The evolutionary ecology of extremophile fishes from hypersaline lakes, caves, and toxic streams

by

Elizabeth Jane Wilson

B.S., Truman State University, 2018
B.A., Truman State University, 2018

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Abstract

Transitions of organisms across habitat boundaries are widespread in nature. Studying such transitions can provide insight into how biodiversity arises as organisms undergo phenotypic and ecological changes when they adapt to novel environmental pressures. Furthermore, systems in which populations independently colonize novel environments with similar selective forces allow for investigating the predictability of biological processes. For my dissertation, I leveraged four systems of extremophile fishes to make comparisons across multiple populations that have overcome similar environmental stressors. (1) I analyzed patterns of differential gene expression between a freshwater and a hypersaline population of a livebearing fish, *Limia perugiae*, and compared these gene expression patterns to other pairs of freshwater and saltwater populations of other fish species to understand if mechanisms of osmoregulation are shared among teleost fishes. I found key differentially expressed genes underlying salinity transitions in *L. perugiae*, but there was little evidence for convergence in gene expression across lineages, suggesting that—at least at the level of gene expression—fish use different mechanisms to overcome salinity barriers. (2) I conducted a study of the natural history and trophic ecology of *Astyanax mexicanus*, a tetra that has colonized multiple caves in Mexico. I found that cavefish had substantial variation among cave populations in body size, sex ratios, and patterns of resource exploitation. But even though cavefish consume a broad array of resources, they still experience nutrient limitation during both the rainy and dry seasons, contradicting the previously held assumption that they experience intermittent starvation. (3) I tested how repeated colonization of toxic sulfide springs has driven changes in host-microbiome associations of a species of livebearing fish, *Poecilia mexicana*. I found convergence in the microbiomes of sulfidic *P. mexicana* across four replicate drainages, as well as shared core microbes among the sulfidic

lineages that constitute candidate microbes for future research on their potential role in adaptive symbioses. (4) I leveraged 23 populations of livebearing fishes that broadly vary in their ecology to investigate what factors shape host-associated microbiomes and to test in an explicit phylogenetic context whether repeated adaptation to sulfide springs results in microbiome convergence. I found that the microbiome dissimilarity between lineages was correlated with the environmental microbiome dissimilarity and the host's phylogenetic distance, indicating that the environment and the host's evolutionary history contribute to host microbiome assembly. I further found microbiome convergence among the sulfidic lineages, demonstrating that shared physiochemical stressors drive convergent associations between hosts and their microbial communities. Overall, my dissertation provides insight into the commonalities and differences in how populations respond to novel sources of selection in extreme environments, which has broad implications for understanding patterns of biodiversity and the predictability of biological processes.

The evolutionary ecology of extremophile fishes from hypersaline lakes, caves, and toxic streams

by

Elizabeth Jane Wilson

B.S., Truman State University, 2018
B.A., Truman State University, 2018

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Michael Tobler

Copyright

Abstract

Transitions of organisms across habitat boundaries are widespread in nature. Studying such transitions can provide insight into how biodiversity arises as organisms undergo phenotypic and ecological changes when they adapt to novel environmental pressures. Furthermore, systems in which populations independently colonize novel environments with similar selective forces allow for investigating the predictability of biological processes. For my dissertation, I leveraged four systems of extremophile fishes to make comparisons across multiple populations that have overcome similar environmental stressors. (1) I analyzed patterns of differential gene expression between a freshwater and a hypersaline population of a livebearing fish, *Limia perugiae*, and compared these gene expression patterns to other pairs of freshwater and saltwater populations of other fish species to understand if mechanisms of osmoregulation are shared among teleost fishes. I found key differentially expressed genes underlying salinity transitions in *L. perugiae*, but there was little evidence for convergence in gene expression across lineages, suggesting that—at least at the level of gene expression—fish use different mechanisms to overcome salinity barriers. (2) I conducted a study of the natural history and trophic ecology of *Astyanax mexicanus*, a tetra that has colonized multiple caves in Mexico. I found that cavefish had substantial variation among cave populations in body size, sex ratios, and patterns of resource exploitation. But even though cavefish consume a broad array of resources, they still experience nutrient limitation during both the rainy and dry seasons, contradicting the previously held assumption that they experience intermittent starvation. (3) I tested how repeated colonization of toxic sulfide springs has driven changes in host-microbiome associations of a species of livebearing fish, *Poecilia mexicana*. I found convergence in the microbiomes of sulfidic *P. mexicana* across four replicate drainages, as well as shared core microbes among the sulfidic

lineages that constitute candidate microbes for future research on their potential role in adaptive symbioses. (4) I leveraged 23 populations of livebearing fishes that broadly vary in their ecology to investigate what factors shape host-associated microbiomes and to test in an explicit phylogenetic context whether repeated adaptation to sulfide springs results in microbiome convergence. I found that the microbiome dissimilarity between lineages was correlated with the environmental microbiome dissimilarity and the host's phylogenetic distance, indicating that the environment and the host's evolutionary history contribute to host microbiome assembly. I further found microbiome convergence among the sulfidic lineages, demonstrating that shared physiochemical stressors drive convergent associations between hosts and their microbial communities. Overall, my dissertation provides insight into the commonalities and differences in how populations respond to novel sources of selection in extreme environments, which has broad implications for understanding patterns of biodiversity and the predictability of biological processes.

Table of Contents

List of Figures xii
List of Tables xvii
Acknowledgements xxi
Dedication xxiii
Preface xxiv
Chapter 1 - Gene expression signatures of salinity transitions in *Limia perugiae* (Poeciliidae), with comparisons to other teleosts 1
Abstract 1
Introduction 2
Methods 6
Sample collection 6
RNA-seq library preparation 6
Mapping 7
Differential gene expression 7
Weighted gene co-expression network analysis 8
Comparisons of L. perugiae with other species 10
Results 12
Comparative analysis of freshwater and hypersaline L. perugiae 12
Comparisons of Limia with phylogenetically disparate teleosts 14
Discussion 15
Responses to variation in salinity in a freshwater and hypersaline population of L. perugiae 16
Regulation of transmembrane transport and gill epithelial permeability 16
Cell cycle regulation 18
Regulation of cell signaling 19
Comparisons among disparate lineages 20
Lineage-specific responses to variation in salinity 20
Shared responses involving ion transport and immune function 22
Tables 24

Figures 29
Chapter 2 - Natural history and trophic ecology of three populations of the Mexican cavefish, *Astyanax mexicanus* 33
Abstract 33
Introduction 34
Methods 37
Study area 37
Biometrics 38
Gut content analysis 39
Results 40
Biometrics 40
Gut content analysis 41
Discussion 44
Sex ratios, body size, and energetics in subterranean environments 44
Nutrient limitation and starvation 47
Trophic ecology and seasonal variation in diet 49
Conclusions 52
Tables 53
Figures 56
Chapter 3 - Host-microbiome associations in livebearing fishes adapted to toxic environments rich in hydrogen sulfide 60
Abstract 60
Introduction 61
Methods 65
Sample collection 65
DNA extraction, amplification, and sequencing 66
Processing reads, assigning taxonomy, and filtering taxa 67
Analyzing commonalities and differences in gut microbiomes between sulfidic and nonsulfidic ecotypes and across river drainages 68
Comparing gut microbiomes to environmental microbiomes and analyzing host-specific microbial communities 70

Identifying potentially adaptive host-associated microbes in sulfide springs 70
Testing for evidence of host regulation of the microbiome 71
Results 72
What are the commonalities and differences in the gut microbiomes of P. mexicana populations? 72
How do gut microbiomes compare to environmental microbiomes, and how do patterns of diversity vary among microbial communities that are primarily host-associated? 73
What are potentially adaptive host-associated microbes of sulfide spring fishes? 75
Is there any evidence that hosts can control the composition of their gut microbiomes? 77
Discussion 78
Convergent shifts in microbiome composition in sulfide springs 78
Potentially adaptive host-associated microbes of sulfidic fish 80
Conclusions 84
Tables 85
Figures 86
Chapter 4 - Comparative phylogenetic analysis of Poeciliid microbiomes: effects of evolutionary history and the environment 88
Abstract 88
Introduction 89
Methods 92
Sample collection 92
DNA extraction and 16S rRNA amplicon sequencing 93
Processing the sequencing reads 95
Analyzing variation in microbiomes 95
Testing for effects of evolutionary history and environment 96
Whole-microbiomes vs. primarily host-associated microbiomes 97
Testing for convergence in host-associated microbiomes using phylogenetic comparative methods 98
Results 100
Variation in environmental and host-associated microbiomes 100
Effects of evolutionary history and environment 101

Honing in on primarily host-associated microbes 102
Microbiome convergence among lineages of extremophile livebearing fishes 103
Discussion 105
How are host-associated microbiomes shaped by the host's evolutionary history and environment? 106
Is there evidence of convergence in the microbiomes of hosts adapted to extreme environments? 109
Conclusion 112
Tables 113
Figures 115
References 119
Appendix A - Gene expression signatures of salinity transitions in *Limia perugiae* (Poeciliidae), with comparisons to other teleosts 142
Appendix A Tables 142
Appendix B - Natural history and trophic ecology of three populations of the Mexican cavefish, *Astyanax mexicanus* 143
Appendix B Tables 143
Appendix C - Host-microbiome associations in livebearing fishes adapted to toxic environments rich in hydrogen sulfide 144
Appendix C Tables 144
Appendix C Figures 145
Appendix D - Comparative phylogenetic analysis of Poeciliid microbiomes: effects of evolutionary history and the environment 152
Appendix D Tables 152
Appendix D Figures 153

List of Figures

Figure 1-1. A. Volcano plot depicting differentially expressed genes between hypersaline and freshwater *Limia perugiae*. Genes that were significantly differentially expressed between hypersaline and freshwater populations (FDR < 0.05) are indicated by the blue and red points—blue points represent genes downregulated in the hypersaline populations, while red points represent upregulated genes. B. Multi-dimensional scaling (MDS) plot of hypersaline and freshwater *L. perugiae* gene expression profiles. MDS axis 1 separated samples by freshwater vs. hypersaline environments. 29

Figure 1-2. Weighted gene co-expression network analysis. A. Average linkage clustering tree based on topological overlap distances in gene expression patterns of *L. perugiae* from freshwater and saltwater habitats. Branches of the dendrogram correspond to modules, as shown in the color bars below. B. Correlation between module eigenvalues and habitat type (freshwater *vs*. saltwater). Each row corresponds to a module of coexpressed genes, and values are Pearson correlation coefficients (left column) and *P*-values (right column in parentheses). Color coloration scales with the correlation coefficient according to the scale bar to the right. 30

Figure 1-3. A. Multi-dimensional scaling plot (MDS) of the general expression patterns of all the populations included in our analysis. B. Similarity of gene expression profiles of saltwater (SW) and freshwater (FW) populations across different lineages. The majority of variation in gene expression reflects phylogenetic divergence among lineages. C. Shared differentially expressed genes across lineages. The large, central number in each section represents the total number of shared differentially expressed genes among the lineages in that intersection. The top number in each section represents the number of shared up-regulated genes, and the bottom number is the number of shared down-regulated genes in that intersection. Only 10 genes were consistently differentially expressed between all of the SW and FW populations. 31

Figure 1-4. Examples of expression variation in shared differentially expressed genes. Nine of the ten shared differentially expressed genes have annotations, and those genes are included here. Magnitude and direction of differential expression is lineage specific..... 32

Figure 2-1. A. The location of the study region in northeastern Mexico is indicated by the gray square. B. Caves studied here are indicated by green circles. Black circles indicate the location of other caves inhabited by *A. mexicanus*. Maps were created in R using data associated with the rnaturalearth (South 2017) and elevatr packages (Hollister 2021)..... 56
Figure 2-2. Large circles with bars represent marginal means and standard errors of standard length across the Pachón, Sabinos, and Tinaja caves, as inferred by the interaction model: standard length ~ sex * cave + (1|Date). Small circles represent the raw data used in the model. Males are indicated in orange, females in blue..... 57
Figure 2-3. A. Frequency of individuals with empty guts in the Pachón and Sabinos populations across multiple sampling times. B. The average number (and standard error) of diet categories per individual for the Pachón and Sabinos populations across multiple sampling times..... 58
Figure 2-4. A. NMDS plot showing the gut composition (means and standard deviations) of caves and sampling times for the Pachón and Sabinos populations. Ordination ranking was conducted using Jaccard distance, and the stress with 3 dimensions is 0.079. The plot points for Pachón and Sabinos are numbered in order corresponding to their sampling times as follows: for Pachón, P1 = 2/3/2001, P2 = 4/19/2001, P3 = 6/24/2001, P4 = 11/3/2001, P5 = 2/21/2002, P6 = 3/22/2002; for Sabinos, S1 = 4/18/2001, S2 = 6/23/2001, S3 = 11/2/2001. B. Precipitation (in mm) during a typical year near the Sierra de El Abra region (monthly average rainfall data from Tampico, Mexico; World Weather Online 2021). Points P1–P6 and S1–S3 correspond to the sampling times in Figure 2-4A. 59
Figure 3-1. NMDS plot based on Bray-Curtis dissimilarity of sulfidic and nonsulfidic *P. mexicana* gut microbial communities A) before environmental correction and B) after environmental correction. 86
Figure 3-2. Venn diagram showing the shared and unique core microbes among all four drainages when comparing A) the whole microbiome (before environmental correction) and B) the primarily host-associated microbiome (after environmental correction). The numbers of core microbial taxa from the nonsulfidic and sulfidic microbiomes combined are the top numbers in the Venn diagram, while the numbers of core microbial taxa from just the nonsulfidic populations are the middle numbers, and the numbers of core microbial taxa from just the sulfidic populations are the bottom numbers..... 86

Figure 3-3. The top microbial taxa driving differences between sulfidic and nonsulfidic *P. mexicana* gut microbial community composition A) before environmental correction and B) after environmental correction. Taxa in nonsulfidic lineages are indicated by blue bars, and taxa in sulfidic lineages are indicated by yellow bars. Taxa are labeled to the genus level or the lowest identifiable taxonomic level, and family names are listed in parentheses next to each genus. 87
Figure 3-4. NMDS plot showing gut microbiome variation of sulfidic and nonsulfidic ecotypes of *P. mexicana* raised in a common-garden environment. 87
Figure 4-1. A) Geographic region of all the sampling sites. Orange points represent nonsulfidic sites, and yellow points represent sulfidic sites. The sulfidic cave site is represented by a gray point. B) Phylogenetic tree showing the evolutionary relatedness between host lineages. Sulfidic lineages are highlighted in yellow, and the sulfidic and cave lineage is highlighted in gray. 115
Figure 4-2. NMDS based on Bray-Curtis dissimilarity showing gut and environmental microbiome variation. Samples from sulfidic environments are marked as yellow points, and samples from nonsulfidic environments are marked as blue points. Gut samples are represented by circles, and environmental samples are represented by triangles. 116
Figure 4-3. NMDS based on Bray-Curtis dissimilarity showing the gut microbiome variation of all the lineages. The points are NMDS coordinate averages of each lineage's microbiome composition, and the lines connecting the points reflect the host phylogenetic relationships. 117
Figure 4-4. Correlations between gut microbiome dissimilarity and A) environmental microbiome dissimilarity (Mantel correlation: 0.271; $P = 0.0008$) and B) host phylogenetic distance (Mantel correlation: 0.238; $P = 0.0017$). 118

Appendix Figure C-1. Map of the sampling sites. Sampling occurred along the Pichucalco, Ixtapangajoya, Puyacatengo, and Tacotalpa river drainages in the Río Grijalva basin. Water, sediment, and fish gut microbiome samples were taken from a pair of sulfidic and nonsulfidic streams in each river drainage. Sulfidic sampling sites are indicated by yellow triangles, and nonsulfidic sampling sites are indicated by blue triangles. 145

Appendix Figure C-2. Rarefaction curves generated in QIIME2 of the (A) gut microbiome samples from the field, (B) environmental samples from the field, and (C) gut samples from the laboratory showing the number of observed features (sequence variants) relative to sequencing depth. 146
Appendix Figure C-3. Alpha diversity metrics of the gut microbiomes of all *P. mexicana* population pairs, as measured by: A) Shannon diversity index, B) Pielou evenness, C) Chao1 richness, and D) relative dominance 147
Appendix Figure C-4. NMDS plot based on Bray-Curtis dissimilarity of all gut and environmental samples 148
Appendix Figure C-5. Alpha diversity metrics of the gut microbiomes of all *P. mexicana* population pairs after environmental correction, as measured by: A) Shannon diversity index, B) Pielou evenness, C) Chao1 richness, and D) relative dominance. 149
Appendix Figure C-6. The top microbial taxa driving differences between sulfidic and nonsulfidic *P. mexicana* gut microbial community composition in A) the Ixtapangajoya drainage, B) the Pichucalco drainage, C) the Puyacatengo drainage, and D) the Tacotalpa drainage. Taxa in nonsulfidic lineages are indicated by blue bars, and taxa in sulfidic lineages are indicated by yellow bars. 150
Appendix Figure C-7. The top microbial taxa driving differences between sulfidic and nonsulfidic *P. mexicana* gut microbial community composition after environmental correction in A) the Ixtapangajoya drainage, B) the Pichucalco drainage, C) the Puyacatengo drainage, and D) the Tacotalpa drainage. Taxa in nonsulfidic lineages are indicated by blue bars, and taxa in sulfidic lineages are indicated by yellow bars. 151
Appendix Figure D-1. QIIME2 rarefaction curves for A) the gut microbiome data and B) the environmental microbiome data showing the number of observed features (sequence variants) relative to sequencing depth. 153
Appendix Figure D-2. Scatterplot showing gut microbiome dissimilarity vs. environmental distance (based on bioclimatic variables), for which there was no significant correlation (Mantel correlation: 0.033; $P = 0.5959$). 154
Appendix Figure D-3. NMDS based on Bray-Curtis dissimilarity showing the primarily host-associated gut microbiome variation of all the lineages (after filtering out environmental microbes). The points are NMDS coordinate averages of each lineage's microbiome

composition, and the lines connecting the points reflect the host phylogenetic relationships. 155
Appendix Figure D-4. Correlations between gut microbiome dissimilarity of the primarily host-associated microbiomes and A) environmental microbiome dissimilarity (Mantel correlation: 0.358; $P < 0.0001$) and B) host phylogenetic distance (Mantel correlation: 0.204; $P = 0.0065$). C) There was no significant correlation between gut microbiome dissimilarity and environmental distance (Mantel correlation: –0.028; $P = 0.7201$). 156

List of Tables

Table 1-1. Species included in the analysis, including environment (freshwater [FW] or saltwater [SW]), collection location, sample size (N), NCBI Sequence Read Archive (SRA) accession numbers, and study reference. 24

Table 1-2. GO process terms with significant enrichment in genes upregulated and downregulated in the hypersaline ecotype of *Limia perugiae* (FDR < 0.05). The table includes the GO term ID, description, the degree of enrichment, the number of differentially expressed genes associated with the GO term (*N*), as well as *P* and FDR-corrected *q*-values. 25

