.479 and -4.398). Note that the sites were fixed in *G. eurystoma*, *P. sulphuraria*, and *P. mexicana* (Tacotalpa); in *P. bimaculatus* and *P. mexicana* (Puyacatengo) the sites were at high frequency (allele frequencies between 0.8 and 0.9) in the sulfidic population.

HEMOBAL exhibited evidence for positive selection at one codon (position 26) with high support (Fig. 6). Codon position 26 is occupied by a proline in a majority of lineages and a glutamine in the others, and both variants occur in sulfidic and non-sulfidic populations. Evidence for positive selection was primarily driven by a shift toward a glutamine in two sulfidic *P. mexicana* lineages (Tacotalpa: allele fixed; Puyacatengo: allele frequency of 0.9). PROVEAN results indicated that a shift from a proline to a glutamine has significant consequences for the function of heme proteins (PROVEAN score: -5.745).

DISCUSSION

Organisms inhabiting environments with similar sources of selection provide the opportunity to investigate convergence in evolutionary adaptation, particularly when selection is strong, as in

environments characterized by extreme physiological conditions. H₂S binds to oxygen transport proteins and drastically impedes their function by reducing their oxygen binding affinity. Organisms inhabiting H₂S-rich environments are consequently predicted to compensate for the reduced capacity to transport oxygen, potentially leading to repeatable patterns of gene expression variation and molecular evolution. Our analyses of poeciliid fishes in sulfidic environments have revealed evidence for shifts in the expression of and positive selection on oxygen-transport genes. However, repeated changes in gene expression and evidence for positive selection was exceedingly rare and mostly restricted to a small number of sulfide spring lineages. Myoglobin and 7 of 11 hemoglobin genes exhibited evidence for significant up-regulation upon sulfide spring colonization, but all expression shifts were restricted to a single lineage (most of them in *P. sulphuraria*). Similarly, repeated positive selection was only found in HEMOAA (two codons) and HEMOBAL (one codon), and putatively adaptive amino acid substitutions at these codons were restricted to *P. bimaculatus*, *G. eurystoma*, and lineages within the *P. mexicana* species complex. Hence, neither changes in gene expression nor signatures of relaxed or positive selection were consistently present among all lineages in H₂S-rich environments, and four lineages inhabiting sulfide springs (*X. hellerii*, *G. holbrooki*, *G. sexradiata*, and *P. latipinna*) exhibited no evidence for significant shifts in the expression of oxygen-transport genes or for positive selection on them. Hence, while first principles may suggest that oxygen transport genes are predictable targets of selection in sulfidic environments and there is indeed lineage-specific evidence supporting this prediction, modulation and modification of oxygen-transport genes are not necessarily repeatable across replicated lineages.

Shared and unique patterns of expression and evolution of oxygen transport genes

As predicted, our results indicated that life in H₂S-rich environments can elicit up-regulation of and drive positive selection on oxygen transport genes. However, few of the documented patterns were

shared among evolutionarily replicated lineages, suggesting that they may contribute to the evolution of increased H₂S tolerance but are not required for it. Several non-mutually exclusive hypotheses can potentially explain the lack of consistent, convergent modulations and modifications of oxygen transport genes (see Kaeuffer et al. 2012 for an overview).

(1) There may be variation in the nature and strength of selection across independent lineages of sulfide spring fish, which could explain the lack of consistency in evolutionary responses. In our analyses, we have treated the presence and absence of H₂S as a dichotomous variable, because all sulfide springs investigated here exhibit H₂S concentrations above what is considered acutely toxic for animals (Riesch et al. 2014). Nonetheless, there is significant variation in H₂S concentrations even among geographically proximate sulfide spring complexes. For example, the Baños del Azufre in Mexico that harbor *P. sulphuraria* and *G. eurystoma* exhibit peak H₂S concentrations in excess of 1000 µM, while nearby Mogote del Puyacatengo with *G. sexradiata* only reaches low double digits (Culumber et al. 2016; Riesch et al. 2014; Tobler et al. 2011). Previous studies in the genera *Gambusia* and *Poecilia* have found that phenotypic divergence between adjacent sulfidic and non-sulfidic population increased with average H₂S concentrations (Riesch et al. 2016), highlighting the possibility that among-lineage variation in the modulation and modification of oxygen transport genes may be a function of unquantified environmental variation (also see Stuart et al. 2017). In addition, the presence of H₂S may also interact with hypoxia and other environmental factors to drive unique, lineage-specific adaptive solutions across the taxa we examined.

(2) Variation in evolutionary responses among different sulfide spring lineages may be a consequence of variation in the degree of divergence from non-sulfidic ancestors. Some sulfide spring populations (e.g., *P. sulphuraria*, *P. mexicana* from the Tacotalpa river drainage, *L. sulphurophila*, and *G. eurystoma*) exhibit low rates of gene flow with populations from adjacent non-sulfidic habitats (Plath et al. 2013; Riesch et al. 2016), and *P. sulphuraria* is considerably more divergent than other

sulfidic lineages, having diverged over 200,000 years ago (Pfenninger et al. 2014). Interestingly, *P. sulphuraria* was the only lineage exhibiting significant up-regulation of multiple oxygen-transport genes, and all of the older lineages were characterized by having at least one amino acid substitution at a codon with evidence for positive selection in HEMOAA and HEMOBAL. In other lineages (e.g., *P. latipinna*, *G. holbrooki*, *G. sexadiata*, and *X. hellerii*), population genetic differentiation is much lower (Riesch et al. 2016), and they exhibit no evidence for differential expression or positive selection on oxygen transport genes. These observations may point toward a possible role of gene flow in preventing adaptive divergence between adjacent populations in sulfidic and non-sulfidic habitats (Moore et al. 2007). It is noteworthy though that some young lineages exhibit amino acid substitutions in HEMOAA and HEMOBAL (*P. bimaculatus* and *P. mexicana* from the Puyacatengo drainage) and evidence for differential expression of hemoglobin genes (HEMOESA2 in *P. mexicana* from the Puyacatengo river drainage).

(3) Variation in evolutionary responses among sulfide spring fish lineages may be a consequence of functional redundancy, where multiple molecular modifications lead to equivalent performance of organisms in H₂S-rich habitats. A recent study on birds indicated that high-altitude species consistently evolved hemoglobins with higher oxygen-binding affinity, yet amino acid substitutions underlying variation in oxygen-binding substantially varied among lineages (Natarajan et al. 2016). If there are multiple amino acid substitutions with equivalent functional consequences, we would expect widespread positive selection on hemoglobin genes, but not necessarily congruence in amino acid substitutions among lineages exposed to the same sources of selection. This could explain why we see elevated rates of non-synonymous to synonymous amino acid substitutions for HEMOBAL2, but no consistent substitutions among lineages. Lineage-specific branch tests indicated evidence for relaxed selection in *G. eurystoma*, *P. sulphuraria*, and *P. mexicana* (Tacotalpa), but potentially relevant amino acid substitutions have occurred at different codon positions (see

Supplementary Fig. S1), leaving the possibility that non-convergent changes of the same gene might confer similar benefits in presence of H₂S. Patterns of consistent selection on oxygen transport genes – yet at different loci and through different amino acid substitutions – have also been documented in weakly electric fishes exposed to chronic hypoxia (Tian et al. 2017).

Open questions

Our study uncovered evidence for differential expression of and positive selection on oxygen transport genes of fishes inhabiting H₂S-rich environments, but several key questions remain unaddressed. First and foremost, the functional consequences of the observed expression and amino acid sequence variation in hemoglobin genes remain unknown. Up-regulation of oxygen-transport genes in *P. sulphuraria* and positive selection inferred from branch-site models predominantly occurred in the hemoglobin genes that are consistently highly expressed across all lineages (see Fig. 2), suggesting that putative adaptive modulations and modifications have occurred at loci that are important for the acquisition and distribution of oxygen in adult fish. Assessment of functional consequences of amino acid substitutions at codons with evidence for positive selection indicated that at the shift from a proline to a glutamine at codon position 26 of HEMOBAL as well as the shift from arginine to histidine at position 143 in HEMOAA should impact hemoglobin functionality. In the human homolog of HEMOAA (HBA1), the substitution of arginine to histidine causes an increase in oxygen binding affinity and a decrease in cooperativity and Bohr effect (Poyart and Bursaux 1980; Poyart et al. 1976). Thus, future research needs to explicitly test whether and how oxygen-binding affinity of hemoglobin varies among sulfidic and non-sulfidic lineages with evidence for non-synonymous amino acid substitutions. This is also required for assessing potential functional redundancies among different combinations of amino acid substitutions observed among sulfide spring lineages, especially in HEMOBAL2. Evaluating oxygen-binding affinities both in presence

and absence of H₂S would further provide insights into the relative roles of H₂S and hypoxia in shaping the evolution of hemoglobin genes. If modifications of hemoglobins are an adaptation to H₂S, oxygen-binding affinity should decrease at a greater rate with increasing H₂S concentrations in non-sulfidic than in sulfidic lineages. In contrast, if hypoxia is the primary source of selection shaping hemoglobin evolution, sulfidic lineages are expected to have hemoglobins with a higher oxygen-binding affinity than non-sulfidic lineages, but the rate of decline in binding-affinity in presence of H₂S should be equal.

It also remains unclear whether variation in oxygen transport gene expression documented here is a consequence of evolutionary change, or whether it merely reflects population-specific exposure histories to ambient environmental conditions. Gene expression in general – and hemoglobin expression specifically – is highly plastic and can be modulated over short periods of time upon exposure to environmental stressors (e.g., Dangre et al. 2010; Roesner et al. 2006; Roesner et al. 2008; Wawrowski et al. 2011). Laboratory experiments using common-garden-reared fish consequently need to test whether the colonization of H₂S-rich environments has resulted in evolved changes in regulatory mechanisms underlying oxygen transport gene expression. Laboratory studies of gene expression variation in two lineages of *P. mexicana* from sulfidic environments have shown that ancestral plasticity plays a minor role in explaining gene expression differences between sulfidic and non-sulfidic populations (Passow et al. 2017). Instead, differential expression in fish from different habitat types appears to be primarily a consequence of evolved changes in baseline expression (in absence of H₂S) and canalization (loss of expression responses to H₂S exposure in the derived, sulfide spring populations), suggesting that regulatory evolution plays an important role during adaptation to H₂S-rich environments (Passow et al. 2017). Whether these findings are mirrored in the expression variation of hemoglobin genes across the broad sampling of sulfide spring fish lineages investigated here remains to be tested.

Finally, our study also has no possibility to detect recent gene duplication events that potentially play a role in adaptation to H₂S-rich environments. Focal genes were identified through the inspection of the *Xiphophorus maculatus* genome (Schartl et al. 2013), and potential hemoglobin copy variation in specific lineages therefore go unnoticed. This could affect both our analyses of gene expression variation, because reads from recently duplicated genes would be mapped to the same locus, and our inferences about positive selection, because variation in the amino acid sequences from recently duplicated genes may either be lost during the inference of consensus sequences for each population or artificially combined into the same locus. Gene duplications have played an important role in the evolution of oxygen transport genes. Over broad evolutionary timescales, such duplications have given rise to gene complexes encoding for proteins with different functional properties that are expressed at different life stages or under different environmental conditions (Grispo et al. 2012; Hardison 2012). It remains unclear, however, whether gene duplications also played a role during microevolutionary changes that could mediate adaptation to particular environmental conditions, and analysis of whole-genome sequences will be required to detect recent gene duplication events in sulfide spring fishes.

Conclusions

The extreme environmental conditions in freshwater sulfide springs with their acutely toxic concentrations of H₂S and widespread hypoxia has led to pronounced patterns of convergent evolution in phenotypic traits. Sulfide spring fishes of the family Poeciliidae consistently exhibit elevated tolerance to H₂S as well as behavioral and morphological modifications that facilitate efficient oxygen acquisition (Riesch et al. 2016; Tobler and Hastings 2011; Tobler et al. 2011; Tobler et al. 2009). Our knowledge of potentially convergent modifications of the molecular mechanisms underlying tolerance to the extreme environmental conditions remains much more fragmentary,

although there appear to be convergent changes of toxicity targets (cytochrome c oxidase in the mitochondrial respiratory chain) and physiological pathways involved in enzymatic H₂S detoxification in at least a subset of sulfide spring fish lineages (Kelley et al. 2016; Pfenninger et al. 2014). This study identified another group of functionally important targets of selection in H₂S-rich environments. As predicted by theory, some oxygen transport genes exhibit evidence for shifts in expression and positive selection upon colonization of sulfide springs. Patterns of expression variation and amino acid substitutions were nonetheless highly idiosyncratic among different sulfide spring lineages, indicating that we may be able to predict the targets of selection during adaptation, but not the nature of modifications at those targets (Natarajan et al. 2016; Reid et al. 2016). Convergent evolution in oxygen-transport genes in response to H₂S is consequently far less prominent as other examples that investigated the evolutionary effects of physiochemical stressors with explicit biochemical consequences (Dobler et al. 2012; Reid et al. 2016; Tarvin et al. 2016). The current study also highlights that adaptation to the extreme environmental conditions in sulfide springs likely involves complex changes in the genome, including toxicity targets in the mitochondrial respiratory chain, toxicity targets involved in oxygen transport, and pathways associated with H₂S sequestration and detoxification (Tobler et al. 2016). A key challenge for future work in this and other systems will be to determine whether different targets of selection represent alternative or complementary mechanisms of adaptation, which will require integrative analyses from genomes to fitness across evolutionary replicated lineages exposed to the same sources of selection.

ACKNOWLEDGEMENTS

We thank the communities of Teapa and Tapijulapa, as well as local landowners and Villa Luz Nature Park for providing access to study sites. We are indebted to Centro de Investigación e Innovación para la Enseñanza y Aprendizaje (CIIEA) and Universidad Juárez Autónoma de

Tabasco (UJAT) for their hospitality and support over many years of research. We thank I. Schlupp and J. B. Johnson for providing samples of *Limia*, and Rori Rohlf for advice during the analysis of gene expression data. Permits were provided by the Mexican Federal Agencies SEMARNAT and CONAPESCA (DGOPA.09004.041111.3088, SGPA/DGVS/04315/11, PRMN/DGOPA-003/2014, PRMN/DGOPA-009/2015). This work was supported by grants from the National Science Foundation (IOS-1463720, IOS-1557795, and IOS-1557860), the Army Research Office (W911NF-15-1-0175), and the Defense University Research Instrumentation Program of the U.S. Office of Naval Research (W911NF-16-1-0225) to MT and JLK. In addition, NB was support by a Graduate Assistance in Areas of National Need (GAANN) Fellowship, RG was support by an NSF Graduate Research Fellowship (GRFP), and CNP was supported by the University of Minnesota Grand Challenges Postdoctoral Program.

REFERENCES

Amsterdam, A., Nissen, R.M., Sun, Z., Swindell, E.C., Farrington, S., and Hopkins, N. 2004. Identification of 315 genes essential for early zebrafish development. Proceedings of the National Academy of Sciences of the United States of America **101**(35): 12792-12797.

Auwer, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Rozen, D., and Thibault, J. 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics: 11.10. 11-11.10. 33.

Bagarinao, T. 1992. Sulfide as an environmental factor and toxicant: tolerance and adaptations in aquatic organisms. Aquatic Toxicology **24**: 21-62.

Beauchamp, R.O., Bus, J.S., Popp, C.S., Boreiko, C.J., and Andjelkovich, D.A. 1984. A critical review of the literature on hydrogen sulfide. CRC Critical Reviews in Toxicology **13**: 25-97.

- 540 Blount, Z.D., Borland, C.Z., and Lenski, R.E. 2008. Historical contingency and the evolution of a
541 key innovation in an experimental population of *Escherichia coli*. Proceedings of the National
542 Academy of Sciences USA **105**(23): 7899-7906.
- 543 Calvo, S.E., Clauser, K.R., and Mootha, V.K. 2015. MitoCarta2. 0: an updated inventory of
544 mammalian mitochondrial proteins. Nucleic acids research: gkv1003.
- 545 Choi, Y., and Chan, A.P. 2015. PROVEAN web server: a tool to predict the functional effect of
546 amino acid substitutions and indels. Bioinformatics **31**(16): 2745-2747.
- 547 Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. 2012. Predicting the functional effect
548 of amino acid substitutions and indels. PloS one **7**(10): e46688.
- 549 Cooper, C.E., and Brown, G.C. 2008. The inhibition of mitochondrial cytochrome oxidase by the
550 gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical
551 mechanism and physiological significance. Journal of Bioenergy and Biomembranes **40**: 533-
552 539.
- 553 Culumber, Z.W., Hopper, G.W., Barts, N., Passow, C.N., Morgan, S., Brown, A., Arias-Rodriguez,
554 L., and Tobler, M. 2016. Habitat use of two extremophile, highly endemic, and critically
555 endangered fish species (*Gambusia eurystoma* and *Poecilia sulphurophila*; Poeciliidae). Aquatic
556 Conservation-Marine and Freshwater Ecosystems **26**: 1155-1167.
- 557 Dangre, A.J., Manning, S., and Brouwer, M. 2010. Effects of cadmium on hypoxia-induced
558 expression of hemoglobin and erythropoietin in larval sheepshead minnow, *Cyprinodon*
559 *variegatus*. Aquatic Toxicology **99**: 168-175.
- 560 Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. 2012. jModelTest 2: more models, new
561 heuristics and parallel computing. Nature methods **9**(8): 772-772.

- 562 DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A.,
563 Del Angel, G., Rivas, M.A., and Hanna, M. 2011. A framework for variation discovery and
564 genotyping using next-generation DNA sequencing data. *Nature genetics* **43**(5): 491-498.
- 565 Dobler, S., Dalla, S., Wagschal, V., and Agrawal, A.A. 2012. Community-wide convergent evolution
566 in insect adaptation to toxic cardenolides by substitutions in the Na,K-ATPase. *Proceedings*
567 *of the National Academy of Sciences USA* **109**: 13040-13045.
- 568 Elmer, K.R., and Meyer, A. 2011. Adaptation in the age of genomics: insights from parallelism and
569 convergence. *Trends in Ecology & Evolution* **26**(6): 298-306.
- 570 Endler, J.A. 1986. *Natural selection in the wild*. Princeton University Press, Princeton, NJ.
- 571 Evans, D.H., and Claiborne, J.B. (eds). 2006. *The physiology of fishes*. 3rd ed. Taylor & Francis,
572 Boca Raton, FL.
- 573 Evans, S.V., Sishta, B.P., Mauk, A.G., and Brayer, G.D. 1994. Three-dimensional structure of
574 cyanomet-sulfmyoglobin C. *Proceedings of the National Academy of Sciences USA* **91**:
575 4723-4726.
- 576 Evans, T.G., Hammill, E., Kaukinen, K., Schulze, A.D., Patterson, D.A., English, K.K., Curtis,
577 J.M.R., and Miller, K.M. 2011. Transcriptomics of environmental acclimatization and
578 survival in wild adult Pacific sockeye salmon (*Oncorhynchus nerka*) during spawning migration.
579 *Molecular Ecology* **20**: 4472-4489.
- 580 Feldman, C.R., Brodie Jr., E.D., Brodie III, E.D., and Pfrender, M.E. 2012. Constraint shapes
581 convergence in tetrodotoxin-resistant sodium channels of snakes. *Proceedings of the*
582 *National Academy of Sciences USA* **109**: 4556-4561.
- 583 Flores, J.F., Fisher, C.R., Carney, S.L., Green, B.N., Freytag, J.K., Schaeffer, S.W., and Royer Jr.,
584 W.E. 2005. Sulfide binding is mediated by zinc ions discovered in the crystal structure of a

585 hydrothermal vent tubeworm hemoglobin. Proceedings of the National Academy of
586 Sciences USA **102**: 2713-2718.

587 Fraser, J., de Mello, L.V., Ward, D., Rees, H.H., Williams, D.R., Fang, Y.C., Brass, A., Gracey, A.Y.,
588 and Cossins, A.R. 2006. Hypoxia-inducible myoglobin expression in nonmuscle tissues.
589 Proceedings of the National Academy of Sciences of the United States of America **103**(8):
590 2977-2981.

591 Fukuto, J.M., Carrington, S.J., Tantillo, D.J., Harrison, J.G., Ignarro, L.J., Freeman, B.A., Chen, A.,
592 and Wink, D.A. 2012. Small molecule signaling agents: the integrated chemistry and
593 biochemistry of nitrogen oxides, oxides of carbon, dioxygen, hydrogen sulfide, and their
594 derived species. Chemical Research in Toxicology **25**: 769-793.

595 Gould, S.J., and Lewontin, R.C. 1979. The spandrels of San Marco and the panglossian paradigm: a
596 critique of the adaptationist programme. Proceedings of the Royal Society B **205**: 581-598.

597 Greenway, R., Arias-Rodriguez, L., Diaz, P., and Tobler, M. 2014. Patterns of macroinvertebrate and
598 fish diversity in freshwater sulphide springs. Diversity **6**: 597-632.

599 Grieshaber, M.K., and Völkel, S. 1998. Animal adaptations for tolerance and exploitation of
600 poisonous sulfide. Annual Review of Physiology **60**: 33-53.

601 Grispo, M.T., Natarajan, C., Projecto-Garcia, J., Moriyama, H., Weber, R.E., and Storz, J.F. 2012.
602 Gene duplication and the evolution of hemoglobin isoform differentiation in birds. Journal
603 of Biological Chemistry **287**: 37647-37658.