Table 1-3. Shared differentially expressed genes among all four population pairs and their associated proteins. Based on the SWISS-PROT database, nine of the ten genes have experimental evidence for the existence of the protein associated with each gene. The tenth gene did not have a name match in the database, so it is not included. 27

Table 1-4. GO process terms with significant enrichment in orthogenes with significant differential expression in at least three lineages (FDR < 0.05). The table include the GO term ID, description, the degree of enrichment, the number of differentially expressed genes associated with the GO term (*N*), as well as *P* and FDR-corrected *q*-values. 28

Table 2-1. Individuals collected for biometrics and gut content analysis at time points between February 2001 and March 2002. $N_{biometrics}$ represents all the individuals included in the biometrics analyses, and the sex ratios are reported as the proportion of the population that is male. N_{gut} is the subset of individuals that were then investigated for their gut contents. Finally, N_{empty} is the number of individuals with empty guts, which is also represented as a percentage ($100* N_{empty}/N_{gut}$). 53

Table 2-2. Results of GLMM analyses that examined variation in (A) body size (measured as standard length) and (B) the frequency of males and females (sex ratios) and (C) results of GLM analysis that examined variation in the frequency of individuals with empty guts. Note that collection date is designated as a random effect (1|Date) in analyses of standard length and sex ratios. 54

Table 2-3. PERMANOVA results of gut content composition relative to cave of origin (Pachón and Sabinos), sampling time (February 2001-March 2002), and their interactions. 55

Table 3-1. Coordinate locations of all the sampling sites, as well as the number of samples (gut, water, and sediment) collected at each site. Numbers in parentheses indicate the final number of samples included in the analyses after rarefying the data. 85
Table 4-1. Coordinates of the sampling sites and the number of water, sediment, and intestine samples collected from all 13 sites. In total, there were 199 samples collected, but 195 samples remained in the analysis after removing 4 samples with fewer than 4,500 reads. In cases where not all of the collected samples were included in the analysis, the number of samples included in the analysis is listed in parentheses. There were 128 total intestine samples and 66 environmental samples (40 water samples and 26 sediment samples) included in the analysis. 113
Table 4-2. Microbial taxa showing convergent shifts among the sulfidic lineages. The convergent microbes were identified both before and after filtering the gut microbiome dataset to exclude microbes that were detected in the environment. Each microbe in the table was listed down to its lowest known taxonomic level. After environmental correction, five additional taxa showed convergent abundance shifts: *Terrimicrobium*, LKC2.127.25, *Paludibacter*, *Annamia* HOs24, and uncultured Phormidiaceae. 114

Appendix Table A-1. Mapping statistics for all four population pairs included in the analysis. Reads for *Limia perugiae* were mapped against the *Poecilia mexicana* genome, for *Leuciscus waleckii* against the *Cyprinus carpio* genome, for *Odontesthes argentinensis* and *Odontesthes bonariensis* against the *Menidia menidia* transcriptome, and for *Gasterosteus aculeatus* against the *Gasterosteus aculeatus* genome. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.”..... 142
Appendix Table A-2. Results of weighted gene co-expression network analysis for all 18,659 genes identified in our RNA-seq analysis of *Limia perugiae*. Each gene's corresponding module is included in the moduleColor column. The correlation coefficient between each gene's expression and salinity is also included (GS.enviro), as well as the *P*-value for each correlation coefficient (p.GS.enviro). The remaining columns contain the correlation coefficients between each gene and each module, as well as the P-values for each of those correlations. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.” 142

Appendix Table A-3. GO terms for the turquoise module from the WGCNA results. The GO processes highlighted in yellow are unique to the WGCNA results and were not associated with up-regulated genes in the hypersaline ecotype. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.” 142
Appendix Table A-4. GO terms for the royalblue module from the WGCNA results. The GO processes highlighted in yellow are unique to the WGCNA results and were not associated with down-regulated genes in the hypersaline ecotype. For comparison with other WGCNA module results, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.” 142
Appendix Table A-5. GO terms for the blue module from the WGCNA results. The GO processes highlighted in yellow are unique to the WGCNA results and were not associated with down-regulated genes in the hypersaline ecotype. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.”..... 142
Appendix Table B-1. Gut contents per fish at sampling times between February 2001 and March 2002 for the Pachón, Sabinos, and Tinaja caves. Absence of diet item categories are indicated by 0’s, and presence of diet item categories are indicated by 1’s. For the “Full” category, 1=full and 0=empty. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.” 143
Appendix Table C-1. Merging and filtering statistics for all samples. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.”..... 144
Appendix Table C-2. Read counts after trimming and merging for all of the field and laboratory samples..... 144
Appendix Table C-3. Core microbes shared among all four drainages before and after environmental filtering. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.” 144
Appendix Table D-1. Merging and filtering statistics for all samples. Due to large size, this table is included in a Microsoft Excel file called “ElizabethWilson2024_Appendix.xlsx.”..... 152
Appendix Table D-2. Read counts after trimming and merging for all of the intestine and environmental samples..... 152

Appendix Table D-3. Convergence metrics of the primarily host-associated microbiomes, indicating the degree of convergence in sulfidic fish gut microbiomes relative to nonsulfidic lineages. 152

Acknowledgements

I first would like to thank my advisor, Michi Tobler, for being such a great mentor during my PhD program. He helped develop my confidence in research and teaching, pushed me to think creatively, supported me in any opportunity I was interested in pursuing, and was always so encouraging and understanding when things were hard. I am so grateful to know him and work with him. I would also like to thank my committee: Sonny Lee, Lydia Zeglin, and Sarah DuRant. My committee members have helped me grow so much as a scientist over the past several years, and they have been so generous with their time and laboratory resources.

My time at Kansas State University has been greatly shaped by the fun, supportive community that I found in the Tobler lab. Thank you so much to Nick Barts, Henry Camarillo, John Coffin, Bryan Frenette, Ryan Greenway, Cassie Delich, Madison Nobrega, Nichole Perez-Nieves, Rachel McNemee, Hannah Hoffman-Colburn, Quinlyn LaFon, Noelle Schumann, Ayden Wilroy, Aaron George, Ethan Christopher, Erik Johnson, Soren Johnson, and Bethany Williams. Outside of the Tobler lab, I also had such an amazing support system of friends including Joel Steyer, Bliss Betzen Coffin, Sam Sharpe, Dustin Haskell, Jordan Blake Banks, Lauryn Scott, Alicia Dunton, and Claire Utzman. I have loved all the time we have spent together over the years, and knowing I always have someone to talk to and laugh with means the world to me.

My family has also been so important to me throughout my PhD program. I especially want to thank my parents, Dan and Erin Wilson, and my sister, Suzie Wilson, for being there for me at every step of my life and career. Beyond their love and support, my parents have taught me so much about following my passions and challenging myself. Suzie provides such a major source of guidance in my life and inspires me every day through her creativity and compassion, and having her by my side throughout graduate school has been invaluable for me. I am so

thankful for a family always willing to drop anything for each other, and I truly could not have done this without them. I would also like to thank my partner, Mathew Keith, for his constant encouragement over the past decade, for bringing so much peace and joy into my life, and for growing my own trust in myself through his belief in me.

Furthermore, thank you so much to the Universidad Juárez Autónoma de Tabasco, the Centro de Investigación e Innovación para la Enseñanza y el Aprendizaje, the Instituto Tecnológico de Ciudad Victoria, Lenin Arias-Rodriguez, and local communities and landowners for providing support and access to research sites. I am so grateful for these groups of people, and I have really valued working with them throughout my research experiences.

Finally, this work would not be possible without funding from the National Science Foundation, the Army Research Office, the Friends of the Sunset Zoo, the Kansas State University College of Arts and Sciences, the Kansas State University Graduate Student Council, and the Kansas State University Division of Biology.

Dedication

I would like to dedicate this work to my sister, Suzie Wilson, and my parents, Dan and Erin Wilson. They have supported every path that has led me to my current research journey, and they inspire me to keep pursuing what I love.

Preface

A major objective of evolutionary biology is to understand factors influencing patterns of biodiversity (Mahler et al. 2017). Consequently, transitions of organisms across habitat boundaries have been a focus of evolutionary research because they can give rise to phenotypic, ecological, and genetic diversification (Schluter 2000; Santini et al. 2013). As populations colonize novel environments, they are exposed to selective pressures that can drive changes in their phenotypic traits and ecology (Härer et al. 2023). Furthermore, cases in which independent populations repeatedly colonize similar environments provide an opportunity to investigate whether shared selective pressures drive similar changes in traits as populations adapt to the novel environmental conditions (Mahler et al. 2017). Idiosyncrasies in the outcomes of separate populations that colonized similar environments provide a foundation for investigating how factors such as variation among populations and random processes can contribute to biodiversity (Powell 2007). Conversely, patterns in which similar traits arise as a result of independent colonization of similar environments (i.e., patterns of convergence) provide unique opportunities to understand the repeatability of biological processes and the nature of biodiversity (Gould 1989; Lachapelle et al. 2015; Mahler et al. 2017).

Novel environments that have well-defined sources of selection make excellent systems for studying how populations respond to crossing a habitat boundary. Extreme environments are useful cases of such systems because they have strong physiochemical selective pressures, making them uninhabitable for most organisms (Tobler et al. 2018; Bang et al. 2018). As a result, populations that do manage to survive in extreme environments must overcome specific environmental conditions that are in stark contrast to the benign environments of their ancestral populations (Tobler et al. 2018). Studying how selection is driving trait divergence between

extremophile and benign populations is therefore useful for understanding how transitioning across an environmental gradient affects a population's phenotypic traits and ecology. Cases where replicated lineages inhabit similar extreme environments provide an added opportunity to compare outcomes of selection between different populations (Tobler et al. 2018).

Objectives

I leveraged systems in which populations of fishes have repeatedly colonized extreme environments with the aim of understanding how selection can drive shared or different outcomes in phenotypic traits and ecology. I studied four extremophile systems: livebearing fish of the species *Limia perugiae* that have colonized a hypersaline lake in the Dominican Republic, Mexican tetras (*Astyanax mexicanus*) that have colonized multiple caves throughout the Sierra de El Abra region in Mexico, livebearing fish of the *Poecilia mexicana* species complex that have adapted to toxic streams rich in hydrogen sulfide (H_2S) throughout the Río Grijalva basin in Mexico, and an assemblage of livebearing fishes that are diverse in their ecology and habitats (including lineages inhabiting toxic H_2S-rich streams) throughout the Río de la Sierra basin in Mexico. All of these populations experience strong selective pressures—such as high salinity, perpetual darkness, low nutrient availability, hypoxia, and the presence of toxic hydrogen sulfide—and the availability of evolutionarily replicated lineages allows for identifying shared and unique changes among populations, thereby shedding light on whether such changes are predictable. Among these systems, patterns of convergence have been explored at some biological levels, but there are several unexplored questions about how selection drives shared and unique changes at other levels. For example, distantly related fish species that have crossed salinity barriers show evidence for convergence in genes involved in osmoregulation (Velotta et al. 2022), but it is unknown whether patterns of gene expression are similarly shared. In different

populations of *A. mexicanus* cavefish, there is convergence in feeding posture associated with more effective foraging in the nutrient-limited cave environment (Kowalko et al. 2013), but it not well known what cavefish eat in nature or if there is variation in diet among cavefish populations. In populations of livebearing fishes that have adapted to toxic, H_2S-rich streams in southern Mexico, there is convergence in their H_2S tolerance at molecular and functional levels (Greenway et al. 2020), but an unexplored facet of their repeated colonization of sulfide springs is their host-associated microbial communities, despite well-described host-microbe symbioses with animals in other H_2S-rich environments (Dubilier et al. 1995; Beinart et al. 2019). My dissertation overall aimed to fill these knowledge gaps through four main objectives:

Chapter 1: Divergent selection regimes between extreme and benign environments often give rise to differential gene expression (Wang and Guo 2019), and investigating the changes in gene expression as populations colonize novel environments has provided insight into mechanisms underlying adaptation and acclimation in several systems (Fisher and Oleksiak 2007; Xu et al. 2013; Wang and Guo 2019). However, it remains unknown in many systems whether patterns of gene expression are shared or unique across multiple lineages that have undergone similar habitat transitions. In my first chapter, I investigated differential gene expression between a freshwater and hypersaline population of *Limia perugiae* to better understand how this livebearing fish species has colonized a hypersaline lake in the Dominican Republic. Transitions across salinity clines are widespread among fishes, but comparisons in gene expression patterns across distantly related fishes that have crossed salinity barriers are lacking. I therefore identified shared differentially expressed genes among freshwater and saltwater population pairs of distantly related fish species to understand if fishes modulate their gene expression in similar ways when crossing salinity barriers. I found differentially expressed

genes underlying the transition of *L. perugiae* from a freshwater to a hypersaline environment. I found upregulation of genes in the hypersaline population that are involved in ion transport and cell signaling, as well as downregulation of genes involved in the cell cycle and protein folding. These findings provide insight into how *L. perugiae*, and species of the family Poeciliidae as a whole, make transitions across salinity barriers. After comparing patterns of gene expression in the *L. perugiae* population pair with other pairs of freshwater and saltwater fish populations, I found little evidence of convergence among the distantly related lineages. This finding suggests that factors such as variation in other sources of selection associated with salinity transitions and the underlying genetic divergence of the lineages investigated may lead to differences in gene expression patterns.

Chapter 2: The physiochemical stressors in extreme environments often change or limit the nutrient availability for organisms inhabiting them (Bang et al. 2018). The cavefish *Astyanax mexicanus* is a model organism in evolutionary biology that has several well-documented adaptations for surviving in nutrient-poor subterranean environments (Kowalko et al. 2013; Aspiras et al. 2015; Riddle et al. 2018; Olsen et al. 2023). It has historically been well-documented that since cavefish can withstand long-term starvation in the laboratory (Salin et al. 2010; Aspiras et al. 2015), cavefish in nature likely experience intermittent periods of starvation that coincide with seasonal nutrient influx, but there have been few studies investigating what cavefish eat in nature (Wilkens and Burns 1972; Espinasa et al. 2017). It is also not known how diet may vary across cavefish populations, despite documented variation in environmental and geographic characteristics of the caves in the Sierra de El Abra region (Keene et al. 2015). I therefore analyzed variation in body size, sex ratios, proportions of individuals with empty guts, and diet composition across the rainy and dry seasons for populations of *A. mexicanus* that have

independently colonized three caves (Pachón, Sabinos, and Tinaja) to shed light on commonalities and differences in their natural history and trophic ecology across populations. I found variation in *A. mexicanus* cavefish populations that may be attributed to variation between the extreme environments. For example, I found differences in body size between cave populations that may reflect differences in flooding and nutrient availability in the caves, or they may reflect different strategies for coping with combined stressors in subterranean environments. However, I also found that cavefish overall showed commonalities in their variety of diet items, supporting that they are opportunistic feeders. Throughout the rainy and dry seasons, individuals also consistently had food in their guts, suggesting that cavefish may experience continual nutrient limitation.

Chapter 3: In extreme environments, selection may also act on host-microbiome interactions, which can facilitate host adaptation (Bang et al. 2018; Beinart et al. 2019; Breusing et al. 2020). Sulfide springs in the Río Grijalva basin are characterized by high levels of H_2S, the presence of hypoxia, and primary production from bacteria (Tobler et al. 2015, 2018; Greenway et al. 2020). Livebearing fish in the *P. mexicana* species complex have evolved a variety of strategies to cope with the stressors and ecological differences (Tobler et al. 2015, 2018; Greenway et al. 2020), but nothing is known about their host-associated microbial communities. Fish gut microbiomes have several beneficial effects on the host (Diwan et al. 2023), and host-associated bacteria in other sulfidic environments help their hosts survive by detoxifying H_2S and contributing to the host's nutrition (Cavanaugh et al. 2006), raising the question of whether adaptive symbioses could arise between fish and bacteria in sulfide springs. I compared the host-microbiome associations across four population pairs of sulfide-tolerant and sulfide-intolerant *P. mexicana* to identify how microbiomes were similar and different between ecotypes and across

drainages. I identified shared core microbes among the drainages, and I identified the top microbial taxa from the sulfidic fish, highlighting microbes to investigate further as potentially adaptive symbionts of extremophile fishes. I found that sulfidic and nonsulfidic fish have different gut microbiomes from each other, and I found that sulfidic fish share core microbes among the drainages. Furthermore, the top microbial taxa in the fish guts associated with sulfide spring colonization include taxa that have implications in hypoxia tolerance, nutrition, and H_2S detoxification in other animal systems.

Chapter 4: Finally, to understand the role of the microbiome in host adaptation, it is important to first identify factors that contribute to microbiome assembly. There are additional lineages of livebearing fishes in the same geographic region as *P. mexicana*, including some in the genera *Gambusia*, *Pseudoxiphophorus*, and *Xiphophorus* that also inhabit sulfide streams. I therefore investigated how evolutionary and environmental factors shape host-associated microbiomes among 23 populations of livebearing fishes in this region, and I explored how sulfide spring colonization has shaped the gut microbiomes of extremophile lineages. I specifically tested how gut microbiome dissimilarity was associated with host phylogenetic distance, environmental microbiome dissimilarity, and environmental distance based on bioclimatic variables. I then used phylogenetic comparative methods to investigate whether sulfide-adapted lineages of livebearing fishes showed convergence in their gut microbiomes. I found that the host-associated microbiomes were correlated with the host phylogeny and environmental microbiomes, but the gut microbiome dissimilarity was most strongly correlated with the environmental microbiome dissimilarity. Furthermore, sulfidic fish lineages showed a high degree of convergence among their gut microbiomes. These findings shed light on the factors shaping host-microbiome diversification in fishes, and they show how extreme selective

pressures can drive convergence in the host-associated microbiome composition of independent host lineages.

Conclusions

My research examined a variety of responses to selection, including changes in gene expression, trophic ecology, and host-microbiome associations, uncovering multiple ways in which extreme sources of selection act on populations. Furthermore, findings from my research have revealed cases of both shared and different responses of populations upon colonizing extreme environments. The strongest evidence for convergence was found in Chapters 3 and 4, in which I identified convergence in the host-associated microbiomes of extremophile fishes. Furthermore, while I did find some commonalities in gene expression and resource use in Chapters 1 and 2, there were also idiosyncrasies among populations investigated in these chapters. These findings highlight that whether or not populations respond in predictable ways to natural selection varies across traits.