604 Hardison, R.C. 2012. Evolution of hemoglobin and its genes. Cold Spring Harbor Perspectives in
605 Medicine **2**: a011627.

606 Hildebrandt, T.M., and Grieshaber, M. 2008. Three enzymatic activities catalyze the oxidation of
607 sulfide to thiosulfate in mammalian and invertebrate mitochondria. FEBS Journal **275**: 3352-
608 3361.

- 609 Hourdez, S., Weber, R.E., Green, B.N., Kenney, J.M., and Fisher, C.R. 2002. Respiratory
610 adaptations in a deep-sea orbinid polychaete from Gulf of Mexico brine pool NR-1:
611 metabolic rates and hemoglobin structure/function relationships. *Journal of Experimental*
612 *Biology* **205**: 1669-1681.
- 613 Hrbek, T., Seckinger, J., and Meyer, A. 2007. A phylogenetic and biogeographic perspective on the
614 evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution* **43**: 986-998.
- 615 Huang, J., Zhang, L., Li, J., Shi, X., and Zhang, Z. 2013. Proposed function of alternative oxidase in
616 mitochondrial sulphide oxidation detoxification in the Echiuran worm, *Urechis unicinctus*.
617 *Journal of the Marine Biological Association* **93**: 2145-2154.
- 618 Johnson, E.A. 1970. The reversion to haemoglobin of sulphhaemoglobin and its coordination
619 derivatives. *Biochimica Et Biophysica Acta* **207**: 30-40.
- 620 Jourdan, J., Kaese, S.T., Lazar, V.M., Zimmer, C., Sommer-Trembo, C., Arias-Rodriguez, L., Klaus,
621 S., Riesch, R., and Plath, M. 2016. Shared and unique patterns of phenotypic diversification
622 along a stream gradient in two congeneric species. *Scientific Reports* **6**: 38971.
- 623 Kabil, O., and Banerjee, R. 2010. Redox biochemistry of hydrogen sulfide. *Journal of Biological*
624 *Chemistry* **285**: 21903-21907.
- 625 Kaeuffer, R., Peichel, C.L., Bolnick, D.I., and Hendry, A.P. 2012. Parallel and nonparallel aspects of
626 ecological, phenotypic, and genetic divergence across replicate population pairs of lake and
627 stream stickleback. *Evolution* **66**: 402-418.
- 628 Kelley, J.L., Arias-Rodriguez, L., Patocsil Martin, D., Yee, M.C., Bustamante, C., and Tobler, M.
629 2016. Mechanisms underlying adaptation to life in hydrogen sulfide rich environments.
630 *Molecular Biology and Evolution* **33**(6): 1419-1434.

- 631 Kelley, J.L., Passow, C., Plath, M., Arias-Rodriguez, L., Yee, M.C., and Tobler, M. 2012. Genomic
632 resources for a model in adaptation and speciation research: the characterization of the
633 *Poecilia mexicana* transcriptome. BMC Genomics **13**: 652.
- 634 Khan, A.A. 1990. Biochemical effects of hydrogen sulfide toxicity. In Proceedings of the
635 international conference on hydrogen sulfide toxicity, Banff, Alberta, Canada, June 18-21,
636 1989. Edited by M. Prior and S. Roth and F. Green and W. Hulbert and R. Reiffenstein.
637 Sulfide Research Network, Edmonton. pp. 79-89.
- 638 Kraus, D.W., and Wittenberg, J.B. 1990. Hemoglobins of the *Lucina pectinata*/bacteria symbiosis: 1.
639 Molecular properties, kinetics and equilibria of reactions with ligands. Journal of Biological
640 Chemistry **265**: 16043-16053.
- 641 Krueger, F. 2014. Trim Galore! version 0.3.7.
642 http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
- 643 Landry, L., and Bernatchez, L. 2010. Role of epibenthic resource opportunities in the parallel
644 evolution of lake whitefish species pairs (*Coregonus* sp.). Journal of Evolutionary Biology **23**:
645 2602-2613.
- 646 Langerhans, R.B., and DeWitt, T.J. 2004. Shared and unique features of evolutionary diversification.
647 American Naturalist **164**(3): 335-349.
- 648 Larkin, M.A., Blackshields, G., Brown, N., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin,
649 F., Wallace, I.M., Wilm, A., and Lopez, R. 2007. Clustal W and Clustal X version 2.0.
650 Bioinformatics **23**(21): 2947-2948.
- 651 Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
652 transform. Bioinformatics **25**: 1754-1760.

- 653 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and
654 Durbin, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*
655 **25**(16): 2078-2079.
- 656 Liu, X., Zhang, Z., Ma, X., Li, X., Zhou, D., Gao, B., and Bai, Y. 2015. Sulfide exposure results in
657 enhanced *sqr* transcription through upregulating the expression and activation of HSF1 in
658 echiuran worm *Urechis unicinctus*. *Aquatic Toxicology* **170**: 229-239.
- 659 Losos, J.B. 2009. *Lizards in an evolutionary tree: ecology and adaptive radiation of Anoles*.
660 University of California Press, Berkeley, CA.
- 661 Losos, J.B. 2011. Convergence, adaptation, and constraint. *Evolution* **65**: 1827-1840.
- 662 Löytynoja, A., and Goldman, N. 2005. An algorithm for progressive multiple alignment of
663 sequences with insertions. *Proceedings of the National academy of sciences of the United*
664 *States of America* **102**(30): 10557-10562.
- 665 Ma, Y.-B., Zhang, Z.-F., Shao, M.-Y., Kang, K.-H., Shi, X.-L., Dong, Y.-P., and Li, J.-J. 2012.
666 Response of sulfide-quinone oxireductase to sulfide exposure in the echiuran worm *Urechis*
667 *unicinctus*. *Marine Biotechnology* **14**: 245-251.
- 668 Martin, N.M., and Maricle, B.R. 2015. Species-specific enzymatic tolerance of sulfide toxicity in plant
669 roots. *Plant Physiology and Biochemistry* **88**: 36-41.
- 670 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K.,
671 Altshuler, D., Gabriel, S., and Daly, M. 2010. The Genome Analysis Toolkit: a MapReduce
672 framework for analyzing next-generation DNA sequencing data. *Genome research* **20**(9):
673 1297-1303.
- 674 Moore, J., Gow, J., Taylor, E.B., and Hendry, A.P. 2007. Quantifying the constraining influence of
675 gene flow on adaptive divergence in the lake-stream threespine stickleback system.
676 *Evolution* **61**: 2015-2026.

- 677 Natarajan, C., Hoffmann, F.G., Weber, R.E., Fago, A., Witt, C.C., and Storz, J.F. 2016. Predictable
678 convergence in hemoglobin function has unpredictable molecular underpinnings. *Science*
679 **354**: 336-339.
- 680 Olson, K.R. 2011. The therapeutic potential of hydrogen sulfide: separating hype from hope.
681 *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **301**:
682 R297-312.
- 683 Opazo, J.C., Butts, G.T., Nery, M.F., Storz, J.F., and Hoffmann, F.G. 2013. Whole-genome
684 duplication and the functional diversification of teleost fish hemoglobins. *Molecular Biology*
685 *and Evolution* **30**: 140-153.
- 686 Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.-E., Walford, G.A., Sugiana,
687 C., Boneh, A., and Chen, W.K. 2008. A mitochondrial protein compendium elucidates
688 complex I disease biology. *Cell* **134**(1): 112-123.
- 689 Palacios, M., Voelker, G., Arias-Rodriguez, L., Mateos, M., and Tobler, M. 2016. Phylogenetic
690 analyses of the subgenus *Mollienesia* (*Poecilia*, Poeciliidae, Teleostei) reveal taxonomic
691 inconsistencies, cryptic biodiversity, and spatio-temporal aspects of diversification in Middle
692 America. *Molecular Phylogenetics and Evolution* **103**: 230-244.
- 693 Passow, C., Arias-Rodriguez, L., and Tobler, M. in press. Convergent evolution of reduced energy
694 demands in extremophile fish. *PLOS ONE*.
- 695 Passow, C.N., Henpita, C., Shaw, J.H., Quakenbush, C., Warren, W.C., Scharl, M., Arias-Rodriguez,
696 L., Kelley, J.L., and Tobler, M. 2017. The roles of plasticity and evolutionary change in
697 shaping gene expression variation in natural populations of extremophile fish. *Molecular*
698 *Ecology* **in press**.
- 699 Perutz, M.F. 1983. Species adaptation in a protein module. *Molecular Biology and Evolution* **1**: 1-28.

- 700 Pfenninger, M., Lerp, H., Tobler, M., Passow, C., Kelley, J.L., Funke, E., Greshake, B., Erkoc, U.K.,
701 Berberich, T., and Plath, M. 2014. Parallel evolution of *cox* genes in H₂S-tolerant fish as key
702 adaptation to a toxic environment. *Nature Communications* **5**: 3873.
- 703 Pfenninger, M., Patel, S., Arias-Rodriguez, L., Feldmeyer, B., Riesch, R., and Plath, M. 2015. Unique
704 evolutionary trajectories in repeated adaptation to hydrogen sulphide-toxic habitats of a
705 neotropical fish (*Poecilia mexicana*). *Molecular Ecology* **24**: 5446-5459.
- 706 Pietri, R., Roman-Morales, E., and Lopez-Garriga, J. 2011. Hydrogen sulfide and heme proteins:
707 knowledge and mysteries. *Antioxidants & Redox Signaling* **15**: 393-404.
- 708 Plath, M., Pfenninger, M., Lerp, H., Riesch, R., Eschenbrenner, C., Slattery, P., Bierbach, D.,
709 Herrmann, N., Schulte, M., Arias-Rodriguez, L., Indy, J.R., Passow, C., and Tobler, M. 2013.
710 Genetic differentiation and selection against migrants in evolutionarily replicated extreme
711 environments. *Evolution* **67**: 2647-2661.
- 712 Pollux, B.J.A., Meredith, R.W., Springer, M.S., Garland, T., and Reznick, D.N. 2014. The evolution
713 of the placenta drives a shift in sexual selection in livebearing fish. *Nature* **513**: 233-236.
- 714 Poyart, C., and Bursaux, E. 1980. Structural and functional studies of hemoglobin suresnes (arg 141
715 alpha 2 replaced by His beta 2). Consequences of disrupting an oxygen-linked anion-binding
716 site. *Journal of Biological Chemistry* **255**: 9465-9473.
- 717 Poyart, C., Krishnamoorthy, R., Bursaux, E., Gacon, G., and Labie, D. 1976. Structural and
718 functional studies of hemoglobin suresnes or α_2 141(HC3) Arg $\rightarrow$ His β_2 , a new high oxygen
719 affinity mutant. *FEBS Letters* **69**: 103-107.
- 720 Reid, N.M., Prostou, D.A., Clark, B.W., Warren, W.C., Colbourne, J.K., Shaw, J.R., Karchner, S.I.,
721 Hahn, M.E., Nacci, D., Oleksiak, M.F., Crawford, D.L., and Whitehead, A. 2016. The
722 genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish.
723 *Science* **354**: 1305-1308.