There are several reasons why convergence may be more likely in some traits than others (Rosenblum et al. 2014). Patterns of convergence can be identified at different biological levels (by focusing on traits at molecular, cellular, physiological, or other levels), and convergence in traits is expected to be more common at higher levels (Rosenblum et al. 2014). In Chapter 1, for example, while there was little convergence at the gene expression level, the functional outcome of salinity tolerance is shared among fish lineages. Non-convergence, especially at molecular levels, can also be due to the underlying genetic divergence between lineages, which can create differences in how populations respond to selection (Kaeuffer et al. 2012). I compared distantly related lineages in Chapter 1, so the higher divergence between lineages likely plays a role in the little convergence in gene expression I found. Furthermore, while extreme environments provide

an opportunity to focus on clearly-defined sources of selection, there are other abiotic and biotic sources of selection that accompany the extreme stressors of interest and may be less clearly defined. These additional sources of selection contribute to variation between populations and are often a source of non-convergence (Kaeuffer et al. 2012; Rosenblum et al. 2014). In all of the extremophile systems I investigated, other sources of selection likely underlie the idiosyncrasies I found throughout all of my chapters. For example, differences in the microbial communities that fish encounter when crossing salinity barriers may affect the host's gene expression (Hughes et al. 2017), and differences in the way nutrients flow into cave habitats may change the frequency and quality of food in the diets of cavefish (Keene et al. 2015). Even in the sulfidic fish gut microbiomes I investigated, there were microbes unique to certain host lineages, and variation in the geography and chemistry between sulfide springs likely also plays a role in these idiosyncrasies, among other sources of selection (Hotaling et al. 2019).

While patterns of convergence can provide support for natural selection driving similar trait changes in populations, it is important to consider the multiple levels at which selection acts and the role of stochastic effects in observing convergence (Kaeuffer et al. 2012; Rosenblum et al. 2014). The convergence I found in sulfidic fish microbiomes provides an opportunity to explore how multiple factors shape host microbiome assembly, especially since selection acts on both the host and its associated microbes (Bang et al. 2018). For example, host-associated microbiomes of sulfidic fish may be similar to each other because sulfidic fish have evolved to recruit adaptive microbes, or this pattern could be because sulfide springs themselves are very distinct from nonsulfidic streams in their microbial communities, resulting in different microbes colonizing the host from the external environment. These findings of convergence identify ways in which natural selection drives similar responses to shared environmental pressures, providing

a foundation for exploring the evolutionary outcomes of such changes, such as investigating whether the convergent host-associated microbes in sulfide springs are adaptive for the host.

Overall, my dissertation shows how populations change in a variety of traits in response to extreme sources of selection, which sheds light on how organisms survive in extreme environments and how populations cross habitat boundaries. Additionally, by exploring the commonalities and differences in how populations respond to natural selection across multiple systems, my dissertation provides insight into the conditions under which natural selection can act in predictable ways, which has implications for understanding how biodiversity arises.

Chapter 1 - Gene expression signatures of salinity transitions in *Limia perugiae* (Poeciliidae), with comparisons to other teleosts

Elizabeth J. Wilson, Nick Barts, John Coffin, James B. Johnson, Carlos M. Rodríguez Peña, Joanna L. Kelley, Michael Tobler, and Ryan Greenway

Abstract

Salinity gradients act as strong environmental barriers that limit the distribution of aquatic organisms. Changes in gene expression associated with transitions between freshwater and saltwater environments can provide insight into organismal responses to variation in salinity. We used RNA-sequencing (RNA-seq) to investigate genome-wide variation in gene expression between a hypersaline population and a freshwater population of the livebearing fish species *Limia perugiae* (Poeciliidae). Our analyses of gill gene expression revealed potential molecular mechanisms underlying salinity tolerance in this species, including the enrichment of genes involved in ion transport, maintenance of chemical homeostasis, and cell signaling in the hypersaline population. We also found differences in gene expression patterns associated with cell cycle and protein folding processes between the hypersaline and freshwater *L. perugiae*. Bidirectional freshwater-saltwater transitions have occurred repeatedly during the diversification of fishes, allowing for broad-scale examination of repeatable patterns in evolution. We compared transcriptomic variation in *L. perugiae* with other teleosts that have made freshwater-saltwater transitions to test for convergence in gene expression. Among the four distantly related population pairs from high- and low-salinity environments that we included in our analysis, we found only ten shared differentially expressed genes, indicating little evidence for convergence. However, we found that differentially expressed genes shared among three or more lineages

were functionally enriched for ion transport and immune functioning. Overall, our results—in conjunction with other recent studies—suggest that different genes are involved in salinity transitions across disparate lineages of teleost fishes.

Introduction

Variation in salinity imposes osmoregulatory challenges on aquatic organisms, and contact zones between freshwater and saltwater environments act as barriers that limit the ability of animals to move from one habitat to the other (Davis et al. 2012). Many aquatic taxa have consequently failed to cross natural salinity gradients (Lee and Bell 1999). Those that achieve such habitat shifts overcome osmoregulatory challenges through plasticity or adaptation, and both responses have greatly shaped aquatic species distributions (Corush 2019; Lee et al. 2003; Whitehead et al. 2011). Due to the ecological expansion that accompanies colonization of novel habitats, invasions from freshwater to saltwater environments, or vice versa, that result in adaptation are of particular interest for elucidating the genetic mechanisms underlying salinity tolerance and the diversification of aquatic organisms (Davis et al. 2012; Lee and Bell 1999; Betancur-R 2010).

Among fishes, diversification has coincided with repeated transitions between freshwater and saltwater habitats (Lee and Bell 1999). Many saltwater-freshwater transitions in fishes have occurred over long evolutionary timescales, and deep evolutionary divergences have resulted in many species only tolerating a narrow range of salinities, restricting them to either freshwater or saltwater environments (Betancur-R 2010; Lee and Bell 1999; Bloom et al. 2013; Bloom and Lovejoy 2017; Carrete Vega and Wiens 2012). A considerably smaller portion of fish species can survive in both freshwater and saltwater environments (Betancur-R et al. 2015). Among species that tolerate a broad range of salinities, movement along salinity clines is characteristic

of diadromous lineages with life histories involving migration between freshwater and saltwater environments during an individual's lifetime (Betancur-R et al. 2015; Bloom and Lovejoy 2014). While some species cannot cross the saltwater-freshwater boundary and some do so in their lifetimes, there are also lineages between these two extremes that have made transitions between saltwater and freshwater environments at microevolutionary scales (Betancur-R 2010; Kusakabe et al. 2017; Xu et al. 2013). Few studies have investigated mechanisms of salinity adaptation in species where closely related populations experience different salinity regimes, and the evolutionary repeatability of these mechanisms across lineages remains to be explored.

Transitions between freshwater and saltwater environments are challenging for animals that actively maintain internal solute homeostasis. The many physiological processes involved in stable state osmotic and ionic balance necessitate changes in multiple interdependent processes when crossing a salinity barrier (Kültz 2015). To maintain homeostasis, fishes in freshwater must actively absorb salt and excrete water in the form of dilute urine to counteract their passive loss of salt and absorption of water (Kültz 2015). In contrast, fishes in saltwater environments must remove salt and retain water (Kültz 2015). As a result, crossing a salinity barrier requires a shift between absorption and excretion of ions and water in multiple organs, involving both active and passive processes (Kültz 2015; Greenwell et al. 2003). Remodeling of gill epithelia and regulation of ion transporters, aquaporins, and tight junctions in the gill are particularly central to this process (Foskett et al. 1983; Greenwell et al. 2003; Hwang 1987; Velotta et al. 2017). One way to quantify such complex physiological responses to variation in salinity is to compare patterns of gene expression across populations in different habitat types.

Several studies have investigated the physiological and transcriptomic responses to changes in salinity between populations of the same species or among closely related species of

fish (Gibbons et al. 2017; Hughes et al. 2017; Velotta et al. 2017; Xu et al. 2013). Differential gene expression associated with osmoregulation and ion transport is commonly found between populations inhabiting environments of different salinities (Hughes et al. 2017; Velotta et al. 2017; Xu et al. 2013). In addition, differential expression has also been documented in genes associated with other biological functions, including immune processes (Gibbons et al. 2017; Hughes et al. 2017), cell communication (Xu et al. 2013), stress tolerance (Xu et al. 2013), and gill membrane permeability (Gibbons et al. 2017). These studies highlight candidate pathways that play important roles in divergence between saltwater and freshwater ecotypes within species, but there have been few comparisons investigating the repeatability of gene expression responses to variation in salinity across phylogenetically disparate taxa. Investigating evidence of convergence in gene expression patterns across taxa that have undergone similar salinity transitions will provide insight into possible shared and unique pathways involved in adaptation to different salinity regimes.

In this study, we investigated patterns of gene expression in a freshwater and a hypersaline population of a livebearing fish species, *Limia perugiae* (Poeciliidae), and compared transcriptomic variation between these populations to those observed in other species that have undergone similar salinity transitions. Freshwater fishes in the genus *Limia* are endemic to the islands of the West Indies (Rauchenberger 1988; Weaver, Cruz, et al. 2016). *Limia perugiae* is a widespread species across the southern portion of Hispaniola, occurring in freshwater artesian springs and low-order creeks, as well as hypersaline inland lakes and coastal lagoons (Weaver, Tello, et al. 2016; Erbelding-Denk et al. 1994; Haney and Walsh 2003). Exposure to high salinities in *L. perugiae* has been shown to decrease metabolic rate (Haney and Walsh 2003), increase the production of Na^+/K^+-ATPase and oxidative phosphorylation proteins in the gills

(Weaver, Tello, et al. 2016), and reduce adult body size (Weaver, Tello, et al. 2016). Though predominantly associated with freshwater habitats, many fishes in the family Poeciliidae are able to tolerate a broad range of salinities, a factor potentially responsible for facilitating their dispersal across a wide geographic range (Myers 1949; Smith and Bermingham 2005; Rosen and Bailey 1963). While the mechanisms and consequences of salinity tolerance at the biochemical and physiological levels have been a focus of research in poeciliids (Chervinski 1984; Gonzalez et al. 2005; Tsai et al. 2018; Weaver, Tello, et al. 2016; Yang et al. 2011), the genetic underpinnings of salinity tolerance have yet to be investigated in this family. We used a natural system with conspecific populations occurring in both a freshwater and hypersaline habitat to characterize the potential molecular mechanisms underlying high salinity tolerance in poeciliid fishes.

We used RNA-sequencing (RNA-seq) to study genome-wide patterns of gene expression between freshwater and hypersaline *L. perugiae*. From this analysis, we aimed to identify genes and physiological pathways associated with salinity tolerance in this species. We then compared the *L. perugiae* population pair to other population pairs of freshwater and saltwater ecotypes in disparate taxa to understand if mechanisms of osmoregulatory capability are shared across divergent lineages of teleost fishes. We utilized a comparative transcriptomics approach that leverages new and pre-existing gene expression datasets to address the following questions: 1) What genes are differentially expressed between freshwater and saltwater *L. perugiae* populations, and with what physiological processes are they associated? 2) Is there evidence for commonalities in gene expression among phylogenetically disparate teleosts with freshwater and saltwater populations?

Methods

Sample collection

Limia perugiae were collected using a seine from a hypersaline lagoon (Laguna Oviedo: 17.801 °N, 71.363 °W) and a geographically proximate freshwater stream (Los Cocos: 17.905 °N, 71.286 °W) in the Dominican Republic. Following capture, adult females (*N*=6 per site) were euthanized, and both sets of gills were extracted using sterilized scissors and forceps. Tissues were immediately preserved in RNAlater (Ambion Inc., Austin, TX, USA).

RNA-seq library preparation

RNA extraction, library preparation, and sequencing of samples followed procedures previously employed for related poeciliid species (Kelley et al. 2012, 2016). Briefly, 10–30 mg of tissue from each individual was frozen in liquid nitrogen, pulverized, and total RNA was extracted using the NucleoSpin RNA kit (Machery-Nagel, Düren, Germany). mRNA isolation and cDNA library preparation were completed with 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), with minor modifications to the manufacturers' protocol (Kelley et al. 2012, 2016; Passow et al. 2017). cDNA libraries were individually barcoded, quantified with Qubit and an Agilent 2100 Bioanalyzer High Sensitivity DNA chip, and then pooled with cDNA samples from other projects in sets of 11–12 samples based on nanomolar concentrations. Samples were split across pools such that samples from each habitat type were not all sequenced together, and there was no evidence for lane effects. Libraries were sequenced on an Illumina HiSeq 2500 using paired-end 101-base-pair (bp) reads at the Washington State University Spokane Genomics Core.

Mapping

All raw reads were trimmed twice (quality 0 to remove Illumina adapters, followed by quality 24) using Trimgalore! (v.0.4.0; Krueger 2014). Trimmed reads were mapped to the *Poecilia mexicana* reference genome (RefSeq accession number: GCF_001443325.1; Warren et al. 2018) with an appended mitochondrial genome (GenBank Accession Number: KC992998.1) using BWA-MEM v.0.7.12 (Li and Durbin 2009). We annotated genes from the *P. mexicana* reference genome by extracting the longest transcript for each gene (with the perl script gff2fast.pl: https://github.com/ISUgenomics/common_scripts/blob/master/gff2fasta.pl) and comparing them against entries in the human SWISS-PROT database (critical E-value 0.001; access date 04/15/2017) using BLASTx (Camacho et al. 2009). Each gene was annotated with the best BLAST hit from the human database based on the top high-scoring segment pair.

Differential gene expression

We used STRINGTIE (v.1.3.3b; Pertea et al. 2015, 2016) to quantify the number of reads mapped to each gene for each individual (measured in counts per million mapped reads) and then used the prepDE.py script (provided with STRINGTIE) to generate a read count matrix (Pertea et al. 2016). We removed genes that did not have at least two counts per million in three or more individuals across both populations, resulting in 18,659 genes that were included in differential gene expression analysis. We identified differentially expressed genes using generalized linear models (GLMs) in R, as implemented in the Bioconductor package edgeR (Robinson et al. 2010). We fit a negative binomial GLM to the normalized read counts of each gene based on tagwise dispersion estimates and a design matrix describing the comparison between the

saltwater and freshwater population using glmFit. We assessed statistical significance using the GLM likelihood-ratio test with a false discovery rate (FDR) of q-value < 0.05, calculated with the Benjamini-Hochberg correction (Benjamini and Hochberg 1995). After identifying the set of differentially expressed genes between the saltwater and freshwater population, we used a Gene Ontology (GO) enrichment analysis to explore putative biological functions of these genes. We annotated all differentially expressed genes that had a match in the human SWISS-PROT database with GO IDs (Harris et al. 2004) and tested for the enrichment of specific GO IDs separately in up and downregulated genes relative to the full background set of 18,659 genes using GOrilla (FDR q-value < 0.05, accessed May 26, 2022; Eden et al. 2009). A total of 10,935 genes in the background set were associated with a GO term in the database.

Weighted gene co-expression network analysis

We constructed weighted gene co-expression networks to identify clusters of genes that were co-expressed across our samples (Zhang and Horvath 2005). To prepare the gene expression data for this analysis, we applied a variance-stabilizing transformation to the filtered reads using the varianceStabilizingTransformation function from the DESeq2 package (v.1.36.0; Love et al. 2014) in R, which allowed us to normalize the read counts relative to library size and appropriately scale the data for clustering (Langfelder and Horvath 2008, 2012). After transforming and normalizing the read counts, we used the cpm function in the edgeR package (v.3.38.1; Robinson et al. 2010) to generate a gene matrix of scaled read counts ($\log_2$-cpm, counts per million mapped reads) from the transformed read counts (Coffin et al. 2022).

As outlined in the WGCNA package documentation (Langfelder and Horvath 2008), we used hierarchical clustering to cluster the samples based on their gene expression profiles. We

used the hclust function from the flashClust package (v.1.1.2; Langfelder and Horvath 2012) to cluster the samples. We then created a weighted network adjacency matrix using the adjacency function from the WGCNA package (v.1.71; Langfelder and Horvath 2008, 2012). The adjacency matrix was constructed by calculating pairwise co-expression similarities (Pearson correlation coefficients) and raising them to a power of β, a soft thresholding power (Langfelder and Horvath 2008). We used the pickSoftThreshold function in the WGCNA package (Langfelder and Horvath 2008) to assist in selecting a value of β that ensured our network fit the approximate scale free topology criterion while retaining the highest possible mean connectivity between the network genes. Based on the scale free topology model fit and mean connectivity of our network, we selected $\beta = 7$.

From our correlation network, we then generated a topological overlap dissimilarity matrix to identify modules of co-expressed genes. We calculated dissimilarity between the genes by converting the adjacencies into topological overlap similarities using the TOMsimilarity function in the WGCNA package (Langfelder and Horvath 2008) and then subtracting these topological overlap measures from 1. To identify modules of co-expressed genes, we used the hclust function from the flashClust package (Langfelder and Horvath 2012) for hierarchical clustering of the genes and then created a hierarchical clustering dendrogram. We used the cutreeDynamic function from the dynamicTreeCut package (Langfelder and Horvath 2008; Langfelder et al. 2008) to extract the modules from the dendrogram. To summarize the gene expression variation in each module, the first principal component of each module in the expression matrix (the eigengene) was calculated using the moduleEigengenes function from the WGCNA package (Langfelder and Horvath 2008). We used the function mergeCloseModules in the WGCNA package to merge eigengenes that were highly correlated. We included eigengenes

with a correlation greater than 0.9 for merging. Finally, we identified modules that were significantly associated with the presence or absence of salinity by calculating correlation coefficients between the eigengenes and the habitat type. *P*-values of the correlation coefficients were calculated using the corPvalueStudent function from the WGCNA package, and we retained modules with *P*-values less than 0.01 for functional enrichment analyses. Similar to our differential gene expression analysis, we used GOrilla (Eden et al. 2009) for Gene Ontology enrichment analysis of genes contained in modules exhibiting significant correlations with habitat type.

Comparisons of L. perugiae with other species

We mined previously published datasets to identify gene expression patterns commonly associated with salinity tolerance in disparate taxa, including South American silversides (*Odontesthes bonariensis* and *Odontesthes argentinensis*) from Hughes et al. (2017), three-spine sticklebacks (*Gasterosteus aculeatus*) from Gibbons et al. (2017), Amur ide (*Leuciscus waleckii*) from Xu et al. (2013), and *L. perugiae* (this study; Table 1-1). Each of these experiments generated paired-end RNA-seq raw reads from gill tissue in two ecotypes (one freshwater and the other saltwater) of the same lineage.

Raw RNA-seq reads from each transcriptomics project were downloaded in FASTQ format, and reference genomes or transcriptomes for each species were downloaded in FASTA format from Genbank (see Table 1-1 for accession numbers). Reads were trimmed and mapped to their respective reference genomes (*Cyprinus carpio*: GCF_000951615.1, Xu et al. 2014; *Poecilia mexicana*: GCF_001443325.1, Warren et al. 2018; *Gasterosteus aculeatus*: Broad S1 v. 93, Jones et al. 2012) or reference transcriptome (*Menidia menidia*: GEVY00000000.1,

Therkildsen and Palumbi 2017) following the same methods described above for *L. perugiae*. We then quantified the number of reads mapped to each gene in the annotation file for each reference genome and created a read counts matrix for each species, which were used for subsequent expression analyses. Expression analyses were performed in R version 4.1.2. The 10,000 genes with the highest standard deviation between freshwater and saltwater samples were abstracted from each read counts matrix, and overall expression patterns were visualized with multi-dimensional scaling (MDS) plots.

To make comparisons across species, we used OrthoFinder v2.2.6 to identify orthologous genes among the reference genomes (Emms and Kelly 2015, 2019), and 18,419 orthogroups were identified. To calculate counts per orthogroup, we used the gene counts matrix of each species to sum up the counts across all loci contained in an orthogroup. Based on this orthogroup counts matrix, we retained only the orthogroups that were expressed in all individuals (cpm > 0 per individual), resulting in 12,743 retained orthogroups.