- 724 Richards, J.G. 2009. Metabolic and molecular responses of fish to hypoxia. *In* Fish physiology:
725 hypoxia. *Edited by* J.G. Richards and A.P. Farrell and C.J. Brauner. Academic Press, New
726 York, NY. pp. 443-485.
- 727 Riesch, R., Plath, M., Garcia de Leon, F.J., and Schlupp, I. 2010. Convergent life-history shifts: toxic
728 environments result in big babies in two clades of poeciliids. *Naturwissenschaften* **97**(2):
729 133-141.
- 730 Riesch, R., Plath, M., Schlupp, I., Tobler, M., and Langerhans, R.B. 2014. Colonization of toxic
731 springs drives predictable life-history shift in livebearing fishes (Poeciliidae). *Ecology Letters*
732 **17**: 65-71.
- 733 Riesch, R., Tobler, M., Lerp, H., Jourdan, J., Dumas, L.T., Nosil, P., Langerhans, R.B., and Plath,
734 M. 2016. Extremophile Poeciliidae: multivariate insights into the complexity of speciation
735 along replicated ecological gradients. *BMC Evolutionary Biology* **16**: 136.
- 736 Riesch, R., Tobler, M., and Plath, M. 2015. Hydrogen sulfide-toxic habitats. *In* Extremophile Fishes:
737 Ecology, Evolution, and Physiology of Teleosts in Extreme Environments. *Edited by* R.
738 Riesch and M. Tobler and M. Plath. Springer, Heidelberg, Germany. pp. 137-159.
- 739 Roesner, A., Hankeln, T., and Burmester, T. 2006. Hypoxia induces a complex response of globin
740 expression in zebrafish (*Danio rerio*). *Journal of Experimental Biology* **209**(11): 2129-2137.
- 741 Roesner, A., Mitz, S.A., Hankeln, T., and Burmester, T. 2008. Globins and hypoxia adaptation in the
742 goldfish, *Carassius auratus*. *FEBS Journal* **275**: 3633-3643.
- 743 Rohlf, R., Harrigan, P., and Nielsen, R. 2013. Modeling gene expression evolution with an extended
744 Ornstein-Uhlenbeck process accounting for within-species variation. *Molecular Biology and*
745 *Evolution* **31**: 201-211.
- 746 Rohlf, R.V., and Nielsen, R. 2015. Phylogenetic ANOVA: the expression variance and evolution
747 (EVE) model for quantitative trait evolution. *Systematic Biology* **64**: 695-708.

- 748 Rosenblum, E.B., Parent, C.E., and Brandt, E.E. 2014. The molecular basis of convergent evolution.
749 Annual Review of Ecology Evolution and Systematics **45**: 203-226.
- 750 Schartl, M., Walter, R.B., Shen, Y., Garcia, T., Catchen, J., Amores, A., Braasch, I., Chalopin, D.,
751 Volff, J.-N., Lesch, K.-P., Bisazza, A., Minx, P., Hillier, L., Wilson, R.K., Fuerstenberg, S.,
752 Boore, J.L., Searle, S., Postlethwait, J.H., and Warren, W.C. 2013. The genome of the
753 platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several
754 complex traits. Nature Genetics **45**: 567-572.
- 755 Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.
- 756 Shen, Y.-Y., Liang, L., Zhu, Z.-H., Zhou, W.-P., Irwin, D.M., and Zhang, Y.-P. 2010. Adaptive
757 evolution of energy metabolism genes and the origin of flight in bats. Proceedings of the
758 National Academy of Sciences **107**(19): 8666-8671.
- 759 Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
760 phylogenies. Bioinformatics **30**(9): 1312-1313.
- 761 Stern, D.L. 2013. The genetic causes of convergent evolution. Nature Reviews Genetics **14**: 751-764.
- 762 Storz, J.F., Scott, G.R., and Cheviron, Z.A. 2010. Phenotypic plasticity and genetic adaptation to
763 high-altitude hypoxia in vertebrates. Journal of Experimental Biology **213**: 4125-4136.
- 764 Stuart, Y.E., Veen, T., Weber, J.N., Hanson, D., Ravinet, M., Lohman, B.K., Thompson, C.J.,
765 Tasneem, T., Doggett, A., Izen, R., Ahmed, N., Barrett, R.D.H., Hendry, A.P., Peichel, C.L.,
766 and Bolnick, D.I. 2017. Contrasting effects of environment and genetics generate a
767 continuum of parallel evolution. Nature Ecology & Evolution **1**: 0158.
- 768 Tarvin, R.D., Santos, J.C., O'Connell, L.A., Zakon, H.H., and Cannatella, D.C. 2016. Convergent
769 substitutions in a sodium channel suggests multiple origins of toxin resistance in poison
770 frogs. Molecular Biology and Evolution **33**: 1068-1081.

- 771 Tian, R., Losilla, M., Lu, Y., Yang, G., and Zakon, H. 2017. Molecular evolution of globin genes in
772 Gymnotiform electric fishes; relation to hypoxia tolerance. *Bmc Evolutionary Biology* **17**: 51.
- 773 Tobler, M., and Culumber, Z.W. 2016. Swimming in polluted waters. *Science* **354**: 1232-1233.
- 774 Tobler, M., and Hastings, L. 2011. Convergent patterns of body shape differentiation in four
775 different clades of poeciliid fishes inhabiting sulfide springs. *Evolutionary Biology* **38**: 412-
776 421.
- 777 Tobler, M., Henpita, C., Bassett, B., Kelley, J.L., and Shaw, J. 2014. H₂S exposure elicits differential
778 expression of candidate genes in fish adapted to sulfidic and non-sulfidic environments.
779 *Comparative Biochemistry and Physiology A: Molecular & Integrative Physiology* **175**: 7-14.
- 780 Tobler, M., Palacios, M., Chapman, L.J., Mitrofanov, I., Bierbach, D., Plath, M., Arias-Rodriguez, L.,
781 Garcia de Leon, F.J., and Mateos, M. 2011. Evolution in extreme environments: replicated
782 phenotypic differentiation in livebearing fish inhabiting sulfidic springs. *Evolution* **65**: 2213-
783 2228.
- 784 Tobler, M., Passow, C.N., Greenway, R., Kelley, J.L., and Shaw, J.H. 2016. The evolutionary ecology
785 of animals inhabiting hydrogen sulfide-rich environments. *Annual Review of Ecology,*
786 *Evolution and Systematics* **47**: 239-262.
- 787 Tobler, M., and Plath, M. 2011. Living in extreme habitats. *In Ecology and evolution of poeciliid*
788 *fishes. Edited by J. Evans and A. Pilastro and I. Schlupp. University of Chicago Press,*
789 *Chicago. pp. 120-127.*
- 790 Tobler, M., Riesch, R., Tobler, C.M., and Plath, M. 2009. Compensatory behaviour in response to
791 sulfide-induced hypoxia affects time budgets, feeding efficiency, and predation risk.
792 *Evolutionary Ecology Research* **11**: 935-948.
- 793 Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L.,
794 Wold, B.J., and Pachter, L. 2010. Transcript assembly and quantification by RNA-Seq reveals

- unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**(5): 511-515.
- van der Meer, D.L., Van Den Thillart, G., Witte, F., de Bakker, M.A., Besser, J., Richardson, M.K., Spaink, H.P., Leito, J.T., and Bagowski, C.P. 2005. Gene expression profiling of the long-term adaptive response to hypoxia in the gills of adult zebrafish. *American Journal of Physiology - Integrative and Comparative Physiology* **289**: 1512-1519.
- Vermeij, G.J. 2006. Historical contingency and the purported uniqueness of evolutionary innovations. *Proceedings of the National Academy of Sciences USA* **103**: 1804-1809.
- Vivitsky, V., Yadav, P.K., Kurthen, A., and Banerjee, R. 2015. Sulfide oxidation by a noncanonical pathway in red blood cells generates thiosulfate and polysulfides. *Journal of Biological Chemistry* **290**: 8310-8320.
- Waits, E.R., Martinson, J., Rinner, B., Morris, S., Proestou, D., Champlin, D., and Nacci, D. 2016. Genetic linkage map and comparative genome analysis for the Atlantic killifish (*Fundulus heteroclitus*). *Open Journal of Genetics* **6**: 28-28.
- Warrington, J.A., Nair, A., Mahadevappa, M., and Tsyganskaya, M. 2000. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiological genomics* **2**(3): 143-147.
- Wawrowski, A., Gerlach, F., Hankeln, T., and Burmester, T. 2011. Changes of globin expression in the Japanese medaka (*Oryzias latipes*) in response to acute and chronic hypoxia. *Journal of Comparative Physiology B* **181**: 199-208.
- Weaver, P.F., Cruz, A., Johnson, S.C., Dupin, J., and Weaver, K.F. 2016. Colonizing the Caribbean: biogeography and evolution of livebearing fishes of the genus *Limia* (Poeciliidae). *Journal of Biogeography* **43**: 1808-1819.

- 818 Weber, R.E. 2007. High-altitude adaptations in vertebrate hemoglobins. *Respiration Physiology and*
819 *Neurobiology* **158**: 132-142.
- 820 Weber, R.E., and Fago, A. 2004. Functional adaptation and its molecular basis in vertebrate
821 hemoglobins, neuroglobins and cytoglobins. *Respiratory Physiology & Neurobiology* **144**:
822 141-159.
- 823 Wilkens, H., and Strecker, U. 2003. Convergent evolution of the cavefish *Astyanax* (Characidae,
824 Teleostei): genetic evidence from reduced eye-size and pigmentation. *Biological Journal of*
825 *the Linnean Society* **80**(4): 545-554.
- 826 Winemiller, K.O., Kelso-Winemiller, L.C., and Brenkert, A.L. 1995. Ecomorphological
827 diversification and convergence in fluvial cichlid fishes. *Environmental Biology of Fishes* **44**:
828 235-261.
- 829 Yang, Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and*
830 *Evolution* **24**(8): 1586-1591.
- 831 Yang, Z., and Nielsen, R. 2002. Codon-substitution models for detecting molecular adaptation at
832 individual sites along specific lineages. *Molecular Biology and Evolution* **19**: 908-917.
- 833 Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.M. 2000. Codon-substitution models for
834 heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431-449.
- 835 Zal, F., Leize, E., Lallier, F.H., Toulmond, A., Van Dorsselaer, A., and Childress, J.J. 1998. S-
836 sulfohemoglobin and disulfide exchange: the mechanisms of sulfide binding by *Riftia*
837 *pachyptila* hemoglobins. *Proceedings of the National Academy of Sciences USA* **95**: 8997-
838 9002.
- 839 Zhang, F., and Broughton, R.E. 2013. Mitochondrial–Nuclear Interactions: Compensatory
840 Evolution or Variable Functional Constraint among Vertebrate Oxidative Phosphorylation
841 Genes? *Genome biology and evolution* **5**(10): 1781-1791.

842 Zhang, J., Nielsen, R., and Yang, Z. 2005. Evaluation of an improved branch-site likelihood method
843 for detecting positive selection at the molecular level. *Molecular Biology and Evolution* **22**:
844 2472-2479.
845
846

Draft

FIGURE CAPTIONS

Figure 1: Phylogenetic relationships and collection localities of focal lineages of poeciliid fishes that have colonized H₂S-rich environments, along with relevant reference lineages from non-sulfidic habitats. A. Phylogeny of lineages included in analyses of expression variation and evolution of oxygen transport genes. Lineages from H₂S-rich environments (S) are underlined, lineages from non-sulfidic lineages are designated as NS. Asterisks indicate bootstrap support $\geq 99\%$. B. Location of collection sites in the US, Mexico, and the Dominican Republic. C. Detailed view of collection localities in Southern Mexico, which harbors the highest diversity of sulfide spring fishes. Roads (black lines) and major towns (shaded areas) are included for orientation. Note that yellow symbols correspond to sulfidic sites, blue symbols to non-sulfidic sites. Numbers correspond to lineages as listed in Table 1.

Figure 2: Overall expression levels ($\log_{10}$ -transformed) of focal genes (one myoglobin gene and eleven hemoglobin genes) across all lineages investigated. Boxes cover the first through third quartile of the data; vertical black lines indicate the median. Open diamonds indicate the mean for each group, and closed dots indicate outliers.

Figure 3: Results of EVE analyses testing for gene expression differences between sulfidic and non-sulfidic lineages overall, as well as in specific lineages that have colonized sulfidic environments. Depicted are empirical LRT_0 distributions for each comparison, including LRT_0 values for background genes (gray) and oxygen-transport genes (red). The vertical red line represents the top 5% cut-off value; i.e., genes with LRT_0 values above this line are considered having significant evidence for an expression shift in the focal lineage.

Figure 4: Lineage-specific expression variation for oxygen-transport genes with significant evidence for expression shifts associated with the colonization of H₂S-rich environments as established by EVE. Non-sulfidic lineages are in blue, sulfidic lineages in yellow, and lineages with significant expression shifts are highlighted with asterisks (also see Table 2). Boxes cover the first through third quartile of the log₁₀-transformed expression levels of each lineage; vertical black lines indicate the median. Open diamonds indicate the mean for each group, and closed dots indicate outliers. The phylogenetic trees depict the evolutionary relationships across lineages; numbers correspond to lineage IDs in Table 1.

Figure 5: Amino acid sequence alignment of HEMOAA, which had significant evidence for selection on specific codons based on a branch-site test in codeml (see Table 4). Highlighted in red are amino acid positions that exhibited evidence for positive selection based on a Bayes Empirical Bayes (BEB) analysis (position 73, probability of 0.996; position 143, probability of 0.989). Amino acid substitutions in sulfidic lineages are highlighted in bold; substitutions with significant functional effects based on PROVEAN analyses are underlined.

Figure 6: Amino acid sequence alignment of HEMOBAL, which had significant evidence for selection on one codon based on a branch-site test in codeml (see Table 4). Highlighted in red is the amino acid position that exhibited evidence for positive selection based on a Bayes Empirical Bayes (BEB) analysis (position 26, probability of 0.954). Amino acid substitutions in sulfidic lineages are highlighted in bold; substitutions with significant functional effects based on PROVEAN analyses are underlined.

TABLES

Table 1. List of lineages from sulfidic and non-sulfidic habitats included in this study. The table provides descriptions of the collection localities, information about the presence or absence of H₂S, latitude and longitude, as well as the number of individuals (*N*) collected from each site.

ID	Species	Locality	H ₂ S	Lat/Long	<i>N</i>
1	<i>Xiphophorus hellerii</i>	Rio El Azufre, west branch, Rio Pichucalco drainage, Chiapas, Mexico	–	17.556/-93.008	6
2	<i>Xiphophorus hellerii</i>	La Gloria springs, Rio Pichucalco drainage, Chiapas, Mexico	+	17.532/-93.015	6
3	<i>Pseudoxiphophorus bimaculatus</i>	Arroyo Pujil, Rio Ixtapangajoya drainage, Chiapas, Mexico	–	17.476/-92.986	6
4	<i>Pseudoxiphophorus bimaculatus</i>	La Gloria springs, Rio Pichucalco drainage, Chiapas, Mexico	+	17.532/-93.015	6
5	<i>Gambusia holbrooki</i>	Mariner's Cove, Lake Monroe, St. Johns River drainage, Florida, USA	–	28.857/-81.239	6
6	<i>Gambusia holbrooki</i>	Green Springs, St. Johns River drainage, Florida, USA	+	28.863/-81.248	6
7	<i>Gambusia eurystoma</i>	Baños del Azufre, Rio Pichucalco drainage, Tabasco, Mexico	+	17.552/-92.999	6
8	<i>Gambusia sexradiata</i>	Laguna Sitio Grande, Rio Ixtapangajoya drainage, Tabasco, Mexico	–	17.677/-92.997	6
9	<i>Gambusia sexradiata</i>	Mogote del Puyacatengo, Rio Puyacatengo drainage, Tabasco, Mexico	+	17.582/-92.900	6
10	<i>Limia perugiae</i>	Stream in Cabral, Rio Yaque del Sur drainage, Barahona, Dominican Republic	–	18.246/-71.223	6
11	<i>Limia sulphurophila</i>	Balnearios La Zurza, Lago Enriquillo basin, Independencia, Dominican Republic	+	18.398/-71.570	6
12	<i>Poecilia latipinna</i>	Mariner's Cove, Lake Monroe, St. Johns River drainage, Florida, USA	–	28.857/-81.239	6
13	<i>Poecilia latipinna</i>	Green Springs, St. Johns River drainage, Florida, USA	+	28.863/-81.248	6
14	<i>Poecilia limantouri</i>	Rio Garces, Rio Panuco drainage, Hidalgo, Mexico	–	20.940/-98.282	6
15	<i>Poecilia sulphuraria</i>	Baños del Azufre, Rio Pichucalco drainage, Tabasco, Mexico	+	17.552/-92.999	6
16	<i>Poecilia mexicana</i>	El Azufre II, Rio Tacotalpa drainage, Tabasco Mexico	+	17.438/-92.775	6
17	<i>Poecilia mexicana</i>	Rio Puyacatengo, Rio Puyacatengo drainage, Tabasco Mexico	–	17.510/-92.914	6
18	<i>Poecilia mexicana</i>	La Lluvia springs, Rio Puyacatengo drainage, Tabasco, Mexico	+	17.464/-92.895	5
19	<i>Poecilia mexicana</i>	Arroyo Rosita, Rio Pichucalco drainage, Chiapas, Mexico	–	17.485/-93.104	6
20	<i>Poecilia mexicana</i>	Arroyo Bonita, Rio Tacotalpa drainage, Tabasco, Mexico	–	17.427/-92.752	6

900 Table 2. List of focal oxygen transport genes investigated in this study, including eleven hemoglobin genes and one myoglobin gene. The
901 table includes the gene ID from the *X. maculatus* reference genome annotation file (XM ID), the accession ID, percent sequence similarity,
902 and *E*-value from the top BLAST hit in Swiss-Prot. The table also contains the results of branch tests that evaluated shifts in gene
903 expression associated with the colonization of H₂S-rich habitats as implemented in EVE. We list the likelihood ratio test statistic (LRT_{θ})
904 for all sulfidic lineages combined (overall) as well as for each sulfidic lineage individually (labeling as in Figure 1). Genes with evidence for
905 divergence in expression (falling within the top 5% of LRT_{θ} values; Fig. 3) are highlighted in bold.

Gene	XM ID	Accession	% similarity	<i>E</i> -value	Overall	Xhel (2)	Pbim (4)	Ghol (6)	Geur (7)	Gsex (9)	Lsul (11)	Plat (13)	Psul (15)	Pmex (Tac) (16)	Pmex (Puy) (18)
HEMOA1L	005809044.2	P02008	49	1.40E-44	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HEMOAA	005816937.1	P69905	55.2	1.80E-48	1.532	0.000	0.000	0.000	0.000	0.000	0.000	0.000	20.987	0.000	0.000
HEMOADL	005815125.2	P02008	52.9	1.30E-47	3.035	0.172	0.186	0.040	0.527	0.042	0.000	0.323	36.268	0.000	0.083
HEMOAL	014474961.1	P02008	58.9	2.00E-36	1.561	0.000	0.031	0.043	14.501	0.030	0.112	0.000	0.728	0.000	0.000
HEMOAL2	005809047.2	P69905	53.8	7.90E-47	3.064	0.000	0.000	0.000	0.000	0.000	0.000	0.000	14.289	0.000	0.000
HEMOB1L	005815126.1	P69891	50.7	8.40E-49	0.023	0.330	0.640	0.642	3.119	0.072	6.696	0.254	0.023	0.158	0.055
HEMOB1L2	005809045.2	P02100	43.8	3.80E-41	1.308	0.242	0.031	0.085	0.326	0.013	0.299	0.011	27.709	0.235	0.121
HEMOBAL	005816936.2	P02100	52.7	9.60E-51	0.000	0.480	0.245	2.954	2.107	0.664	0.250	1.043	3.998	2.892	0.130
HEMOBAL2	005809046.1	P02100	50.7	7.70E-50	2.550	0.000	0.000	0.000	0.000	0.000	0.000	0.000	32.787	0.000	0.000
HEMOBL	005809043.1	P02100	46.3	1.10E-43	4.297	0.059	1.111	0.945	0.322	0.010	0.014	0.634	0.945	1.111	0.063
HEMOESA2	005809048.1	P02008	52.8	1.30E-45	1.155	0.188	0.410	0.156	0.163	0.240	0.045	0.030	2.280	0.459	2.559
MYO	005797508.1	P02144	43.2	1.00E-32	1.832	0.015	0.876	0.277	0.675	0.405	0.267	0.887	23.503	1.549	0.316

906

Table 3. Results of branch tests evaluating differences in non-synonymous to synonymous amino acid substitution rates (ω) between non-sulfidic (background) and sulfidic (foreground) lineages. Branch test models ($\omega_{\text{non-sulfidic}} \neq \omega_{\text{sulfidic}}$) were evaluated against a null model ($\omega_{\text{non-sulfidic}} = \omega_{\text{sulfidic}}$) using a likelihood ratio test, and we present the log-likelihoods ($\ln L$) for each model, the test statistic ($2\Delta\ln L$), and the corresponding P -value. In addition, we provide ω of the null model or separate estimates of ω for non-sulfidic and sulfidic lineages, depending on which model was supported. Genes with evidence for $\omega_{\text{non-sulfidic}} \neq \omega_{\text{sulfidic}}$ are highlighted in bold.