To evaluate expression differences for each orthogroup, we made pairwise comparisons between ecotypes following the same methods described above for the *L. perugiae* comparisons. Briefly, we normalized reads, created and compared generalized linear models of the normalized read counts, generated a design matrix, estimated tagwise dispersion, and conducted GLM likelihood-ratio tests to test whether differences in expression were statistically different between the freshwater and saltwater population for each orthgroup. To identify orthogroups exhibiting convergent expression patterns across lineages, we intersected the significantly upregulated and downregulated orthogroups from all lineage-specific comparisons, identifying orthogroups that were differentially expressed in the same direction in pairwise, three-, and four-way comparisons among the lineages. After identifying the set of orthogroups with differential expression across

three or more lineages, we used a GO enrichment analysis as described above to explore the putative biological functions of these candidate gene sets.

Results

Comparative analysis of freshwater and hypersaline L. perugiae

We used RNA-seq to characterize the transcriptomes of *L. perugiae* from a freshwater (n=6) and a hypersaline populations (Table 1-1). 73,590,325 total raw reads were obtained across all individuals: 34,998,157 from freshwater *L. perugiae* (n=6) and 38,592,168 from saltwater *L. perugiae* (n=6) before trimming (Appendix Table A-1). After trimming, 95.5% of reads from the freshwater individuals mapped to the *Poecilia mexicana* reference genome, and 95.2% mapped for the saltwater individuals (Appendix Table A-1).

We identified 4,895 differentially expressed genes between saltwater and freshwater ecotypes of *L. perugiae*, 2,437 of which were upregulated and 2,458 of which were downregulated in the saltwater ecotype (Figure 1-1A). The genes upregulated in the saltwater population were largely associated with ion transport, maintaining chemical homeostasis, and cell signaling (FDR < 0.05) (Table 1-2). Processes relevant to chemical homeostasis included several solute carrier genes, such as *SLC9A3*, *SLC8B1*, *SLC12A8*, and *SLC30A9*. Na^+/K^+-ATPase and other ATPase genes, such as *ATP1B1* and *ATP6V1A*, were also upregulated among the genes involved in chemical homeostasis and signal transduction. Genes downregulated in the saltwater population corresponded to GO process terms such as mitotic cell cycle process, protein folding, chromosome segregation, rRNA processing, and mitochondrial translational elongation (FDR < 0.05; Table 1-2).

WGCNA revealed five modules of co-expressed genes that were significantly correlated with salinity (Figure 1-2). The turquoise module was positively correlated with salinity (*P*-value < 0.01), and the black, red, royalblue, and blue modules were all negatively correlated with salinity (Figure 1-2). Out of the 18,659 genes included in our analysis, the positively correlated module (turquoise module) contained 4,056 genes. The negatively correlated modules contained 2,166 genes (black module), 837 genes (red module), 173 genes (royalblue module), and 3,300 genes (blue module). The correlation coefficients and their associated *P*-values between each gene and the environmental condition (salinity), and between each gene and each module, are included in Appendix Table A-2. Each gene's module assignment can also be found in Appendix Table A-2. From the functional enrichment analysis, we found that the turquoise, royalblue, and blue modules were significantly enriched for biological processes, and these modules of co-expressed genes largely corroborated the differential expression results. Like the biological processes that were enriched among the up-regulated genes in the saltwater population, the module that was positively correlated with salinity (turquoise) was functionally enriched for GO terms involved in ion transport and cell signaling (Appendix Table A-3). There were, however, GO terms associated with the turquoise module that were not identified from the differential expression analysis, including regulation of autophagy and lipid transport (Appendix Table A-3). The modules that were negatively correlated with salinity also reflected biological processes that were associated with down-regulated genes in the saltwater *L. perugiae*, including regulation of the cell cycle and protein folding (Appendix Table A-4 and Appendix Table A-5). The royalblue module only had one significantly enriched GO term, chemokine-mediated signaling pathway, which was a GO term also associated with down-regulation of genes in the saltwater environment (Appendix Table A-4). In contrast, the blue module contained several significant

GO terms, including mitotic cell cycle process, protein folding, rRNA metabolic process, and chromosome organization (Appendix Table A-5). There were many GO terms associated with the blue module that were not represented in the differential expression analysis, including sarcomere organization, macromolecule biosynthetic process, regulation of cellular response to heat, and nucleic acid metabolic process (Appendix Table A-5). Several of the GO terms unique to the blue module were related to cellular metabolism and biosynthesis pathways (Appendix Table A-5).

Comparisons of Limia *with phylogenetically disparate teleosts*

We compared transcriptomic differences between *L. perugiae* populations to previously published transcriptome data from teleosts with populations from freshwater and saline habitats (Table 1-1). Mapping statistics for all four population pairs can be found in Table A-1. As expected, MDS plots indicated that orthogroup expression variation was primarily driven by differences among taxonomic groups, with much smaller differences between ecotypes within species (Figure 1-3A). We then compared orthogroup expression profiles of all the freshwater and saltwater lineages based on mean expression values and found that the variation in orthogroup expression largely reflects phylogenetic divergence among lineages (Figure 1-3B). Closely related lineages exhibited more similar expression profiles, irrespective of environmental conditions (saltwater vs. freshwater; Figure 1-3B). Across all groups, 122 differentially expressed orthogroups were shared across at least three lineages, but only 10 shared orthogroups were differentially expressed across all freshwater and saltwater population pairs (Figure 1-3C). Of those 10 orthogroups, 9 had annotations in the SWISS-PROT database (Table 1-3; Figure 1-4), including a Na^+/H^+-exchanger (*SLC9A3*) involved in osmoregulation. The directionality of

differential expression varied among lineages, and none of the 10 shared differentially expressed orthogroups were up-regulated or down-regulated among all four population pairs (Figure 1-4).

To investigate the biological processes reflected in shared differentially expressed orthogroups, we analyzed the differentially expressed orthogroups that were shared among three or more lineages (Figure 1-3C). GO analysis indicated enrichment in biological processes associated mostly with transmembrane transport (particularly ion transport) and some associated with immune function, which were all significant based on *P*-value (*P*-value < 0.001) but not after FDR correction (Table 1-4). Some of the GO processes specific to transmembrane transport included anion transport, inorganic cation import across plasma membrane, potassium ion import, urea transmembrane transport, and pyruvate transmembrane transport. The shared differentially expressed immune genes were reflective of negative regulation of T-helper cell differentiation, negative regulation of leukocyte differentiation, negative regulation of CD4-positive, alpha-beta T cell differentiation, and regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily.

Discussion

To better understand mechanisms of salinity tolerance in the livebearing fish *Limia perugiae*, we compared patterns of gene expression between a freshwater and a hypersaline population. GO enrichment analysis of differentially expressed genes revealed that the genes upregulated in the saltwater population are largely related to ion transport and maintaining chemical homeostasis, while downregulated genes were associated with processes involved in the cell cycle regulation and protein folding. These results provide insight into how *L. perugiae* has colonized a novel environment and maintains homeostasis under extreme salinity stress. Comparisons of our *Limia*

results with pre-existing gene expression data collected from freshwater and saltwater ecotypes of South American silversides (*Odontesthes* spp.: Hughes et al. 2017), three-spine stickleback (*Gasterosteus aculeatus*: Gibbons et al. 2017), and Amur ide (*Leuciscus waleckii*: Xu et al. 2013) indicated that there were few shared differentially expressed genes among all four ecotype pairs. Variation in gene expression was largely shaped by phylogeny rather than environment, and the shared differentially expressed genes among all four pairs showed strong variation in the direction and magnitude of differential expression across lineages. Shared differentially expressed genes were primarily associated with ion transport and immune function. Overall, these results suggest that disparate lineages utilize different mechanisms for overcoming salinity challenges—at least at the level of gene expression. We found that various patterns of gene expression can emerge from crossing saltwater-freshwater boundaries, providing evidence of diverse responses of teleost lineages to a similar environmental challenge. A major question remaining is to what degree variation in gene expression in *L. perugiae* and the other lineages here shaped by phenotypic plasticity and by genetic differences in gene regulation. A previous study in stickleback found evidence for heritable gene expression differences between freshwater and saltwater ecotypes, evidence for shared plastic responses between ecotypes, but only little evidence for ecotype-specific plasticity (Gibbons et al. 2017). Similar studies that combine field studies with laboratory experimentation are highly warranted for other study systems.

Responses to variation in salinity in a freshwater and hypersaline population of L. perugiae

Regulation of transmembrane transport and gill epithelial permeability

Overcoming the physiological challenges associated with transitioning between saltwater and freshwater environments requires modification of ion transport across the gill epithelia (Foskett

et al. 1983). In our analysis of differentially expressed genes between a freshwater and hypersaline population of *L. perugiae*, we found evidence for differential regulation of ion transport. Among the genes that were upregulated in the hypersaline population, GO enrichment analysis revealed that several terms were associated with anion transport, sodium ion transport, metal ion transport, and maintaining chemical homeostasis. Genes corresponding to solute carrier families (e.g., *SLC9A3*, *SLC8B1*, *SLC12A8*, *SLC30A9*, *SLC7A1*) and ATPases (e.g., *ATP1B1*, *ATP6V1A*, *ATP13A3*) were among the upregulated genes that are important in maintaining ion and chemical homeostasis (Li et al. 2020). Increasing the rate of inorganic ion, amino acid, and nucleotide transport via upregulation of solute carriers allows aquatic organisms to maintain osmotic balance in saline environments, potentially facilitating acclimation or adaptation to hypersalinity (Li et al. 2020). Specifically, ATPases—such as Na^+/K^+-ATPase—have been well-studied for their role in salinity tolerance during both acclimation and adaptation (Deane and Woo 2004; Gonzalez et al. 2005; Langdon and Thorpe 1984; Laverty and Skadhauge 2012; Scott et al. 2004; Tipsmark et al. 2002; Weaver, Tello, et al. 2016). Consistent with our study, previous Western blot analyses of freshwater and hypersaline *L. perugiae* populations have revealed an increase in Na^+/K^+-ATPase expression in hypersaline *L. perugiae* when compared to their freshwater conspecifics, which is essential for excreting Na^+ and Cl^- out of the body to maintain homeostasis (Weaver, Tello, et al. 2016).

At very high salinity levels, it is especially difficult to pump Na^+ against the chemical gradient and out of the gill (Gonzalez 2012). In a hypersaline environment, it therefore may be beneficial to maintain a lower epithelial salt permeability to avoid a back-flux of salts across the tight junctions of the gill epithelia (Gonzalez 2012; Lam et al. 2014). We found up-regulation of genes involved in cell-cell junction organization, including claudins involved in the formation of

tight junctions (*CLDN3*, *CLDN4*, *CLDN5*, *CLDN8*, and *CLDN10*). Expression of gill claudin genes has been associated with salinity acclimation in fish (Lam et al. 2014; Tipsmark et al. 2002), and up-regulation of claudin-8 (*CLDN8*) has been shown to reduce the paracellular barrier permeability to Na^+ (Amasheh et al. 2009), suggesting a role for regulation of gill epithelial permeability via tight junctions during osmoregulatory challenges. Our findings support the hypothesis that modifications to transmembrane transport and gill epithelial permeability jointly contribute to the salt exclusion necessary for survival in a hypersaline environment.

Cell cycle regulation

Populations living in hypersaline environments must cope with the adverse effects of high salinity levels, often resulting in strategies for damage repair and energy reallocation (Laverty and Skadhauge 2012; Weaver, Tello, et al. 2016). Functional analysis of the downregulated genes in the hypersaline population of *L. perugiae* showed signatures of downregulation of cell cycle and protein folding processes. Downregulation of genes involved in the cell cycle is expected to occur under high salinity stress to stop the replication of damaged cells and allow time for DNA repair (Evans and Kültz 2020; Kaufmann and Paules 1996). Similarly, locally adapted killifish living in freshwater and brackish environments exhibit evidence for divergence in the expression of genes underlying control of the cell cycle, supporting a central role in cell cycle regulation in adaptation to salinity stress (Brennan et al. 2015; Evans and Kültz 2020). However, our findings may also suggest a general re-allocation of energy resources (Evans and Kültz 2020). In hypersaline *L. perugiae*, the downregulation of genes involved in the cell cycle and protein folding could be reflective of an energetic trade-off between increased demand for processes involved in osmoregulation and the energetic cost of growth under stressful conditions

(Weaver, Tello, et al. 2016). In support of such energetic trade-offs, hypersaline *L. perugiae* have morphological differences from their freshwater counterparts, including a significantly smaller body size and reduced secondary sex features in males (Weaver, Tello, et al. 2016; Haney and Walsh 2003).

Regulation of cell signaling

To respond to osmoregulatory challenges, fishes must perceive their environmental salinity and maintain intracellular signals that modulate ion transport and other processes (Evans 2002; Kültz 2012). We found differentially expressed genes associated with a variety of cell signaling processes. Several upregulated genes in the hypersaline population were enriched in GO terms associated with cell signaling, including signal transduction, regulation of signaling, G-protein-coupled receptor signaling pathway, and second-messenger-mediated signaling. G-protein-coupled receptor signaling is among the pathways involved in allowing aquatic organisms to sense their environmental salinity (Papakostas et al. 2012; Sun et al. 2020). Some of these signaling pathway components may also play a role in regulating ion transport, such as mitogen-activated protein kinases (MAPKs) and other serine/threonine protein kinases (Evans 2010; Herzig and Neumann 2000; Kozak et al. 2014; Kültz and Avila 2001), which were among the upregulated genes. MAPK signaling pathways are also implicated in regulation of the cell cycle, some of which trigger cellular growth arrest and DNA damage repair in response to osmotic stress (Kültz and Avila 2001; Kültz and Burg 1998). The upregulation of signaling pathways, especially MAPK genes, may consequently play a role in the downregulation of cell-cycle genes that we found in the hypersaline *L. perugiae* population.

Comparisons among disparate lineages

Lineage-specific responses to variation in salinity

Convergence in gene expression among independently evolved populations can occur in response to shared environmental stressors (Fisher and Oleksiak 2007; Greenway et al. 2020). Among fishes, cases of convergence in gene expression have been identified in response to selective pressures such as pollution (Fisher and Oleksiak 2007), hydrogen sulfide (Greenway et al. 2020), and the absence of light in caves (Stahl and Gross 2017). In contrast, our comparative transcriptomics analysis indicated there is little evidence for convergence in gene expression patterns among fish lineages in response to variation in salinity. Even among the shared differentially expressed genes, there is considerable variation in the magnitude and direction of expression differences between ecotypes across lineages. This finding mirrors genomic analyses that looked for convergent signatures associated with adaptation to salinity variation, which found osmoregulation genes to be common targets of selection but also variation in the exact genes that were involved across different lineages (Velotta et al. 2022).

There are several non-mutually exclusive hypotheses for why we might not expect convergence in gene expression in response to variation in salinity (Kaeuffer et al. 2012). First, changes in both protein-coding DNA sequences and gene expression may occur during adaptation to a novel environment, or they may occur independently (Brown et al. 2019; Jones et al. 2012; Rosenblum et al. 2004). Under certain conditions, selection may favor changes in protein structure and function via modifications at the sequence level, regardless of whether or not gene expression is affected (Brown et al. 2019). Alternatively, changes in gene expression may be favored without selection acting at the sequence level (Brown et al. 2019). Variation in how selection acts on gene expression could contribute to the lack of convergence among the

lineages in our study, and there also may be stronger signals of convergence at levels other than gene expression.

Secondly, differences in genetic architecture among lineages may cause different responses to selection, leading to diverse evolutionary outcomes (Kaeuffer et al. 2012). Convergence is therefore less likely to occur with increasing genetic divergence between lineages (Conte et al. 2012), so it is not necessarily surprising that we did not find convergence among distantly related fishes. Even if the genetic architectures are similar among lineages, idiosyncratic responses to selection can also arise as a consequence of functional redundancy (Láruson et al. 2020). Specifically, modification of different genes and pathways may have equivalent functional and fitness consequences (Kozak et al. 2014; Velotta et al. 2022). Functional redundancy is particularly common in complex traits like salinity tolerance that involve many genes and physiological pathways (Láruson et al. 2020). While salinity tolerance is a shared outcome among the lineages in our analysis, the molecular mechanisms underlying osmoregulation may be unique to each lineage.

Third, the directionality of habitat transition may impact gene expression responses, especially if there are genetic adaptations to the habitat of origin. For example, a high-activity version of an ion transporter may be downregulated during transitions to freshwater, while a low-activity version of the same enzyme may be upregulated in the opposite direction. Among the four lineages we included in our study, two of them transitioned from saltwater to freshwater environments (*Gasterosteus* and *Odontesthes*), and two transitioned from freshwater to saltwater environments (*Limia* and *Leuciscus*). If the direction of transition elicits similar responses, then we would expect populations that transitioned in the same direction to share more differentially expressed genes than those that did not transition in the same direction. However, we did not find

this to be the case, as the pairs transitioning in the same direction share fewer differentially expressed genes than those that made opposite transitions (Figure 1-3C).

Finally, covariation with other sources of selection—both abiotic and biotic—may cause idiosyncratic gene expression patterns across lineages. Beyond the challenges directly imposed by variation in salinity, such habitat transitions are often accompanied by other environmental challenges, such as variation in temperature, exposure to novel parasites, and restructuring of host-associated microbial communities (Hughes et al. 2017; Lee and Bell 1999; Lokesh and Kiron 2016; Schmidt et al. 2015). For example, hypersaline Amur ide must cope with high alkalinity stress in addition to salinity, and hypersaline *L. perugiae* have a warmer environment than freshwater *L. perugiae* (Weaver, Tello, et al. 2016; Xu et al. 2013). Additionally, there are differences in the gill microbial communities of saline and freshwater populations of South American silversides (Hughes et al. 2017). Such correlated environmental factors can contribute variation in gene expression responses to salinity transitions we observed.

Shared responses involving ion transport and immune function

Although there is little evidence for convergence among all four lineage pairs included in our analysis, there was some functional overlap in the differentially expressed genes shared among three or more lineages. Most of the shared differentially expressed genes were associated with ion transport, as well as some immune system processes. As previously discussed, regulation of ion transport is expected to be crucial when crossing a saltwater-freshwater boundary (Foskett et al. 1983) and osmoregulation genes also exhibit evidence of convergent evolution during salinity transitions (Velotta et al. 2022). Variation in expression of immune genes, particularly those related to inflammation and adaptive immunity, has also been documented in fish during salinity

acclimation (El-Leithy et al. 2019; Jeffries et al. 2019). Under a variety of selection pressures, locally adapted populations also frequently show divergence in immune genes due to other factors such as life history differences, different parasite exposure, and shifts in the microbiome (Eizaguirre et al. 2012; Hughes et al. 2017; Jeffries et al. 2019; Zhang et al. 2015; Miller et al. 2001). Immune loci are consequently evolutionary hotspots in diversification (Holmes 2004; Hughes 2002; Sommer 2005). It needs to be tested whether changes in the expression of immune-related genes are directly linked to variation in salinity or whether these genes generally respond to changes in correlated biotic sources of selection.