Gene	$\ln L_{\text{null}}$	$\ln L_{\text{branch}}$	$2\Delta\ln L$	P	ω_{null}	$\omega_{\text{non-sulfidic}}$	ω_{sulfidic}
HEMOA1L	-258.0	-256.4	3.195	0.074	0.228		
HEMOAA	-1154.1	-1151.0	6.023	0.014		0.252	0.818
HEMOADL	-676.5	-676.5	0.000	0.998	0.036		
HEMOAL2	-994.8	-993.7	2.130	0.144	0.599		
HEMOAL	-15.3	-15.3	0.000	1.000	1.000		
HEMOB1L2	-853.5	-853.4	0.253	0.615	0.183		
HEMOB1L	-425.9	-425.7	0.255	0.613	0.109		
HEMOBAL2	-819.4	-817.2	4.415	0.036		0.153	0.565
HEMOBAL	-941.3	-941.4	0.209	0.648	0.327		
HEMOESA2	-1.4	-1.4	0.000	1.000	1.000		
MYO	-813.2	-812.2	1.825	0.177	0.152		

Table 4. Results of branch-site tests evaluating positive selection on specific codon between sulfidic and non-sulfidic lineages. Branch-site models were evaluated against a null model using a likelihood ratio test, and we present the log-likelihoods ($\ln L$) for each model, the test statistic ($2\Delta\ln L$), and the corresponding P -value. Genes with evidence for significant branch-site models are highlighted in bold.

Gene	$\ln L_{null}$	$\ln L_{branch-site}$	$2\Delta\ln L$	P
HEMOA1L	-868.3	-868.2	0.248	0.619
HEMOAA	-1123.6	-1115.8	15.424	<0.001
HEMOADL	-673.3	-673.3	0.000	0.998
HEMOAL2	-972.8	-970.9	3.745	0.053
HEMOAL	-620.5	-620.5	0.000	1.000
HEMOB1L2	-848.0	-848.0	0.000	1.000
HEMOB1L	-669.3	-669.3	0.000	1.000
HEMOBAL2	-806.0	-805.5	0.823	0.364
HEMOBAL	-919.6	-913.3	12.692	<0.001
HEMOBL	-1046.6	-1046.4	0.511	0.475
HEMOESA2	-1021.7	-1021.5	0.425	0.514
MYO	-813.2	-813.2	0.000	1.000

A

B

C

MYO

HEMOBAL2

HEMOAAL

HEMOAL2

HEMOBAL

HEMOB1L2

HEMOADL

HEMOAL

HEMOA1L

HEMOB1L

HEMOBL

HEMOESA2

<i>X. hellerii</i> NS	MSLSGKDKTVVKSFWDKVAPKAAEMGGEALGRMLTAYPQTKTYFSHWSDLSPESAQVKKHGATIMSAIGDAVAKIDDMPSAMSKLSELHAFKLRVDPANFRILAHNIVLVLALYFPADFTPEVHVSIDKFLQNLAMALSEKYR
<i>X. hellerii</i> S	MSLSGKDKTVVKSFWDKVAPKAAEMGGEALGRMLTAYPQTKTYFSHWSDLSPESAQVKKHGATIMSAIGDAVAKIDDMPSAMSKLSELHAFKLRVDPANFRILAHNIVLVLALYFPADFTPEVHVSIDKFLQNLAMALSEKYR
<i>P. bimaculatus</i> NS	MSLSGKDKTVVKAFWDKMAPKAAELGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMLAIGDAIGKIDDLPGALSKLSELHAFKLRVDPANFRILAHNIIILVLALYFPADFTPEVHVSIDKFLQNLALALSEKYR
<i>P. bimaculatus</i> S	MSLSGKDKTVVKAFWDKMAPKAAELGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMLAIGDAIAKIDDLPGALSKLSELHAFKLRVDPANFRILAHNIIILVLALYFPADFTAENVHVSVDKFLQNLALALSEKYH
<i>G. holbrooki</i> NS	MSLSGKDKTVVKAFWDKIASKAPELGGEALGRMLTSYPQTKTYFAHWSDLSPESAQVKKHGATIMLAIGDAIGKIDDLTGALSKLSELHAFKLRVDPANFRILAHNIIILVLAMYFPADFTPEVHVSVDKFLQNLALGLSEKYR
<i>G. holbrooki</i> S	MSLSGKDKTVVKAFWDKIASKAPELGGEALGRMLTSYPQTKTYFAHWSDLSPESAQVKKHGATIMLAIGDAIGKIDDLTGALSKLSELHAFKLRVDPANFRILAHNIIILVLAMYFPADFTPEVHVSVDKFLQNLALGLSEKYR
<i>G. eurystoma</i> S	MSLSGKDKTVVKAFWDKIASKAPELGGEALGRMLTSYPQTKTYFAHWSDLSPESAQVKKHGATIMLAIGDAIGKIDDLAGALSKLSELHAFKLRVDPANFRVLAHNIIILVLAMYFPADFTPEVHVSVDKFLQNLALALSEKYH
<i>G. sexradiata</i> NS	MSLSGKDKTVVKAFWDKIASKAPELGGEALGRMLTSYPQTKTYFAHWSDLSPESAQVKKHGATIMLAIGDAIGKIDDLAGALSKLSELHAFKLRVDPANFRILAHNIIILVLAMYFPADFTPEVHVSVDKFLQNLALALSEKYR
<i>G. sexradiata</i> S	MSLSGKDKTVVKAFWDKIASKAPELGGEALGRMLTSYPQTKTYFAHWSDLSPESAQVKKHGATIMLAIGDAIGKIDDLAGALSKLSELHAFKLRVDPANFRILAHNIIILVLAMYFPADFTPEVHVSVDKFLQNLALALSEKYR
<i>L. perugiae</i> NS	MSLSGKDKTVVKAFWDKVSPKAAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIGKIDDLTGALSKLSELHAFKLRVDPANFRILAHNIIILVLALYYPADFTPEVHVSIDKFLQNLALALSEKYR
<i>L. sulphurophila</i> S	MSLSGKDKTVVKAFWDKVSPKAAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIGKIDDLTGALSKLSELHAFKLRVDPANFRILAHNIIILVLALYYPADFTPEVHVSIDKFLQNLALALSEKYR
<i>P. latipinna</i> NS	MSLSGKDKTVVKAFWDKVAPKTAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIGKIDDLQGALSKLSELHAFKLRVDPANFRILAHNIVLVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYR
<i>P. latipinna</i> S	MSLSGKDKTVVKAFWDKVAPKTAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIGKIDDLQGALSKLSELHAFKLRVDPANFRILAHNIIILVLAMYYPADFTPEVHVSIDKFLQNLALALSEKYR
<i>P. limantouri</i> NS	MSLSGKDKTVVKAFWDKVAPKAAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIGKIDDLQGALAKLSELHAFKLRVDPANFRILAHNIIILVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYR
<i>P. sulphuraria</i> S	MSLSGKDKSVVKAFWDKVAPKAAEMGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGAFIGMAIGDAIAKIDDLQGALAKLSELHAFKLRVDPANFRILAHNIIILVLAMYYPADFTPEVHVSIDKFLQNLALALSEKYR
<i>P. mexicana</i> (Tac) S	MSLSGKDKTVVKAFWDKVAPKAAEMGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIAKIDDLQGALAKLSELHAFKLRVDPANFRILAHNIIILVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYR
<i>P. mexicana</i> (Puy) NS	MSLSGKDKTVVKAFWDKVAPKAAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDTIGKIDDLQGALSKLSELHAFKLRVDPANFRIFAHNIIILVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYR
<i>P. mexicana</i> (Puy) S	MSLSGKDKTVVKAFWDKVAPKAAEMGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDAIAKIDDLQGALAKLSELHAFKLRVDPANFRILAHNIIILVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYH
<i>P. mexicana</i> (Pich) NS	MSLSGKDKTVVKAFWDKVAPKAAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDTIGKIDDLQGALSKLSELHAFKLRVDPANFRIFAHNIIILVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYR
<i>P. mexicana</i> (Tac) NS	MSLSGKDKTVVKAFWDKVAPKAAEIGGEALGRMLTSYPQTKTYFSHWSDLSPESAQVKKHGATIMGAIGDTIGKIDDLQGALSKLSELHAFKLRVDPANFRIFAHNIIILVLAMYYPADFTPEVHVSVDKFLQNLALALSEKYR

Draft

<i>X. hellerii</i> NS	MVEWTDDEERKAITTLWGQIDVGEIGPQALTRLLIVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSVMHSDKLHVDPDNFRALAE	CISVVVAAKFGPTVFTASFQEA	WQKFLAVVVSALGRQYH	
<i>X. hellerii</i> S	MVEWTDDEERKAITTLWGQIDVGEIGPQALTRLLIVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSVMHSDKLHVDPDNFRALAE	CISVVVAAKFGPTVFTASFQEA	WQKFLAVVVSALGRQYH	
<i>P. bimaculatus</i> NS	MVEWTDDAERKAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTAGFQEA	WQKFLAVVVSALGRQYH	
<i>P. bimaculatus</i> S	MVEWTDDAERKAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTAGFQEA	WQKFLAVVVSALGRQYH	
<i>G. holbrooki</i> NS	MVEWTDDAERKAITTLWGKIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNIDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPKVFTA	EFQEA	WQKFLAVVVSALGRQYH
<i>G. holbrooki</i> S	MVEWTDDAERKAITTLWGKIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNIDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPKVFTA	EFQEA	WQKFLAVVVSALGRQYH
<i>G. eurystoma</i> S	MVEWTDDAERKAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPSVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>G. sexradiata</i> NS	MVEWTDDAERKAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPSVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>G. sexradiata</i> S	MVEWTDDAERKAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPSVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>L. perugiae</i> NS	MVEWTDDAERQAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGSFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPSVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>L. sulphurophila</i> S	MVEWTDDAERQAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGSFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPSVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. latipinna</i> NS	MVEWTDDAERQAITTLWGQIDVGEIGQQALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. latipinna</i> S	MVEWTDDAERQAITTLWGQIDVGEIGQQALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVKNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. limantouri</i> NS	MVEWTDDAERQAITTLWGQIDVGEIGQQALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVRNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. sulphuraria</i> S	MVEWTDDAERQAITTLWGQIDVGEIGPQALTRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVVQHGGKTVMSGLETAVRNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. mexicana</i> (Tac) S	MVEWTDDAERQAITTLWGQIDVGEIGQQALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVRNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. mexicana</i> (Puy) NS	MVEWTDDAERQAITTLWGQIDVGEIGPKALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVRNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. mexicana</i> (Puy) S	MVEWTDDAERQAITTLWGQIDVGEIGQQALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVRNMDNIKNAYAKLSVMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. mexicana</i> (Pich) NS	MVEWTDDAERQAITTLWGQIDVGEIGPKALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVRNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	
<i>P. mexicana</i> (Tac) NS	MVEWTDDAERQAITTLWGQIDVGEIGPKALIRLLVVYPWTQRHFGAFGNLSTYAAIVGNAKVAQHGGKTVMSGLETAVRNMDNIKNAYAKLSIMHSEKLHVDPDNFRALAE	CISVVVAAKFGPTVFTPGFQEA	WQKFLAVVVSALGRQYH	

SUPPLEMENTARY MATERIAL

Molecular evolution and expression of oxygen transport genes in livebearing fishes (Poeciliidae) from hydrogen sulfide rich springs

Nicholas Barts, Ryan Greenway, Courtney N. Passow, Lenin Arias-Rodriguez, Joanna L. Kelley, and Michael Tobler

Table S1. List of background genes used for phylogenetic analyses and to estimate expression variation for EVE analyses. The table includes the gene ID from the *X. maculatus* reference genome annotation file (XM ID), the length (number of base pairs included in phylogenetic analyses), and the best supported model based on analyses of jModelTest. Note that genes with an asterisk were excluded from phylogenetic analyses due to alignment problems.

Gene	XM ID	Length (bp)	Model
AADAT	005807238.2	1278	HKY+G
ABCD3	014475225.1	1980	GTR+I+G
ABHD10	005797896.2	870	HKY+G
ABHD11	014474520.1	1077	HKY+G
ABHD112	005813987.1	954	HKY+G
ACADS	005811490.2	1368	GTR+I+G
ACSL5	005794720.1	2064	K80+I+G
ACSS2	005801880.1	2070	GTR+G
ACTC1	005806151.1	1131	SYM+I+G
ADSL	005811732.2	1449	GTR+G
ALAS2	005808942.2	2349	HKY+I+G
ALDH1L2	005808146.1	3375	GTR+I+G
ALDH2	005816457.1	1877	HKY+G
AMACR	005798702.1	1212	GTR+G
ANXA2	014475428.1	1017	HKY+G
ARMC4	014475557.1	2637	GTR+I+G
AURKAIP1	014474989.1	636	HKY+G
CASP8	005797663.1	2091	HKY+G
CCT3	005805489.1	1623	GTR+I+G
CCT7	005799601.1	2082	GTR+I+G
CGA	005801281.1	375	K80+G
Chaperonin	005816344.1	1737	GTR+I+G
CLIC4	005799133.2	756	GTR+G
CLYBL	014468401.1	1086	GTR+G
COMT	005813825.2	792	HKY+G
COPA	005796554.1	3861	GTR+I+G
CPSF1	005809367.1	4464	GTR+I+G
CPSF4	005808808.1	798	GTR+G
CSTF3	005813724.1	2560	GTR+G
CTSD	005808104.1	1206	GTR+I+G
CTSZ	005804262.2	909	GTR+G
CYP11B1	014473016.1	1626	HKY+I+G
CYP27A1	014468470.1	1554	HKY+I+G
DAZAP2	005815697.2	507	HKY+I
DDOST	005805756.2	1401	SYM+G
DDX19A	005797975.2	1467	GTR+G
DDX27	005812949.1	3246	GTR+G

DDX55	005815541.2	1770	GTR+I+G
ECH1	005796846.2	933	HKY+G
ECHS1	005813416.2	1122	GTR+I
ECI2	005796012.2	1212	HKY+G
ECT2	005816235.1	3072	HKY+I+G
EEF1A1P1	005805749.1	1386	GTR+I+G
EEF2	005797097.1	2577	GTR+G
EIF2S3	005812701.2	1497	GTR+G
ENDOG	014470985.1	903	HKY+G
FAM36A	005816157.2	444	K80+I
FOXRED1	005809315.1	1713	GTR+G
G6PD	005801747.2	1543	GTR+G
GAC	005807371.1	1842	GTR+I+G
GAPDH	005798762.2	1002	GTR+G
GCG	005808066.2	807	HKY+G
GK	005807820.1	2184	GTR+I+G
GLE1	005803874.2	2190	GTR+G
GLRX	005814739.2	318	K80+G
GLRX2	005800858.2	741	GTR+I
GTPBP3	014469449.1	1563	HKY+I+G
HADHB	005803415.1	1422	HKY+G
HDAC1	005799138.2	1467	GTR+G
HDAC3	005813186.2	1287	GTR+G
HIGD1A	005809444.2	279	HKY+G
HK1	005804231.1	3930	GTR+G
HSP90AB1	005803118.1	2181	GTR+I+G
HSPB7	005805755.2	468	K80+G
HSPE1	005816343.1	297	K80+G
HTRA2	005801130.2	1344	GTR+G
IDH3B	005808461.1	1578	GTR+G
IGSF8	005796988.1	1905	HKY+I+G
IMMT	005816057.1	2529	GTR+G
KARS	005810419.1	1734	GTR+G
L2HGDH	005799262.1	1344	GTR+G
LDHB	014473567.1	1002	GTR+G
LOC102231964*	005809220.1		
MCEE	005812538.2	519	HKY+G
MCM2	005808547.1	2703	GTR+G
MCM7	014469884.1	2703	GTR+I+G
ME1	005798064.2	1863	GTR+G
METTL17	005796597.2	1395	GTR+G
MICU1	005810989.2	1452	GTR+I+G
MMAB	005804029.2	690	HKY+G
MR1	005807674.1	1068	K80+I+G
MRPL30	005798047.2	480	HKY+G
MRPL42	005797188.2	435	K80+G
MRPL52	005807216.2	384	HKY+G
MRPS2	005797990.2	807	HKY+I
MRPS25	005801682.2	516	HKY+G
MRPS9	005802575.2	1149	HKY+G
MTERF2*	005803267.1		
MTX2	005816134.1	861	GTR+G
NARS	005813657.2	1701	GTR+I+G
NDC80	014476171.1	2643	GTR+G
NFYC	005797337.2	1119	GTR+G
NME1	005811683.1	456	SYM+I+G
NOP56	005815162.1	1686	GTR+G

ODC1	005799115.2	1380	SYM+G
OGDH	005806877.1	3432	GTR+I+G
OGG1	005801686.2	1203	HKY+G
OPA3	005810552.2	468	HKY+G
OXCT1	005795042.2	1584	GTR+G
PABPC1	005797394.1	1905	GTR+G
PANK2	005806844.2	1299	GTR+G
PDK2	005815972.2	1227	GTR+G
PDSS2	005798304.2	1158	HKY+G
PGK1	014474319.1	1257	GTR+I+G
PGS1	005794622.1	1647	HKY+G
PKD2	005797595.2	2685	GTR+G
PLRG1	005810225.2	1947	GTR+G
POLR1A	014471940.1	5073	GTR+I+G
POLR2D	005808231.2	558	HKY+I
PPIE	005799140.2	906	GTR+G
PPIF	005794591.2	582	K80+G
PRDX5	005800759.2	573	GTR+I
PRODH	005810076.2	1884	GTR+I+G
PRODH2	005810843.1	1896	GTR+G
PRPF3	014473017.1	2091	GTR+G
PRTFDC1	005796013.2	675	K80+G
PSMC6	005806032.2	1170	GTR+G
PSPC1	005810205.2	1683	GTR+G
PTS	014473362.1	492	GTR+G
PYCR1	005801446.2	972	GTR+G
RAE1	005800637.1	1110	GTR+I+G
RFC5	005803838.2	1008	GTR+G
RPL13A	005801275.1	615	GTR+I+G
RPL24	005798066.1	471	GTR+I
RPL28	005798660.1	414	GTR+I+G
RPL3	005810301.1	1800	HKY+I+G
RPL35	005814550.1	369	GTR+G
RPL7	005797870.1	738	GTR+I+G
RPL8	005802986.1	771	GTR+I+G
RPS11*	005796715.2		
RPS15	NM_001286309.1	435	HKY+I+G
RPS18	005804726.1	456	GTR+G
RPS4XP21	005806410.1	789	GTR+I+G
RPS5	005796424.1	609	HKY+I+G
SEPT4Iso2	005801071.1	1869	HKY+G
SEPT4Iso3	005801071.1	1497	GTR+I+G
SFXN3	005801574.1	969	GTR+G
SHMT1	005808810.1	1449	GTR+G
SKP1	005807193.2	489	GTR+G
SLC25A14	005800355.1	1278	K80+G
SLC25A22*	005806958.1		
SLC25A27	005804585.2	858	GTR+G
SLC25A4	005807138.1	897	GTR+I+G
SLC30A6	014471433.1	1405	HKY+G
SLC35D1	005817323.1	972	GTR+I
SLIRP	005797730.1	330	K80+G
SRP68	005797485.2	2514	GTR+G
STAR	005808973.1	858	GTR+I+G
STOML2	014473560.1	1077	GTR+G
STXBP3	005805216.2	1782	GTR+I+G
SUCLA2	005811324.1	1401	GTR+I+G

TATDN3	005799224.2	834	K80+I+G
TBP	005804211.2	924	GTR+G
TDRKH	005811695.2	1752	GTR+G
TIMM9	005812066.1	267	K80
TRAP1	005809292.1	2148	GTR+G
TRIAP1	005813844.1	222	F81
TRNT1	005808484.2	1275	GTR+G
TXN2	005802800.1	510	HKY+I
TXNRD1	005807810.2	1803	GTR+I+G
U2AF1	005812217.2	684	GTR+I+G
UCP1	014476189.1	915	GTR+I+G
UFD1L	005794850.2	936	GTR+G
VAMP8	005802151.1	372	HKY+G
VAR5	014468686.1	5490	GTR+G
VPS18	005808318.1	2949	GTR+G
WEE1	005807460.2	1869	GTR+G
YBX1	005800098.1	897	GTR+I+G
YME1L1	014474304.1	2184	GTR+G
YWHAQ	005805652.1	732	GTR+I
ZBED5	014475757.1	4365	HKY+G

Draft

Table S2. Descriptive statistics (mean FPKM $\pm$ standard error) of expression levels across species for all oxygen-transport genes.