Overall, our analysis of gene expression patterns between locally adapted freshwater and hypersaline populations of *L. perugiae* revealed upregulation of genes involved in ion regulation, maintaining chemical homeostasis, and cell signaling in the hypersaline population, as well as downregulation of genes involved in the cell cycle and protein folding. These results provide insight into how this livebearing fish maintains homeostasis in a hypersaline environment. Several other lineages of teleosts have similarly made transitions between freshwater and saline environments, providing an opportunity to test the predictability of responses to variation in salinity. We therefore investigated whether disparate lineages show evidence of convergence in gene expression in response to saltwater-freshwater transitions. Our comparisons between four population pairs of freshwater and saline ecotypes in disparate teleost lineages showed little evidence for convergence, as there were only ten differentially expressed genes that were shared among them all. Despite this, we found that the differentially expressed genes shared in three or more of the lineages reflected biological processes related to ion transport and immune functioning.

Tables

Table 1-1. Species included in the analysis, including environment (freshwater [FW] or saltwater [SW]), collection location, sample size (N), NCBI Sequence Read Archive (SRA) accession numbers, and study reference.

Species	Environment	Collection Location	N	SRA Accessions	Study Reference
Limia perugiae	FW	Los Cocos (17.905°N, 71.286°W), Dominican Republic	6	*Pending*	This study
Limia perugiae	SW	Laguna Oviedo (17.801°N, 71.363°W), Dominican Republic	6	*Pending*	This study
Gasterosteus aculeatus	FW	Trout Lake (49°30029″N, 123°52029″W), British Columbia, Canada	5	SRX2544970, SRX2544969, SRX2544968, SRX2544967, SRX2544966	(Gibbons et al. 2017)
Gasterosteus aculeatus	SW	Oyster Lagoon (49°36043.53″N, 124°01052.12″W), British Columbia, Canada	5	SRX2544985, SRX2544984, SRX2544983, SRX2544982, SRX2544981	(Gibbons et al. 2017)
Odontesthes bonariensis	FW	Lake Chascomús (35°34′S, 58°01′W), Argentina	3	SRX1681471, SRX1681473, SRX1681474	(Hughes et al. 2017)
Odontesthes argentinensis	SW	Mar del Plata (38°02′S, 57°31′W), Argentina	3	SRX1671790, SRX1681012, SRX1681017	(Hughes et al. 2017)
Leuciscus waleckii	FW	Ganggeng Nor Lake (43°17'48"N, 116°53'27"E), Mongolia	1	SRX333071	(Xu et al. 2013)
Leuciscus waleckii	SW	Dali Nor Lake (43°22′43"N, 116°39'24"E), Mongolia	1	SRX1410650	(Xu et al. 2013)

Table 1-2. GO process terms with significant enrichment in genes upregulated and downregulated in the hypersaline ecotype of *Limia perugiae* (FDR < 0.05). The table includes the GO term ID, description, the degree of enrichment, the number of differentially expressed genes associated with the GO term (*N*), as well as *P* and FDR-corrected *q*-values.

GO term	Description	Enrichment	*N*	*P*-value	FDR *q*-value
Upregulated					
GO:0006820	anion transport	1.8	100	5.67E-10	8.04E-06
GO:0006811	ion transport	1.49	181	3.41E-09	2.42E-05
GO:0007165	signal transduction	1.22	540	4.67E-09	2.21E-05
GO:0051049	regulation of transport	1.36	242	4.50E-08	1.59E-04
GO:0034220	ion transmembrane transport	1.51	129	3.41E-07	9.67E-04
GO:0043269	regulation of ion transport	1.61	98	3.89E-07	9.19E-04
GO:0010959	regulation of metal ion transport	1.85	62	4.53E-07	9.18E-04
GO:0098656	anion transmembrane transport	1.89	57	5.26E-07	9.33E-04
GO:0032879	regulation of localization	1.24	360	1.19E-06	1.87E-03
GO:0015711	organic anion transport	1.71	73	1.21E-06	1.71E-03
GO:0002028	regulation of sodium ion transport	2.72	23	1.62E-06	2.09E-03
GO:0023051	regulation of signaling	1.19	451	3.07E-06	3.62E-03
GO:0048878	chemical homeostasis	1.4	153	3.61E-06	3.94E-03
GO:0010646	regulation of cell communication	1.19	444	7.08E-06	7.17E-03
GO:0003254	regulation of membrane depolarization	3.24	15	7.84E-06	7.41E-03
GO:0051050	positive regulation of transport	1.41	132	1.38E-05	1.23E-02
GO:1902305	regulation of sodium ion transmembrane transport	2.85	17	1.76E-05	1.46E-02
GO:0050794	regulation of cellular process	1.07	1144	2.50E-05	1.97E-02
GO:0031344	regulation of cell projection organization	1.43	111	3.54E-05	2.64E-02
GO:0043270	positive regulation of ion transport	1.78	46	3.61E-05	2.56E-02
GO:0007186	G protein-coupled receptor signaling pathway	1.47	97	3.72E-05	2.51E-02
GO:0010975	regulation of neuron projection development	1.51	84	3.89E-05	2.51E-02
GO:0045664	regulation of neuron differentiation	1.44	105	4.04E-05	2.49E-02
GO:0042908	xenobiotic transport	3.3	12	5.15E-05	3.04E-02
GO:0019932	second-messenger-mediated signaling	1.69	53	5.17E-05	2.93E-02
GO:0042592	homeostatic process	1.29	197	5.18E-05	2.83E-02
GO:0120035	regulation of plasma membrane bounded cell projection organization	1.42	109	5.79E-05	3.04E-02
GO:0006814	sodium ion transport	2.04	29	7.33E-05	3.71E-02
GO:0032501	multicellular organismal process	1.19	360	7.62E-05	3.72E-02
GO:0045216	cell-cell junction organization	1.94	33	7.65E-05	3.62E-02
GO:0050767	regulation of neurogenesis	1.38	122	7.78E-05	3.56E-02
GO:0055085	transmembrane transport	1.3	165	9.84E-05	4.36E-02
GO:0030001	metal ion transport	1.47	85	1.05E-04	4.52E-02
GO:0031345	negative regulation of cell projection organization	1.83	37	1.07E-04	4.46E-02
GO:0065008	regulation of biological quality	1.15	453	1.21E-04	4.91E-02
GO:0015849	organic acid transport	1.69	47	1.25E-04	4.91E-02
GO:0046942	carboxylic acid transport	1.69	47	1.25E-04	4.77E-02
GO:0065007	biological regulation	1.06	1274	1.27E-04	4.75E-02
Downregulated					
GO:1903047	mitotic cell cycle process	1.66	149	2.52E-11	3.58E-07
GO:0006457	protein folding	2.14	64	3.03E-10	2.15E-06
GO:0022402	cell cycle process	1.46	191	6.63E-09	3.13E-05

GO:0051301	cell division	1.71	98	1.40E-08	4.98E-05
GO:0006364	rRNA processing	1.91	65	4.51E-08	1.28E-04
GO:0009987	cellular process	1.05	1687	6.13E-08	1.45E-04
GO:0016072	rRNA metabolic process	1.82	69	1.35E-07	2.73E-04
GO:0030198	extracellular matrix organization	1.75	76	2.13E-07	3.77E-04
GO:0007059	chromosome segregation	2.38	33	3.54E-07	5.57E-04
GO:0043062	extracellular structure organization	1.7	80	3.85E-07	5.46E-04
GO:0000070	mitotic sister chromatid segregation	3.61	15	7.56E-07	9.74E-04
GO:0000819	sister chromatid segregation	3.41	16	1.01E-06	1.19E-03
GO:0007051	spindle organization	2.04	42	1.40E-06	1.53E-03
GO:0043933	protein-containing complex subunit organization	1.28	269	2.07E-06	2.10E-03
GO:0098813	nuclear chromosome segregation	3.14	17	2.34E-06	2.21E-03
GO:0006415	translational termination	2.16	35	2.50E-06	2.22E-03
GO:0043624	cellular protein complex disassembly	1.99	42	3.07E-06	2.56E-03
GO:0070125	mitochondrial translational elongation	2.18	33	3.71E-06	2.92E-03
GO:0070098	chemokine-mediated signaling pathway	3.17	16	4.01E-06	2.99E-03
GO:0018208	peptidyl-proline modification	2.65	21	6.13E-06	4.34E-03
GO:0007010	cytoskeleton organization	1.38	153	8.67E-06	5.85E-03
GO:0007088	regulation of mitotic nuclear division	1.89	42	1.29E-05	8.34E-03
GO:0044772	mitotic cell cycle phase transition	1.66	62	1.72E-05	1.06E-02
GO:0006336	DNA replication-independent nucleosome assembly	3.1	14	2.24E-05	1.32E-02
GO:0034622	cellular protein-containing complex assembly	1.36	150	2.47E-05	1.40E-02
GO:0051783	regulation of nuclear division	1.82	44	2.48E-05	1.35E-02
GO:0051383	kinetochore organization	4.16	9	2.59E-05	1.36E-02
GO:0044770	cell cycle phase transition	1.63	63	2.67E-05	1.35E-02
GO:0000278	mitotic cell cycle	1.94	36	2.98E-05	1.46E-02
GO:0000413	protein peptidyl-prolyl isomerization	2.77	16	3.82E-05	1.80E-02
GO:0034724	DNA replication-independent nucleosome organization	2.98	14	4.05E-05	1.85E-02
GO:0070126	mitochondrial translational termination	2.02	31	4.10E-05	1.82E-02
GO:0051276	chromosome organization	1.53	75	5.71E-05	2.45E-02
GO:0071840	cellular component organization or biogenesis	1.11	739	6.92E-05	2.88E-02
GO:0006414	translational elongation	1.85	37	7.46E-05	3.02E-02
GO:0022610	biological adhesion	1.35	134	8.31E-05	3.27E-02
GO:0007155	cell adhesion	1.35	133	8.91E-05	3.41E-02
GO:0042493	response to drug	1.37	119	9.53E-05	3.56E-02
GO:0016074	snoRNA metabolic process	3.46	10	9.90E-05	3.60E-02
GO:0000281	mitotic cytokinesis	2.43	18	1.11E-04	3.95E-02
GO:0000086	G2/M transition of mitotic cell cycle	1.83	36	1.15E-04	3.97E-02
GO:1902850	microtubule cytoskeleton organization involved in mitosis	1.89	33	1.15E-04	3.88E-02
GO:0045943	positive regulation of transcription by RNA polymerase I	3.21	11	1.16E-04	3.84E-02
GO:0006334	nucleosome assembly	2.24	21	1.33E-04	4.30E-02
GO:0044839	cell cycle G2/M phase transition	1.81	36	1.42E-04	4.46E-02
GO:0072321	chaperone-mediated protein transport	4.31	7	1.58E-04	4.86E-02

Table 1-3. Shared differentially expressed genes among all four population pairs and their associated proteins. Based on the SWISS-PROT database, nine of the ten genes have experimental evidence for the existence of the protein associated with each gene. The tenth gene did not have a name match in the database, so it is not included.

Protein name	**Gene name**
Sodium/hydrogen exchanger 3	*SLC9A3*
Sulfide:quinone oxidoreductase, mitochondrial	*SQRDL*
Cytochrome b reductase 1	*CYBRD1*
Zinc finger protein 800	*ZNF800*
Pulmonary surfactant-associated protein D	*SFTPD*
Glucose-fructose oxidoreductase domain-containing protein 1	*GFOD1*
Carcinoembryonic antigen-related cell adhesion molecule 1	*CEACAM1*
1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-1	*PLCD1*
Cirhin	*CIRH1A*

Table 1-4. GO process terms with significant enrichment in orthogenes with significant differential expression in at least three lineages (FDR < 0.05). The table include the GO term ID, description, the degree of enrichment, the number of differentially expressed genes associated with the GO term (*N*), as well as *P* and FDR-corrected *q*-values.

GO term	Description	Enrichment	*N*	*P*-value	FDR *q*-value
GO:0006820	anion transport	3.63	14	2.81E-05	3.71E-01
GO:0098656	anion transmembrane transport	4.99	10	2.90E-05	1.91E-01
GO:0099587	inorganic ion import across plasma membrane	10.07	5	1.17E-04	5.16E-01
GO:0098659	inorganic cation import across plasma membrane	10.07	5	1.17E-04	3.87E-01
GO:0010107	potassium ion import	14.21	4	1.44E-04	3.80E-01
GO:1990573	potassium ion import across plasma membrane	14.21	4	1.44E-04	3.17E-01
GO:0055067	monovalent inorganic cation homeostasis	6.04	7	1.46E-04	2.74E-01
GO:0015711	organic anion transport	3.73	11	1.66E-04	2.73E-01
GO:0055075	potassium ion homeostasis	22.65	3	2.33E-04	3.42E-01
GO:0071918	urea transmembrane transport	60.4	2	2.72E-04	3.58E-01
GO:0006848	pyruvate transport	60.4	2	2.72E-04	3.26E-01
GO:0015840	urea transport	60.4	2	2.72E-04	2.99E-01
GO:1901475	pyruvate transmembrane transport	60.4	2	2.72E-04	2.76E-01
GO:0045623	negative regulation of T-helper cell differentiation	20.13	3	3.46E-04	3.26E-01
GO:0035728	response to hepatocyte growth factor	20.13	3	3.46E-04	3.04E-01
GO:0015718	monocarboxylic acid transport	6.25	6	3.66E-04	3.01E-01
GO:1902106	negative regulation of leukocyte differentiation	7.74	5	4.21E-04	3.26E-01
GO:0034220	ion transmembrane transport	2.67	15	4.51E-04	3.30E-01
GO:0072521	purine-containing compound metabolic process	3.53	10	5.14E-04	3.56E-01
GO:0043371	negative regulation of CD4-positive, alpha-beta T cell differentiation	16.47	3	6.63E-04	4.37E-01
GO:0031167	rRNA methylation	16.47	3	6.63E-04	4.16E-01
GO:0009150	purine ribonucleotide metabolic process	3.67	9	7.42E-04	4.45E-01
GO:0000466	maturation of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	40.27	2	8.07E-04	4.62E-01
GO:0016338	calcium-independent cell-cell adhesion via plasma membrane cell-adhesion molecules	15.1	3	8.74E-04	4.80E-01
GO:0002822	regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	5.25	6	9.34E-04	4.92E-01
GO:0098742	cell-cell adhesion via plasma-membrane adhesion molecules	4.45	7	9.53E-04	4.83E-01

Figures

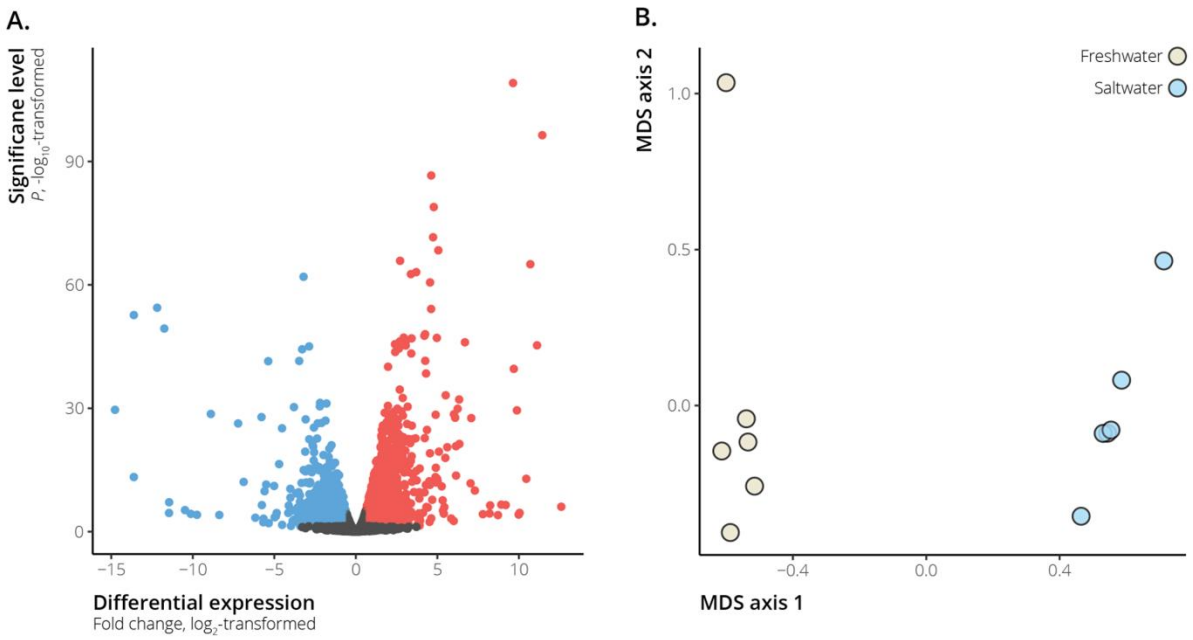


Figure 1-1. A. Volcano plot depicting differentially expressed genes between hypersaline and freshwater *Limia perugiae*. Genes that were significantly differentially expressed between hypersaline and freshwater populations (FDR < 0.05) are indicated by the blue and red points—blue points represent genes downregulated in the hypersaline populations, while red points represent upregulated genes. B. Multi-dimensional scaling (MDS) plot of hypersaline and freshwater *L. perugiae* gene expression profiles. MDS axis 1 separated samples by freshwater vs. hypersaline environments.

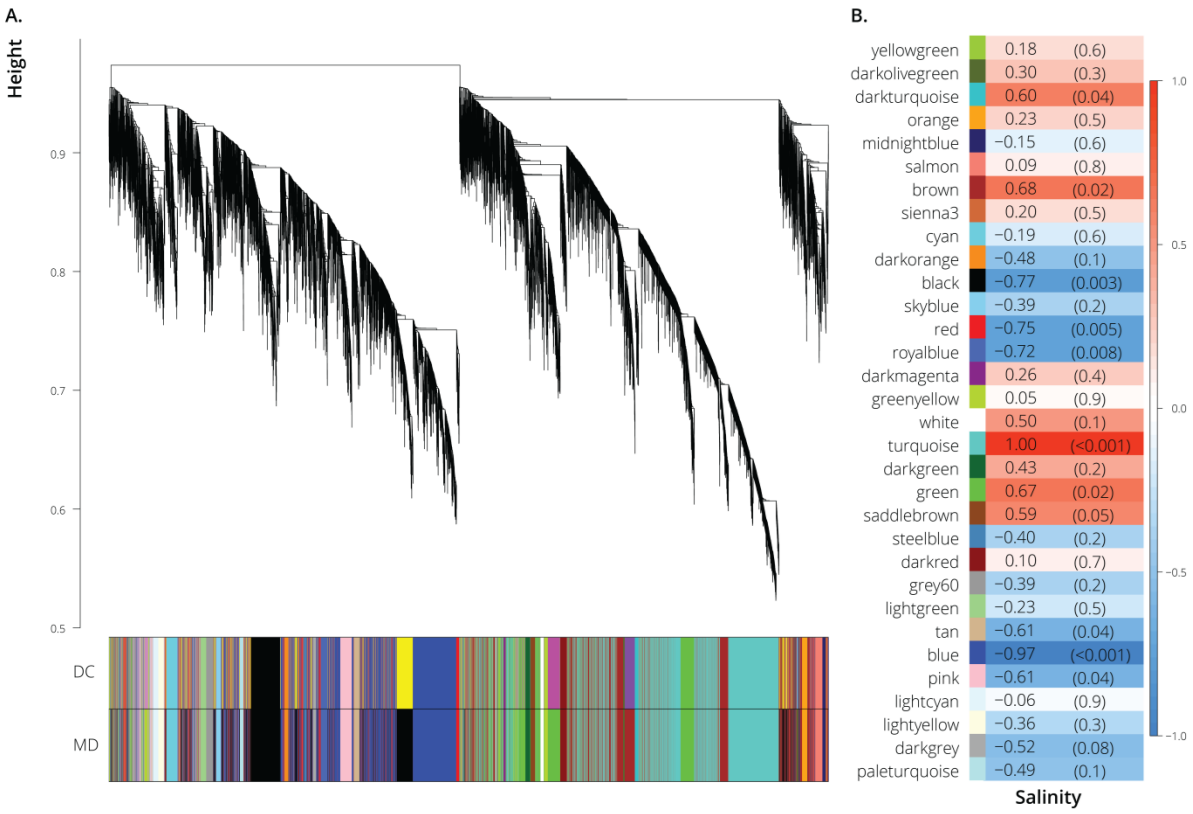


Figure 1-2. Weighted gene co-expression network analysis. A. Average linkage clustering tree based on topological overlap distances in gene expression patterns of *L. perugiae* from freshwater and saltwater habitats. Branches of the dendrogram correspond to modules, as shown in the color bars below. B. Correlation between module eigenvalues and habitat type (freshwater *vs*. saltwater). Each row corresponds to a module of coexpressed genes, and values are Pearson correlation coefficients (left column) and *P*-values (right column in parentheses). Color coloration scales with the correlation coefficient according to the scale bar to the right.

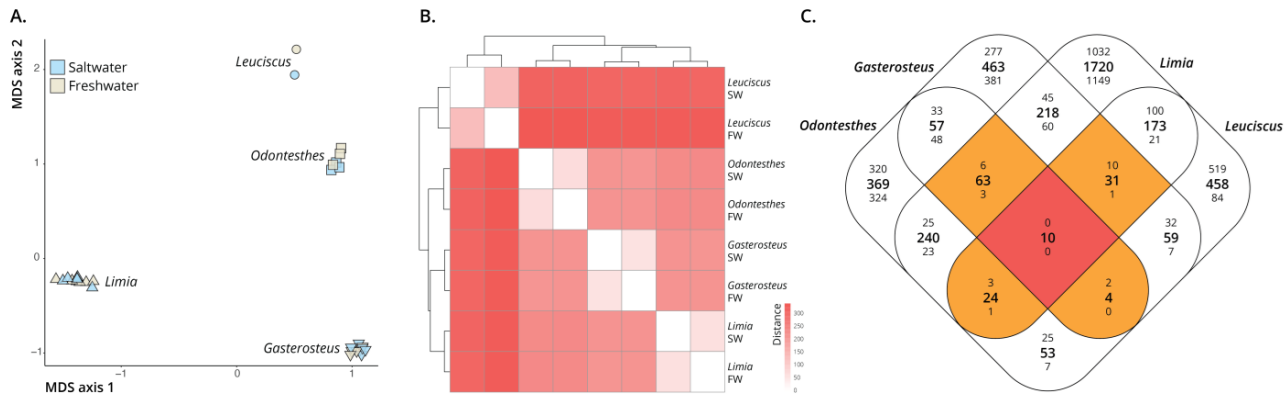


Figure 1-3. A. Multi-dimensional scaling plot (MDS) of the general expression patterns of all the populations included in our analysis. B. Similarity of gene expression profiles of saltwater (SW) and freshwater (FW) populations across different lineages. The majority of variation in gene expression reflects phylogenetic divergence among lineages. C. Shared differentially expressed genes across lineages. The large, central number in each section represents the total number of shared differentially expressed genes among the lineages in that intersection. The top number in each section represents the number of shared up-regulated genes, and the bottom number is the number of shared down-regulated genes in that intersection. Only 10 genes were consistently differentially expressed between all of the SW and FW populations.

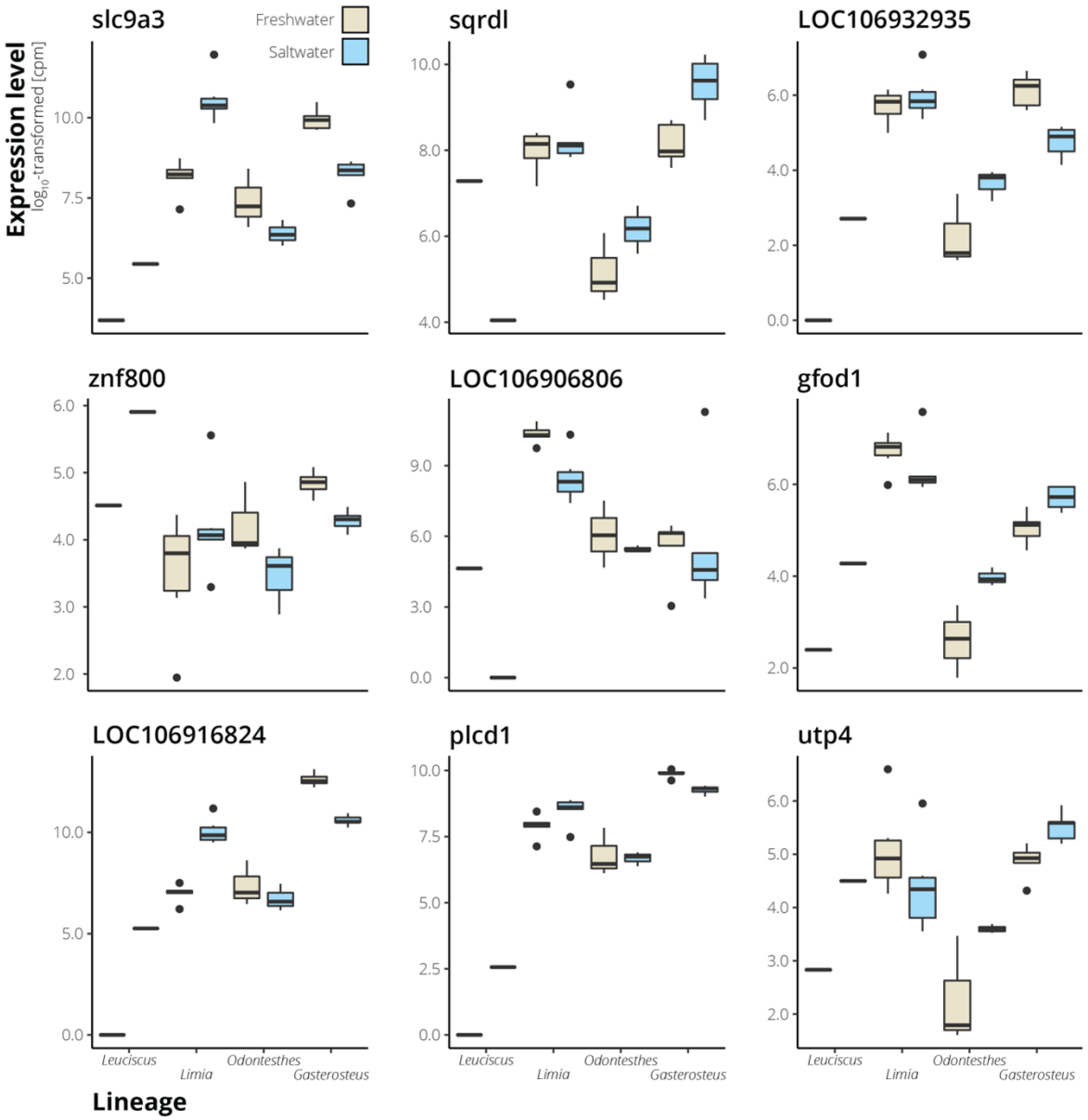


Figure 1-4. Examples of expression variation in shared differentially expressed genes. Nine of the ten shared differentially expressed genes have annotations, and those genes are included here. Magnitude and direction of differential expression is lineage specific.

Chapter 2 - Natural history and trophic ecology of three populations of the Mexican cavefish, *Astyanax mexicanus*

Elizabeth J. Wilson, Michael Tobler, Rüdiger Riesch, Lizeth Martínez-García, Francisco Javier García-De León

Abstract

The Mexican cavefish, *Astyanax mexicanus* (Characidae), has become an important model in evolutionary physiology and developmental biology, providing insights into the evolution of sensory systems, pigmentation, and metabolism. In contrast, comparatively little is known about the natural history and trophic ecology of this elusive cave inhabitant. We investigated cavefish from three independently colonized cave systems (Pachón, Tinaja, and Sabinos), which are located in the Sierra de El Abra of Northeastern Mexico. Samples were collected multiple times throughout the year to investigate variation in body size, sex ratios, proportions of individuals with empty guts, and diet composition. We found consistent differences in body size among caves, and sex ratios were generally female biased, although to varying degrees. Gut content analyses indicated that cavefish consume food throughout the year, and diets are dominated by detritus, plant materials, and aquatic invertebrates. Especially in the Pachón cave, where we had the densest sampling, there was evidence for seasonal changes in diet composition that coincided with the rainy and dry seasons. Our findings potentially suggest that the cave environments in this system are characterized by continual nutrient limitation, rather than intermittent periods of starvation.

*Published as: Wilson, E.J., M. Tobler, R. Riesch, L. Martínez-García & F.J. García-De León (2021): Natural history and trophic ecology of three populations of the Mexican cavefish, *Astyanax mexicanus*. *Environmental Biology of Fishes* 104, 1461–1474. Springer Nature. *Reproduced with permission from Springer Nature.*
https://link.springer.com/journal/10641

Introduction

The Mexican cavefish, *Astyanax mexicanus*, is a well-established model organism for addressing evolutionary questions that span a variety of biological fields, including genomics, developmental biology, metabolic physiology, and behavior (Jeffery 2009; Frøland Steindal et al. 2018; Kowalko 2020). Several populations of *A. mexicanus* independently colonized caves from adjacent surface streams, adapting to the unique challenges that cave environments impose on life (Gross 2012; Pérez-Rodríguez et al. 2021). Caves are often characterized by complete darkness, low food availability, and low temperature variability, driving the convergent evolution of several traits among subterranean fauna (Porter and Crandall 2003; Kowalko et al. 2013; Rétaux and Casane 2013). Comparisons between cave and surface populations of *A. mexicanus*, as well as experiments involving surface-cave hybrids, have provided insight into the molecular mechanisms underlying the evolution of cave-specific traits such as loss of pigmentation, eye regression, and modification of other sensory systems (Gross 2012; Aspiras et al. 2015; Kowalko 2020). The evolution of many traits in *A. mexicanus*, particularly sensory system modifications and metabolic adaptations, are driven by the extreme nutrient limitations brought about by darkness and food scarcity (Moran et al. 2014; Riddle et al. 2018).

To meet nutritional demands while living in darkness, cavefish have evolved ways to discover and capture food non-visually and exhibit extraordinary starvation resistance. Feeding experiments have shown that cave *A. mexicanus* outcompete surface fish for food in darkness (Hüppop 1987; Yoshizawa et al. 2010; Espinasa et al. 2014). Their success in locating prey in darkness is attributed to vibration attraction behavior (VAB), which allows cavefish to move toward a disturbance in the water (Yoshizawa et al. 2010). VAB is largely dependent on the lateral line system, which can detect vibration frequencies at which copepods and other invertebrates move (Yoshizawa et al. 2010). In addition to the sensory enhancements associated with VAB, cavefish have more taste buds than surface fish, particularly on the ventral side of the head, which aid in feeding along the bottom of the cave (Franz-Odendaal and Hall 2006). There is evidence that cavefish have heightened olfaction as well, further contributing to their foraging success in darkness (Kasumyan and Marusov 2015; Hinaux et al. 2016). Multiple cavefish populations have independently evolved a lower feeding angle relative to their surface conspecifics, and while it is yet unclear how this posture affects bottom feeding, convergence in this trait suggests that it could be adaptive (Kowalko et al. 2013).

While it is important to maximize feeding efficiency in caves, cavefish also must tolerate periods of starvation. This is due to the lack of photosynthetic primary producers in caves, as well as inconsistent input of food items from external sources (Xiong et al. 2018). Inconsistencies in food availability are likely a result of seasonality; caves in subtemperate areas have relatively stable temperatures year-round, but the amount of rainfall varies between rainy and dry season (Simon et al. 2017; Xiong et al. 2018). During the rainy season, flash flooding can wash in organic matter from the surface, interrupting periods of limited food in the dry season (Simon et al. 2017). Cavefish have been shown to have lower metabolic rates, a loss of

circadian rhythm in metabolism, elevated appetites, and increased fat storage, allowing them to withstand food shortage for extended periods of time (Hüppop 1987; Moran et al. 2014; Aspiras et al. 2015). The greater body mass of *A. mexicanus* cavefish compared to their surface ancestors has been linked to a mutation that causes insulin resistance, which hinders the body's ability to regulate glucose (Riddle et al. 2018). In humans, the same phenomenon causes health issues that are associated with diabetes (Riddle et al. 2018). Cavefish, however, do not suffer from the same pathological manifestations of disrupted glucose regulation (Riddle et al. 2018). Due to their adaptive outcomes regarding insulin resistance and elevated body fat, cave-adapted *A. mexicanus* have been studied for their relevance to human metabolic diseases as well (Riddle et al. 2018).

While *A. mexicanus* has been an exceptional model for studies in evolutionary developmental biology and metabolic physiology, there is limited research regarding their natural history, including trophic ecology. Espinasa et al. (2017) found that juvenile fish from the Pachón cave in Mexico are well-fed, and their diets are comprised mostly of invertebrates that were consumed as live prey. In contrast, their analysis of five individuals indicated that adult cavefish primarily feed on detritus during both the rainy and dry seasons (Espinasa et al. 2017). An additional study investigating the diet of *A. mexicanus* from another cave, Micos, similarly found that a mixture of a brown substance (guano or detritus), fish parts, and insect exoskeleton parts comprises the majority of the adult cavefish diet (Wilkens and Burns 1972). Consequently, *A. mexicanus* have generally been described generalists with opportunistic feeding habits at maturity (Wilkens and Burns 1972; Trajano 2001; Espinasa et al. 2017). To our knowledge, there has been no other research on temporal change in diet or comparison of gut item composition in *A. mexicanus* between the rainy (June to November) and dry (December to May) seasons (Espino del Castillo et al. 2009).

The objective of this study was to broaden our knowledge of the natural history of cave-adapted *A. mexicanus*, with a particular emphasis on diet composition and diet variation across time and populations. Diet, as well as starvation, influences the well-studied metabolic physiology of cavefish, yet there is limited research on what cavefish actually eat in nature. We assessed body size and sex ratio variation across populations from three caves (Pachón, Sabinos, and Tinaja), and examined the gut contents of populations from the Pachón and Sabinos caves at several time points. Based on previous research, we expected fish from Pachón to be smaller than those from Tinaja (Simon et al. 2017), but we did not have any prior data to inform predictions about how Sabinos cavefish sizes would compare to these populations. Subterranean species have been shown to have female-biased sex ratios, presumably due to limited dispersal and inbreeding causing reductions in genetic variation (Culver and Holsinger 1969; Frank 1990; Premate et al. 2021), and *A. mexicanus* cavefish populations do have less genetic variation than their surface conspecifics (Bradic et al. 2012). We therefore expected to find female-biased sex ratios in all populations. Consistent with an opportunistic feeding style, we also predicted that the cavefish would have a variety of food items in their diet, including detritus. Furthermore, we predicted that all three populations would have a greater proportion of individuals with food in their guts during the rainy season when flooding can wash in organic debris from the surface. Understanding what cave *A. mexicanus* eat in the wild and how their diets change over time will help inform research addressing many aspects of physiology that are heavily impacted by nutrition.

Methods

Study area

The Sierra de El Abra is a limestone ridge in Northeastern Mexico that contains at least 30 caves inhabited by *A. mexicanus* (Figure 2-1) (Mitchell et al. 1977; Espinasa et al. 2018). Throughout the region, surface populations of *A. mexicanus* are also widespread (Miller et al. 2005). The Pachón cave is the northernmost of the three sampled caves, while the Sabinos cave is about 60 km south of Pachón and in close proximity to the Tinaja cave (Figure 2-1) (Avise and Selander 1972; Espinasa and Borowsky 2001). Pachón is also located at the highest elevation in the valley (Avise and Selander 1972). All three caves are comprised of highly branching subterranean tunnels as well as pools that open up into cave chambers (Avise and Selander 1972). The rainy season in this region occurs from June to November, while the dry season is approximately during December to May (Espino del Castillo et al. 2009).

Biometrics

To compare body size variation and sex ratios (quantified as the proportion of males in the population) among the caves, we sampled *A. mexicanus* from the Pachón (N = 611), Sabinos (N = 197), Tinaja (N = 73) caves at several time points between February 2001 and March 2002 (Table 2-1). We sampled the first pools that were accessible upon entering the caves, and we returned to the same pools for each subsequent sampling time. Fish were caught using a seine (2 meters long by 1.5 meters tall) and dip nets. Standard length was measured with calipers to the nearest millimeter, and sex was determined based on the presence of denticles on the anterior rays of the anal fin in males (Borowsky 2008). The fish were then released at the collection location.

All statistical analyses were conducted in R version 4.0.3 (R Core Team 2020). To analyze variation in standard length, we ran generalized linear mixed models (GLMM) with a

Gaussian distribution using R's lme4 package (Bates et al. 2014). We excluded 89 juveniles from Pachón, 15 from Sabinos, and 1 from Tinaja from analyses because their sex could not be determined with confidence. We explored multiple GLMMs to better understand the effects of different predictor variables: (1) a null model, including only the date of collection as a random factor; (2) a model adding sex as a fixed factor; (3) a model adding cave of origin as a fixed factor; (4) an additive model adding sex and cave of origin as fixed factors; and (5) an interaction model adding sex and cave of origin as fixed factors. The small sample unbiased Akaike information criterion (AIC_C) approach was used for model selection (Johnson and Omland 2004), and models with $\Delta AIC_C < 2$ were considered to be equally well-supported (Burnham and Anderson 2001). Similarly, binomial GLMMs were used to test for differences in sex ratios across caves. The null model included the date of collection as a random factor, and it was compared to a model adding cave of origin using AIC_C.

Gut content analysis

Fish were sampled from the Pachón (N = 146), Sabinos (N = 58), and Tinaja (N = 13) caves for gut content analysis (Table 2-1). These individuals were also collected across time points between February 2001 and March 2002 (6 sampling times for Pachón, 3 for Sabinos, and 3 for Tinaja; Table 2-1) using the same capturing methods as previously described. The fish that were collected for gut content analysis were haphazardly selected. We aimed to collect at least 25 fish per sampling time for gut content analysis, unless the population size was small, in which case we aimed for at least 10 fish. During some sampling times, fewer than 10 individuals were collected to avoid depleting the populations. The fish were immediately euthanized in the field, preserved in ethanol, and transported back to the laboratory for dissection and measurements.