Lineage	MYO	HEMOBAL2	HEMOAA	HEMOAL2	HEMOBAL	HEMOB1L2	HEMOADL	HEMOAL	HEMOA1L	HEMOB1L	HEMOBL	HEMOESA2
<i>X. hellerii</i> NS	706.2 $\pm$ 109.8	19980.9 $\pm$ 3662.0	14392.8 $\pm$ 2511.5	9635.8 $\pm$ 2598.3	2718.1 $\pm$ 735.8	161.3 $\pm$ 56.5	26.4 $\pm$ 7.0	0.3 $\pm$ 15.0	43.7 $\pm$ 4.3	1.6 $\pm$ 5.8	7.9 $\pm$ 1.4	0.4 $\pm$ 0.7
<i>X. hellerii</i> S	452.1 $\pm$ 109.8	8910.2 $\pm$ 3662.0	7696.4 $\pm$ 2511.5	4740.2 $\pm$ 2598.3	1809.5 $\pm$ 735.8	159.8 $\pm$ 56.5	17.8 $\pm$ 7.0	0 $\pm$ 15.0	15.4 $\pm$ 4.3	0.8 $\pm$ 5.8	2.2 $\pm$ 1.4	0.1 $\pm$ 0.7
<i>P. bimaculatus</i> NS	35.6 $\pm$ 120.3	9389.4 $\pm$ 4011.5	15124.1 $\pm$ 2751.2	10164.7 $\pm$ 2846.3	2815.0 $\pm$ 806.1	165.3 $\pm$ 61.9	15.1 $\pm$ 7.6	21.8 $\pm$ 16.5	6.8 $\pm$ 4.7	22.9 $\pm$ 6.3	0.1 $\pm$ 1.6	1.6 $\pm$ 0.8
<i>P. bimaculatus</i> S	91.2 $\pm$ 109.8	11883.7 $\pm$ 3662.0	12207.1 $\pm$ 2511.5	10546.4 $\pm$ 2598.3	3479.0 $\pm$ 735.8	228.0 $\pm$ 56.5	12.6 $\pm$ 7.0	0.9 $\pm$ 15.0	0.6 $\pm$ 4.3	0.6 $\pm$ 5.8	0.0 $\pm$ 1.4	0.0 $\pm$ 0.7
<i>G. holbrooki</i> NS	370.5 $\pm$ 109.8	18534.1 $\pm$ 3662.0	11579.9 $\pm$ 2511.5	17786.9 $\pm$ 2598.3	4859.3 $\pm$ 735.8	414.8 $\pm$ 56.5	27.1 $\pm$ 7.0	58.5 $\pm$ 15.0	45.0 $\pm$ 4.3	17.2 $\pm$ 5.8	2.3 $\pm$ 1.4	0.3 $\pm$ 0.7
<i>G. holbrooki</i> S	228.5 $\pm$ 109.8	17101.3 $\pm$ 3662.0	14280.9 $\pm$ 2511.5	12793.4 $\pm$ 2598.3	5098.4 $\pm$ 735.8	329.4 $\pm$ 56.5	31.3 $\pm$ 7.0	0.3 $\pm$ 15.0	33.8 $\pm$ 4.3	0.2 $\pm$ 5.8	0.2 $\pm$ 1.4	0.2 $\pm$ 0.7
<i>G. eurystoma</i> S	66.9 $\pm$ 109.8	16652.4 $\pm$ 3662.0	16322.7 $\pm$ 2511.5	15752.5 $\pm$ 2598.3	4759.0 $\pm$ 735.8	389.5 $\pm$ 56.5	42.3 $\pm$ 7.0	79.1 $\pm$ 15.0	9.9 $\pm$ 4.3	18.5 $\pm$ 5.8	4.5 $\pm$ 1.4	0.2 $\pm$ 0.7
<i>G. sexradiata</i> NS	147.0 $\pm$ 109.8	13914.2 $\pm$ 3662.0	11977.2 $\pm$ 2511.5	10680.3 $\pm$ 2598.3	3104.0 $\pm$ 735.8	206.3 $\pm$ 56.5	22.0 $\pm$ 7.0	5.2 $\pm$ 15.0	4.6 $\pm$ 4.3	3.9 $\pm$ 5.8	5.0 $\pm$ 1.4	0.1 $\pm$ 0.7
<i>G. sexradiata</i> S	167.7 $\pm$ 109.8	24894.6 $\pm$ 3662.0	20033.6 $\pm$ 2511.5	20174.1 $\pm$ 2598.3	3912.8 $\pm$ 735.8	290.8 $\pm$ 56.5	31.6 $\pm$ 7.0	22.9 $\pm$ 15.0	12.8 $\pm$ 4.3	4.4 $\pm$ 5.8	3.2 $\pm$ 1.4	0.1 $\pm$ 0.7
<i>L. perngiae</i> NS	358.3 $\pm$ 109.8	3323.4 $\pm$ 3662.0	2639.4 $\pm$ 2511.5	1925.4 $\pm$ 2598.3	1153.2 $\pm$ 735.8	103.5 $\pm$ 56.5	10.2 $\pm$ 7.0	1.5 $\pm$ 15.0	0.3 $\pm$ 4.3	2.9 $\pm$ 5.8	0.4 $\pm$ 1.4	0.1 $\pm$ 0.7
<i>L. sulphurophila</i> S	233.3 $\pm$ 109.8	5641.9 $\pm$ 3662.0	6652.4 $\pm$ 2511.5	3208.6 $\pm$ 2598.3	3493.6 $\pm$ 735.8	148.2 $\pm$ 56.5	28.0 $\pm$ 7.0	16.7 $\pm$ 15.0	3.4 $\pm$ 4.3	23.2 $\pm$ 5.8	3.2 $\pm$ 1.4	0.5 $\pm$ 0.7
<i>P. latipinna</i> NS	690.4 $\pm$ 109.8	10642.7 $\pm$ 3662.0	6816.8 $\pm$ 2511.5	11654.8 $\pm$ 2598.3	2562.5 $\pm$ 735.8	576.8 $\pm$ 56.5	21.5 $\pm$ 7.0	1 $\pm$ 15.0	6.6 $\pm$ 4.3	0.4 $\pm$ 5.8	9.4 $\pm$ 1.4	3.3 $\pm$ 0.7
<i>P. latipinna</i> S	1038.4 $\pm$ 109.8	8831.7 $\pm$ 3662.0	4533.5 $\pm$ 2511.5	10953.2 $\pm$ 2598.3	1375.3 $\pm$ 735.8	243.8 $\pm$ 56.5	11.1 $\pm$ 7.0	9.2 $\pm$ 15.0	5.9 $\pm$ 4.3	0.8 $\pm$ 5.8	0.7 $\pm$ 1.4	1.6 $\pm$ 0.7
<i>P. limantouri</i> NS	138.0 $\pm$ 109.8	5455.8 $\pm$ 3662.0	6563.3 $\pm$ 2511.5	3978.8 $\pm$ 2598.3	2143.9 $\pm$ 735.8	89.1 $\pm$ 56.5	14.8 $\pm$ 7.0	3.9 $\pm$ 15.0	0.3 $\pm$ 4.3	0.0 $\pm$ 5.8	0.2 $\pm$ 1.4	1.9 $\pm$ 0.7
<i>P. sulphuraria</i> S	2581.5 $\pm$ 109.8	55517.8 $\pm$ 3662.0	33158.6 $\pm$ 2511.5	32174.9 $\pm$ 2598.3	5442.8 $\pm$ 735.8	1081.1 $\pm$ 56.5	111.5 $\pm$ 7.0	21.7 $\pm$ 15.0	1.5 $\pm$ 4.3	2.8 $\pm$ 5.8	0.2 $\pm$ 1.4	2.8 $\pm$ 0.7
<i>P. mexicana</i> (Tac) S	1199.2 $\pm$ 109.8	12240.6 $\pm$ 3662.0	4051.7 $\pm$ 2511.5	8390.6 $\pm$ 2598.3	490.3 $\pm$ 735.8	161.5 $\pm$ 56.5	26.3 $\pm$ 7.0	0.8 $\pm$ 15.0	0.5 $\pm$ 4.3	0.4 $\pm$ 5.8	0.0 $\pm$ 1.4	0.4 $\pm$ 0.7
<i>P. mexicana</i> (Puy) NS	289.8 $\pm$ 109.8	4567.2 $\pm$ 3662.0	7030.5 $\pm$ 2511.5	1788.5 $\pm$ 2598.3	1625.4 $\pm$ 735.8	146.3 $\pm$ 56.5	17.4 $\pm$ 7.0	1.2 $\pm$ 15.0	0.0 $\pm$ 4.3	2.5 $\pm$ 5.8	5.0 $\pm$ 1.4	3.1 $\pm$ 0.7
<i>P. mexicana</i> (Puy) S	833.0 $\pm$ 120.3	12445.7 $\pm$ 4011.5	8878.2 $\pm$ 2751.2	7522.9 $\pm$ 2846.3	2248.9 $\pm$ 806.1	190.5 $\pm$ 61.9	38.1 $\pm$ 7.6	4.3 $\pm$ 16.5	10.7 $\pm$ 4.7	1.5 $\pm$ 6.3	2.2 $\pm$ 1.6	3.1 $\pm$ 0.8
<i>P. mexicana</i> (Pich) NS	506.9 $\pm$ 109.8	6718.5 $\pm$ 3662.0	9995.3 $\pm$ 2511.5	1724.2 $\pm$ 2598.3	1734.0 $\pm$ 735.8	137.8 $\pm$ 56.5	21.7 $\pm$ 7.0	0.5 $\pm$ 15.0	0.6 $\pm$ 4.3	0.5 $\pm$ 5.8	8.0 $\pm$ 1.4	1.3 $\pm$ 0.7
<i>P. mexicana</i> (Tac) NS	267.3 $\pm$ 109.8	4056.4 $\pm$ 3662.0	4825.4 $\pm$ 2511.5	1103.2 $\pm$ 2598.3	815.1 $\pm$ 735.8	98.1 $\pm$ 56.5	14.5 $\pm$ 7.0	1.1 $\pm$ 15.0	3.9 $\pm$ 4.3	3.7 $\pm$ 5.8	3.0 $\pm$ 1.4	0.3 $\pm$ 0.7

<i>X. helleri</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>X. helleri</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. bimaculatus</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. bimaculatus</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>G. holbrooki</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>G. holbrooki</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>G. eurystoma</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>G. eurystoma</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>G. seccardiata</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>L. perugiae</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>L. sulphuriphila</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. latipinna</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. latipinna</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. limantouri</i> NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. sulphuraria</i> S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. mexicana</i> (Tac) S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. mexicana</i> (Puy) NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. mexicana</i> (Puy) S	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. mexicana</i> (Pich) NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA
<i>P. mexicana</i> (Tac) NS	MVEWTD AERTAI STLWSNIDVGEIGPQALSRLLVVFPWTQRYFPTFGDLSTPAAIAANPKVAQHGTKVMGGLETA

Figure S1: Amino acid sequence alignment of HEMOBAL2, which exhibited ω_s was higher than ω_{NS} , but no evidence for a significant a branch-site test (see Table 4). Branch tests with a single lineage in the foreground revealed significantly elevated values of ω for *G. eurystoma*, *P. sulphuraria*, and *P. mexicana* (Tacotalpa) (results not shown), which all exhibit amino acid substitutions at unique codon positions (highlighted in red).