The digestive tracts were dissected, and the contents were examined under a dissecting microscope. The presence or absence of the following diet categories was recorded for each individual: fish parts, invertebrates or invertebrate parts, algae, seeds, other plant material, detritus, and inorganic materials. Note that the detritus category included all unidentifiable organic matter, and the inorganic materials category included rock granules and sandy sediment.

Due to the small sample size of Tinaja cavefish for gut content analyses, individuals from that cave were excluded from quantitative analyses. We first compared the frequency of empty guts between caves using binomial generalized linear models (GLMs) as implemented in the lme4 package in R (Bates et al. 2014). We ran five alternative models that were contrasted using AIC_C: (1) a null model, including only the intercept; (2) a model adding collection date as a fixed factor; (3) a model adding cave of origin as a fixed factor; (4) an additive model adding collection date and cave of origin as fixed factors; and (5) an interaction model adding collection date and cave of origin as fixed factors. Additionally, we compared variation in diet composition over time and across caves. We first removed individuals with empty guts from the dataset and analyzed the remaining data using permutational multivariate analysis of variance (PERMANOVA) with Jaccard distances and 999 permutations using the adonis2 function in the R package vegan (McArdle and Anderson 2001; Oksanen et al. 2020). Sampling date, cave of origin, and their interaction term were included as predictor variables. For data visualization, we used non-metric multidimensional scaling via the metaMDS function in R's vegan package (Oksanen et al. 2020).

Results

Biometrics

After excluding juveniles that could not be sexed, the standard lengths of fish used in the analyses ranged from 1.9-6.2 cm in Pachón (N = 522 adults), 1.9-6.4 cm in Sabinos (N = 182), and 2.1-6.4 cm in Tinaja (N = 72). Analyzing variation in standard length using mixed-effects models yielded two models that were equally supported ($\Delta AIC_c < 2$), including both the additive and interactive models with all variables (Table 2-2A). Hence, variation in standard length was affected by both sex and cave of origin. Visualizing the marginal effects of the interaction model indicated that Tinaja cavefish were larger than both Pachón and Sabinos cavefish (Figure 2-2). This trend was consistent for both males and females. In addition, males were larger than females in the Pachón and Sabinos caves, while there was no difference in body size between the sexes in the Tinaja cave (Figure 2-2).

There were significant differences in sex ratios between the caves and across sampling times, as indicated by the mixed model analysis of the frequency of males and females (Table 2-2B). For all the cave populations and across all sampling times between February 2001 and March 2002, females outnumbered males (Table 2-1). The sex ratio in Pachón ranged from 0.01 (11/4/2001) to 0.38 (2/3/2001) (Table 2-1). In the Sabinos population, it ranged from 0.21 (11/2/2001) to 0.42 (6/23/2001) (Table 2-1). Lastly, the sex ratio in the Tinaja population ranged from 0.27 (4/17/2001) to 0.58 (6/23/2001) (Table 2-1). The average sex ratios (with all sampling times included) were 0.17 for Pachón, 0.29 for Sabinos, and 0.39 for Tinaja, indicating that all three populations have a female-biased sex ratio.

Gut content analysis

At any given sampling time, the percentage of individuals with empty guts ranged from 44.4% to 66.7% in the Pachón population and from 0.0% to 66.7% in the Sabinos population (Figure 2-

3A). Although they were not included in the gut content analysis due to small sample sizes, the percentage of Tinaja individuals with empty guts ranged from 25.0% to 66.7% (Table 2-1). There was no evidence that the frequency of individuals with empty guts significantly varied between caves or sampling periods (Table 2-2C). Among the guts that contained food items, there were a variety of diet categories present overall. The Pachón population had all diet item categories (fish, invertebrates, algae, plant material, seeds, detritus, and inorganic materials) represented, and the Sabinos population had all categories represented except for seeds. The median number of diet categories present in fish from both caves was 2, with a range of 1-6 in Pachón and 1-4 in Sabinos (Figure 2-3B). Tinaja cavefish had a median of 1 diet item category, ranging from 1-4 categories (Appendix Table B-1).

Examining diet composition among caves and sampling periods using a PERMANOVA, we found significant effects for cave (PERMANOVA: $R^2 = 4.6\%$, $P = 0.002$) and time of sampling (PERMANOVA: $R^2 = 3.7\%$, $P = 0.002$), as well as for the interaction of the two factors (PERMANOVA: $R^2 = 2.1\%$, $P = 0.046$; Table 2-3). The ordination plot based on NMDS (Figure 2-4A) indicated that the Pachón population exhibited greater variation in diet item composition across sampling times than the Sabinos population. While not all diet item compositions changed significantly over shorter time scales (such as 1-3 months), there were significant gut content changes between sampling times that were 4 or more months apart. There is some evidence that sampling times toward the end of the dry season cluster together, as the points for February 2001 (P1), April 2001 (P2), and March 2002 (P6) group near each other. This grouping is driven largely by the presence of seeds and organic detritus. The greatest sequential shifts in gut content composition occurred across three time points, from 4/19/2001 (P2) to 6/24/2001 (P3) to 11/3/2001 (P4). These dates represent a period from the end of the dry season (April 2001) to the

early part of the rainy season (June 2001) and then finally to the end of the rainy season (November 2001) (Figure 2-4B). Time point P3, the June 2001 sampling date, particularly stands out in the middle of this shift, as it does not cluster with any other sampling time and is uniquely characterized by the presence of plant material. Following the end of the rainy season in November 2001 (P4), the next sampling point in February 2002 (P5, during the dry season) groups closely with the November 2001 time point (Figure 2-4). The diets at these two sampling times cluster together due to the presence of inorganic materials. As previously mentioned, the final sampling time for Pachón (March 2002, point P6) is more similar to the 2001 sampling times for February and April, so there is a considerable difference in gut content composition over the course of a month between February 2002 (P5) and March 2002 (P6).

In the Sabinos population, there were no significant changes in diet from the rainy to dry season. All 3 sampling times cluster together, and the separation of this grouping from the Pachón samples is largely driven by the abundance of invertebrates in the diet. All the time points for Sabinos most closely reflect points P4 and P5 from Pachón, which are the gut content samples collected in November 2001 and February 2002. November 2001 marks the transition from the rainy season to the dry season, while February 2002 is within the dry season (Figure 2-4B). Invertebrates and inorganic materials drive the clustering of these two Pachón time points with the Sabinos time points.

In the Tinaja population, gut contents were described for individuals collected at three time points—2/4/2001 (n=6), 6/23/2001 (n=3), and 11/2/2001 (n=4) (Table 2-1). While the majority of individuals had empty guts, the gut contents from cavefish sampled in February 2001 included inorganic materials and fish parts (Appendix Table B-1). In June and November of that same year (during the rainy season), plant material was more frequently found in cavefish guts

(Appendix Table B-1). Invertebrates were only identified in cavefish guts during the November sampling time, and one individual from November had seeds and detritus in addition to invertebrates and plant material (Appendix Table B-1).

Discussion

The cavefish *Astyanax mexicanus* has been well-studied for its adaptations to a subterranean lifestyle, but the natural history and diet of cavefish in the wild are less well known. Our analyses found variation in standard length, sex ratios, and diet between cave populations of *A. mexicanus*, as well as variation in sex ratios and diet through time. Across three cave populations, Tinaja cavefish were larger than Pachón and Sabinos cavefish. Males were larger than females in Pachón and Sabinos, while there was no difference in size between the sexes in Tinaja. Additionally, all three cave populations had a female-biased sex ratio. Most importantly, our gut content analyses indicated that the majority of cavefish consumed food in similar proportion throughout the year. Pachón and Sabinos populations differed in their diet, which was primarily driven by the frequency of detritus, invertebrates, seeds, other plant material, and inorganic materials. Furthermore, in Pachón, for which we had the most rigorous sampling, there was also evidence for dietary shifts occurring across the rainy and dry seasons.

Sex ratios, body size, and energetics in subterranean environments

Animals living in subterranean environments experience space limitations and isolation from the surface, often driving changes in population structure (Premate et al. 2021). Reductions in heterozygosity and female-biased sex ratios have been described for many subterranean species and are attributed to inbreeding due to constraints in physical space (Konec et al. 2015;

Behrmann-Godel et al. 2017; Premate et al. 2021). *Astyanax mexicanus* cavefish populations also have reduced genetic variation when compared to their surface ancestors (Bradic et al. 2012), and we found female-biased sex ratios in all three cave populations sampled. As has been suggested for other subterranean species, increased mate competition in a confined space and more frequent mating between siblings, further reinforced by kin selection, can result in female-biased sex ratios because competition between brothers favors a decline in investment in males (Frank 1990; Premate et al. 2021). It is possible that these interacting forces are similarly contributing to the female-biased sex ratios in the Pachón, Sabinos, and Tinaja caves. Differential sex determination between males and females can also drive variation in sex ratios. Interestingly, Pachón cavefish carry B chromosomes that are predominantly found in males and contain two copies of a sex determination gene that is only expressed in male gonads (Imarazene et al. 2021). When this gene was knocked out in a study by Imarazene and colleagues (2021), males were sex-reversed, indicating a role for the gene in driving sex determination in Pachón cavefish. Populations that are biased toward one sex have been shown to carry B chromosomes that are predominant in the more abundant sex, supporting a role for B chromosomes in sex determination (Beladjal et al. 2002; Yoshida et al. 2011; Clark and Kocher 2019). Accordingly, although we might expect male-biased sex ratios from these patterns, skewed sex ratios can emerge from a variety of environmental, genetic, and demographic factors, and variation in sex determination is one of many potential contributors to sex ratio bias (Beladjal et al. 2002; Yoshida et al. 2011; Cornelio et al. 2017; Clark and Kocher 2019). Furthermore, in a study of 200 Pachón cavefish that were born in captivity from laboratory stocks, the sex ratio was found to be 1:1, suggesting that mechanisms of sex determination alone are not the only factors responsible for sex ratio bias, but there are likely environmental effects influencing sex ratios in

wild population (Imarazene et al. 2020). Finally, if female-biased sex ratios are a result of selection against males for any number of reasons, then the presence of a male sex determination gene may be selectively maintained in the population (Natri et al. 2019). Regardless of the selective pressures associated with cave environments, female-biased sex ratios are common in other fish species due to greater survival or longevity of females (Reichard et al. 2014; Fryxell et al. 2015). Further investigations into the genetic, demographic, and environmental interactions in wild cavefish populations, as well as comparisons between cavefish and surface fish life histories and their sex-specific differences, will provide insight into the mechanisms underlying female-biased sex ratios in *A. mexicanus* cavefish.

In addition to limitations on population size and genetic diversity, subterranean environments with their lack of sunlight and nutrients impose significant energetic constraints on the organisms that inhabit them. Cave-adapted animals cope with these challenges by conserving energy in various ways, such as optimizing body stores and lowering metabolic rate relative to ancestral surface populations (Hüppop 1986; Hervant and Renault 2002; Hervant 2012; Xiong et al. 2018). Additionally, body size reduction in cave animals can help lower energetic demands and may occur independently or in conjunction with a decrease in metabolic rate, potentially changing the allometric relationship between body size and metabolic rate (Passow et al. 2015). Like many subterranean fauna, cavefish species are generally smaller than their surface relatives, but differences in nutrient availability between cave populations may drive variation in strategies for conserving energy, including body size modifications (Hüppop 1985; Trajano 2001; Simon et al. 2017). We did find differences in standard length between the Pachón, Sabinos, and Tinaja cave populations, and Tinaja cavefish were larger than Pachón and Sabinos individuals. In support of this finding, there is evidence for a genetic limitation on size for some cave

populations, as Pachón cavefish have been shown to reach smaller sizes than Tinaja cavefish and another population, Subterráneo, in the wild and when reared under identical treatments in the laboratory (Simon et al. 2017). If cave-specific body size variability is dependent on food limitation, it is possible that the Pachón and Sabinos caves are more nutrient-poor than Tinaja and that this drives a more substantial body size reduction in these cave populations. This matches the observation that Tinaja's mud is richer in carbon content than Pachón's (Simon et al. 2017). Alternatively, Tinaja cavefish might indeed face similar nutritional challenges as Pachón and Sabinos cavefish but utilize other strategies for conserving energy or optimizing nutrient uptake and consequently do not rely as much on body size reduction. For example, Pachón cavefish accumulate fat stores earlier in development than Tinaja cavefish, providing them starvation resistance sooner in life, but Tinaja cavefish exhibit hyperphagia when food is available, unlike those from Pachón (Aspiras et al. 2015; Xiong et al. 2018). Accordingly, any combination of physiological or behavioral traits could provide a means of resisting starvation and maintaining energy homeostasis. To understand how body size reduction interacts with other energy conservation strategies in response to nutrient variation across cave environments, it is crucial to better understand the trophic ecology of cavefish in their natural habitats.

Nutrient limitation and starvation

Because consistent or periodic starvation forces cave-dwelling animals to reduce their energy expenditure, understanding to what extent individual populations experience starvation can provide insight into the drivers of behavioral, morphological, and physiological traits of subterranean fauna. Caves are expected to limit food access and cause periods of starvation in cave-adapted *A. mexicanus*. We predicted that during the rainy season, more cavefish would

have food in their guts due to flooding washing in debris. However, we found that Pachón and Sabinos cavefish seem to eat throughout the year, without significant seasonal variation in the frquency of empty guts. There were consistently large portions of the population that had empty guts (44%-67% of the population for both Pachón and Sabinos at all sampling times, with the exception of one collection of 3 Sabinos cavefish that all had full guts) (Figure 2-3A). This indicates that the caves might not experience substantial changes in food availability between the rainy and dry seasons or that shifts in food accessibility are short-lived. For example, the way that water from the surface moves into the caves could play a role in how effectively nutrients reach the cave pools. There is evidence that water influx from flash flooding in Pachón comes more from sump inflow than stream inflow because the Pachón cave is at a higher elevation than the nearby Río Mante, potentially reducing the quantity and frequency of debris entering the cave (Keene et al. 2015; Espinasa et al. 2017; Simon et al. 2017). This would diminish seasonal impacts on the percentage of individuals with empty guts, which was corroborated by our findings in both caves. Flooding in Sabinos, however, may be more likely to push debris in, as it does not have that elevational effect and fragments of organic material have been documented in the cave (Keene et al. 2015). Although we found no evidence that there is enough influx of organic debris to impact the gut contents of cavefish during the rainy season, seasonal trends for Sabinos could potentially emerge with more extensive temporal sampling. In all the populations, sampling at additional time points in the rainy season (particularly in August-October) could reveal greater influxes of food, even if a significant proportion of the populations do have food in their guts during the drier parts of the year.

Alternatively, there could be bursts of nutrient availability after heavy rains in the caves, but the effects of these flash floods might be brief, and thus not detected at our sampling times. If

cavefish do experience consistent low food availability, then that may explain why cave-adapted *A. mexicanus* are smaller than surface *A. mexicanus*. Intermittent starvation is predicted to select for larger individuals because of the greater energy stores associated with larger size, and the periods of food abundance would provide opportunities for larger individuals to gain the energy required to maintain homeostasis (Culver et al. 1995). If food availability is instead consistently low throughout the year, the benefit of using less energy to power a smaller body may outweigh the cost of having reduced energy stores, particularly if buffered by other adaptations for energy storage (such as increased fat accumulation in cave-adapted *A. mexicanus*).

Finally, it is also important to consider that reduced rates of energy acquisition may not only be driven by low resource availability, but by the cost-to-benefit ratio of foraging. In cave populations of *Poecilia mexicana*, 35% of guts are consistently empty, even though these fish occur in a chemoautotrophic cave with high resource availability (Tobler 2008; Roach et al. 2011). In these populations, the high frequency of empty guts arises from intrinsic limitations associated with a trade-off between benthic foraging and surface breathing in hypoxic water (Tobler et al. 2009). Other intrinsic limitations may be associated with the cost of food finding and processing, digestion, and nutrient assimilation (Whelan and Brown 2005; Secor 2009). Hence, intermittent starvation in nutrient poor environments could also be an adaptive strategy to balance costs associated with energy acquisition.

Trophic ecology and seasonal variation in diet

In most subterranean environments, energy sources that are carried in from surface habitats can be in the form of suspended organic particles, invertebrates, plant pieces, microorganisms, bat guano, and carcasses (Hervant 2012). Consequently, hypogean animals are expected to be

generalist feeders (Trajano 2001). Several cavefish species show similar feeding preferences to their epigean relatives (as omnivores or generalist carnivores), suggesting that generalist feeding could be a preadaptation for cave colonization (Trajano 2001). Despite dietary overlap between related populations of surface fish and cavefish, behavioral and sensory adaptations for finding food often arise with cave colonization (Soares and Niemiller 2013). We found a variety of food items present in the diet of cave-adapted *A. mexicanus* from the Pachón and Sabinos caves, including plant material and invertebrates, which comprise a large part of the diet of surface *A. mexicanus* as well (Riddle et al. 2019). Unlike surface *A. mexicanus*, however, the enhanced sensory systems for food detection and lower feeding angle of cavefish suggest that they feed at the bottom as well as consume prey moving in the water column. We commonly found detritus and inorganic materials in the cavefish guts, supporting the hypothesis that adult cavefish feed along the cave floor. Detritus is comprised of a mixture of organic materials and can be an energy-rich food source for hypogean fauna (Fenolio et al. 2006; Tobler et al. 2015). Guano has been identified as a major component of detritus in other caves, and stable isotope analyses of cave-adapted salamander gut contents have revealed that guano could have a similar nutritional value to invertebrate prey (Fenolio et al. 2006). Detritus is a heterogeneous mixture (Moore et al. 2004), and guano itself contains protein-rich exoskeleton fragments (Wilkens and Burns 1972; Fenolio et al. 2006), so further isotope analyses and nutrient assimilation studies will be necessary for understanding whether cavefish selectively uptake certain components of detritus based on their nutritional value (Lin et al. 2007; Enyidi et al. 2013). Our finding of inorganic materials, invertebrates, seeds, and other plant material in addition to detritus is in line with Espinasa et al.'s (2017) finding of a mixture of organic materials and arthropods in adult Pachón cavefish and supports that *A. mexicanus* cavefish are opportunistic feeders.

While it is true that cave *A. mexicanus* seem to eat whatever they have available, we detected trends in what food items are present in their guts based on cave of origin and seasonality. Diet varied between the rainy (June to November) and dry (December to May) seasons for the Pachón population, but this seasonal variation was not as prevalent for Sabinos. In Pachón, there is a very clear pattern in diet change over the course of the year, which is completed with the last sampling time in March 2002 falling into place with collection time points from almost exactly a year prior. Dry season diets were dominated by the presence of seeds, detritus, and inorganic materials at different sampling times. The time point immediately following the April 2001 collection date, in June 2001, represents the wettest part of the year out of all the sampling times, and the diet composition at this time point is categorized by the presence of plant material. This suggests that during the rainy season, macroscopic organic material from the surface can make its way into Pachón, indicating that surface debris can enter the cave and is consumed by cavefish. As the rainy season transitioned into the dry season, the following November 2001 and February 2002 sampling times (points P4 and P5 in Figure 2-4) clustered together based on the presence of inorganic materials. Periods of time during the dry season could make the caves very nutrient poor without rainfall, possibly causing the cavefish to pick up more inorganic materials than anything of nutritional value. The November 2001 and February 2002 sampling times for Pachón are also the most similar to all three Sabinos time points, but the Sabinos collection times are shaped by the presence of invertebrates. It should be noted as well that six sampling times were included for Pachón (n=146), and three were included for Sabinos (n=64), so this difference in resolution could play a role in the lack of seasonal change detected for Sabinos in comparison to Pachón. Future research including multiple sampling points throughout the year, particularly with additional times during the rainy season,

and further gut content analysis to include food item measurements by mass will be necessary for understanding how and what cave *A. mexicanus* eat in nature.

Conclusions

Our findings show that while cave-adapted *A. mexicanus* are known to be smaller than closely-related surface populations, there is variation in body size of adult cavefish between caves. Cave-specific differences in body size could be driven by variation in food availability between cave environments, and body size modifications are likely accompanied by other behavioral and metabolic adaptations for reducing energetic costs. Strategies for maximizing energy use efficiency such as metabolic rate suppression, accumulation of fat stores, and hyperphagia may accompany body size reduction and co-occur to varying degrees. Additionally, *A. mexicanus* cavefish were found to eat during both the rainy and dry seasons, suggesting that the cave environments in this system are potentially characterized by continual nutrient limitation, rather than intermittent periods of starvation. These conditions are expected to select for smaller body size and an opportunistic feeding style, and our diet analyses revealed a variety of food items that provide support for opportunistic feeding. Detritus, inorganic materials, invertebrates, seeds, and other plant material drove differences in diet over time and between two cave populations. Open questions regarding the diet of cave-adapted *A. mexicanus* are centered on understanding how the quantity, quality, and relative proportions of food items in their guts seasonally vary and identifying what nutrients most contribute to their metabolic physiology.

Tables

Table 2-1. Individuals collected for biometrics and gut content analysis at time points between February 2001 and March 2002. $N_{biometrics}$ represents all the individuals included in the biometrics analyses, and the sex ratios are reported as the proportion of the population that is male. N_{gut} is the subset of individuals that were then investigated for their gut contents. Finally, N_{empty} is the number of individuals with empty guts, which is also represented as a percentage ($100* N_{empty}/N_{gut}$).

Cave	Date	$N_{biometrics}$	Sex ratio	N_{gut}	N_{empty}	Percent empty
Pachón	2/3/2001	50	0.38	34	18	52.9%
	2/21/2001	67	0.21			
	4/19/2001	39	0.13	34	18	52.9 %
	6/24/2001	100	0.16	27	12	44.4 %
	11/3/2001	70	0.11	29	15	51.7 %
	11/4/2001	31	0.01			
	2/21/2002			12	8	66.7 %
	2/22/2002	21	0.24			
	2/23/2002	19	0.21			
	3/22/2002	65	0.03	10	5	50.0 %
	3/23/2002	38	0.08			
	3/24/2002	22	0.23			
Sabinos	4/18/2001	35	0.23	25	11	44.0 %
	6/23/2001	76	0.42	30	20	66.7 %
	11/2/2001	71	0.21	3	0	0.0 %
Tinaja	2/4/2001	26	0.35	6	4	66.7 %
	4/17/2001	11	0.27			
	6/23/2001	24	0.58	3	1	33.3 %
	11/2/2001	11	0.36	4	1	25.0 %

Table 2-2. Results of GLMM analyses that examined variation in (A) body size (measured as standard length) and (B) the frequency of males and females (sex ratios) and (C) results of GLM analysis that examined variation in the frequency of individuals with empty guts. Note that collection date is designated as a random effect (1|Date) in analyses of standard length and sex ratios.

	AIC_c	**ΔAIC_c**	**df**	**AIC_c Weight**
(A) Standard length				
(1|Date) (null model)	1950.4	61.9	3	<0.001
Sex + (1|Date)	1931.8	43.3	4	<0.001
Cave + (1|Date)	1902.6	14.1	5	<0.001
Sex * Cave + (1|Date)	1888.8	0.3	8	0.46
Sex + Cave + (1|Date)	1888.5	0.0	6	0.54
(B) Sex ratios				
(1|Date) (null model)	778.7	4.4	2	0.1
Cave + (1|Date)	774.3	0.0	4	0.9
(1|Date) (null model)	778.7	4.4	2	0.1
(C) Empty guts				
(C) Empty gutsIntercept (null model)	284.3	0.0	1	0.679
Intercept + Cave	286.3	2.0	2	0.249
Intercept + Date	290.3	6.0	6	0.034
Intercept + Cave + Date	292.5	8.1	7	0.012
Intercept + Cave * Date	291.9	7.5	9	0.016

Table 2-3. PERMANOVA results of gut content composition relative to cave of origin (Pachón and Sabinos), sampling time (February 2001-March 2002), and their interactions**.**

	df	**Sum of squares**	R^2	***F***	***P***
Cave	1	1.3251	0.04576	4.7501	0.002
Time	1	1.0701	0.03695	3.8360	0.002
Cave:Time	1	0.6179	0.02134	2.2149	0.046
Residual	93	25.9443	0.89595		
Total	96	28.9575	1.00000		

Figures

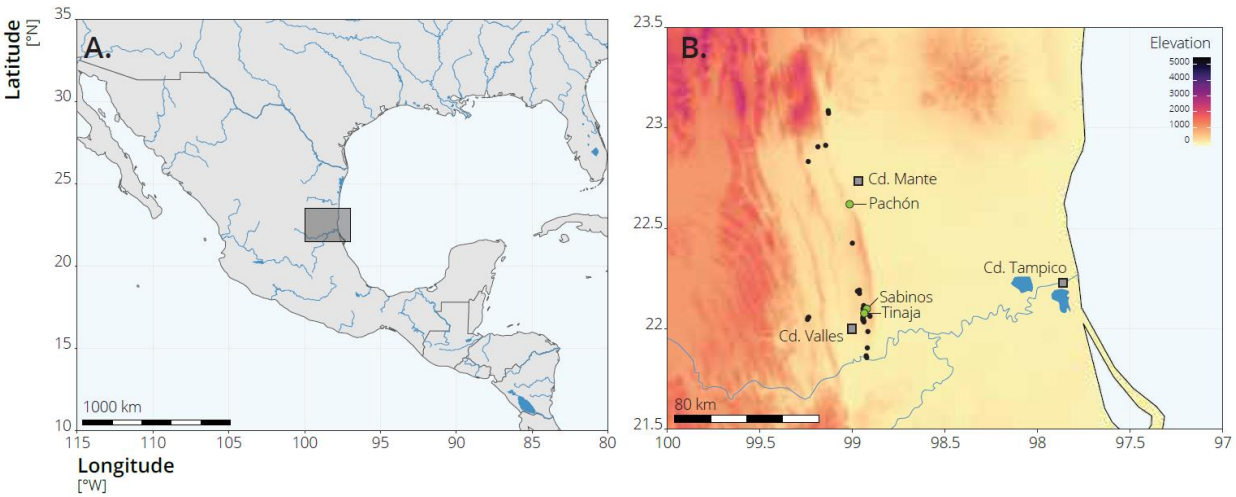


Figure 2-1. A. The location of the study region in northeastern Mexico is indicated by the gray square. B. Caves studied here are indicated by green circles. Black circles indicate the location of other caves inhabited by *A. mexicanus*. Maps were created in R using data associated with the rnaturalearth (South 2017) and elevatr packages (Hollister 2021)

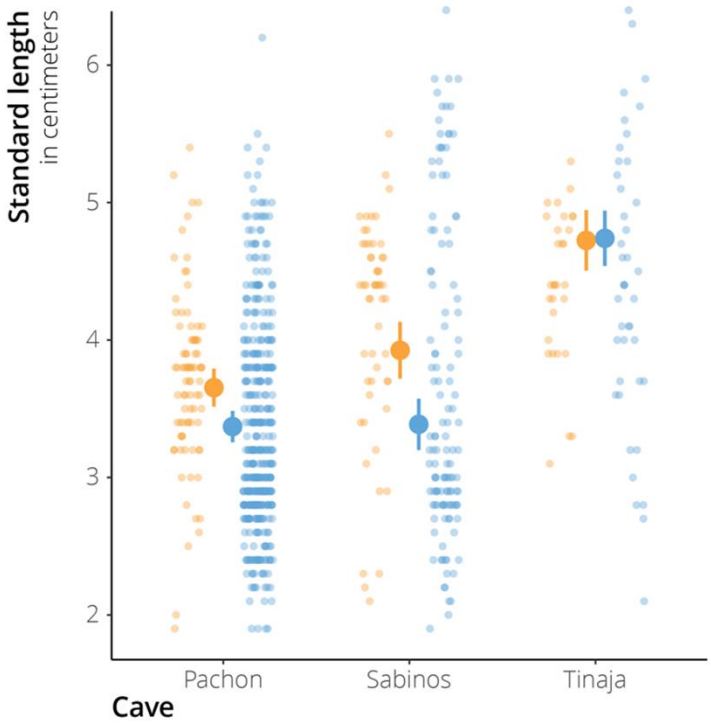


Figure 2-2. Large circles with bars represent marginal means and standard errors of standard length across the Pachón, Sabinos, and Tinaja caves, as inferred by the interaction model: standard length ~ sex * cave + (1|Date). Small circles represent the raw data used in the model. Males are indicated in orange, females in blue

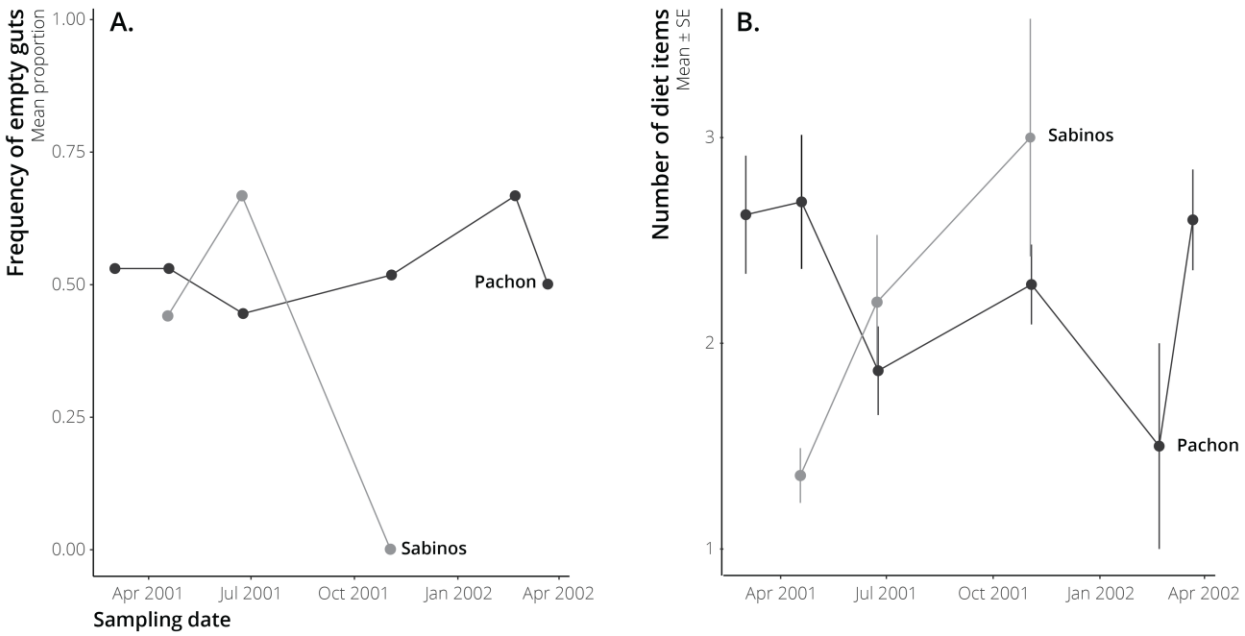


Figure 2-3. A. Frequency of individuals with empty guts in the Pachón and Sabinos populations across multiple sampling times. B. The average number (and standard error) of diet categories per individual for the Pachón and Sabinos populations across multiple sampling times

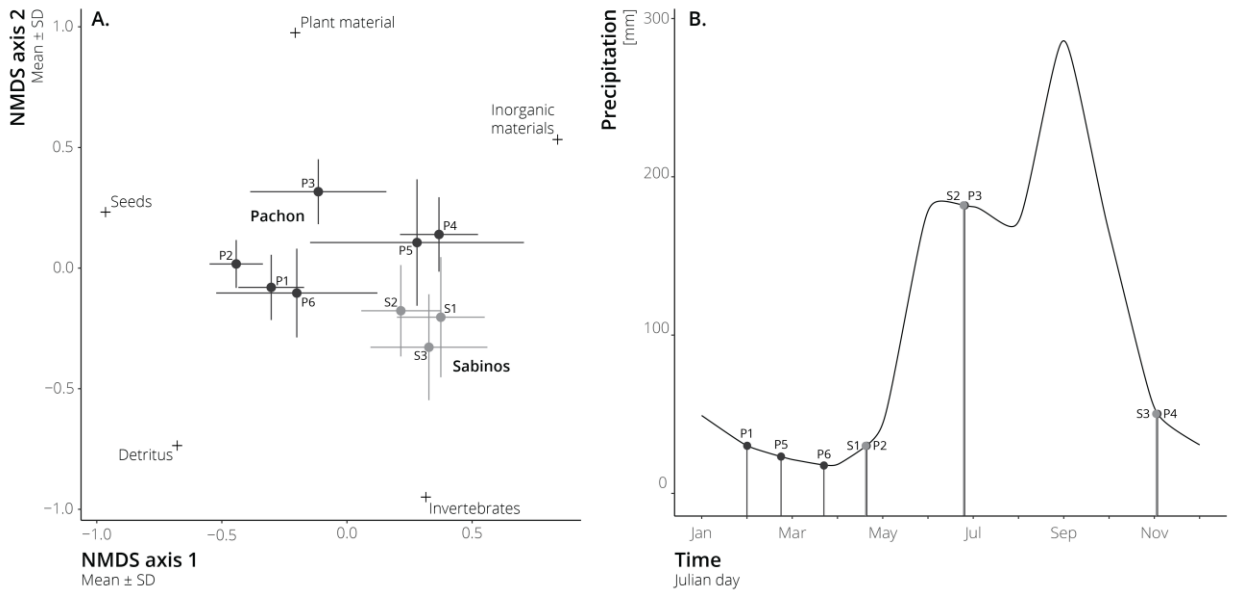


Figure 2-4. A. NMDS plot showing the gut composition (means and standard deviations) of caves and sampling times for the Pachón and Sabinos populations. Ordination ranking was conducted using Jaccard distance, and the stress with 3 dimensions is 0.079. The plot points for Pachón and Sabinos are numbered in order corresponding to their sampling times as follows: for Pachón, P1 = 2/3/2001, P2 = 4/19/2001, P3 = 6/24/2001, P4 = 11/3/2001, P5 = 2/21/2002, P6 = 3/22/2002; for Sabinos, S1 = 4/18/2001, S2 = 6/23/2001, S3 = 11/2/2001. B. Precipitation (in mm) during a typical year near the Sierra de El Abra region (monthly average rainfall data from Tampico, Mexico; World Weather Online 2021). Points P1–P6 and S1–S3 correspond to the sampling times in Figure 2-4A.

Chapter 3 - Host-microbiome associations in livebearing fishes adapted to toxic environments rich in hydrogen sulfide

Elizabeth J. Wilson, Sonny T.M. Lee, Lydia H. Zeglin, Lenin Arias-Rodriguez, and Michael Tobler

Abstract

Organisms inhabiting extreme environments must tolerate a variety of physiochemical stressors. In some cases, host-associated microbial communities facilitate survival of their hosts in extreme environments, but extremophile symbioses have not been identified in vertebrates. We used 16S rRNA amplicon sequencing to investigate commonalities and differences in the gut microbial communities of livebearing fish (*Poecilia mexicana* species complex, Poeciliidae) that have repeatedly colonized toxic sulfide streams in southern Mexico. We found shared core gut microbiomes across habitat types and drainages but also differences in the microbiomes between sulfidic and nonsulfidic populations, both in terms of patterns of diversity and community composition. Most importantly, we documented convergent changes in microbiome composition across independent sulfide spring lineages. These patterns were consistent when we analyzed whole gut microbiomes and primarily host-associated microbiomes that excluded taxa that are commonly found in the environment. Our analyses also revealed several microbial taxa associated with sulfide spring colonization that have been implicated in symbioses before and may influence the host's tolerance to the extreme environmental conditions.

Introduction

Extreme environments—such as hot springs, hypersaline lakes, and deep-sea hydrothermal vents—impose harsh physiochemical stressors on the organisms that inhabit them (Rothschild and Mancinelli 2001; Li et al. 2014; Shu and Huang 2022; Sierra et al. 2022). Diverse archaea, bacteria, and eukaryotes have colonized extreme environments, often tolerating multiple stressors at once (Shu and Huang 2022). Studying how different sources of selection shape the biological communities in extreme environments provides opportunities to address questions about mechanisms of adaptation and drivers of diversification (Jorquera et al. 2016; Tobler et al. 2018; Dick 2019). Most organisms thriving in extreme environmental conditions (i.e., extremophiles) are microorganisms (Rampelotto 2013; Singh et al. 2019) with unique protective and metabolic adaptations (Singh et al. 2019). The microbial communities of extreme environments are often distinct in their composition and diversity, and they can mediate ecological processes as free-living or host-associated communities (Shu and Huang 2022). Host-associated microbiomes of animals are also known to play important roles in the host's functions, and interactions between hosts and their microbes may allow some species to persist in extreme environments (Hoang et al. 2021). Microbiomes facilitating host adaptation to extreme environments have been identified for invertebrates (Hoang et al. 2021), but despite the well-established functional importance of vertebrate microbiomes to host physiology and fitness (Smith et al. 2017; Suzuki 2017; Fontaine et al. 2022), the putative role of microbial symbionts mediating adaptation in vertebrate extremophiles remains unstudied.

Host-associated microbiomes can mediate host tolerance and adaptation to environmental stress (Vrijenhoek 2013). For example, extremophile invertebrates, such as annelids and shrimp in deep-sea hydrothermal vents, depend on symbiotic sulfur-oxidizing bacteria to thrive in

presence of toxic hydrogen sulfide (H_2S) (Powell and Somero 1986; Petersen et al. 2010; Vrijenhoek 2013). Bacteria in the classes Gammaproteobacteria and Epsilonproteobacteria are well-known for their symbioses with marine animals living in H_2S-rich environments (Petersen et al. 2010; Zbinden et al. 2015; Breusing et al. 2020b). By using sulfur compounds from the environment as an energy source, these bacterial symbionts can provide nutrients to the host and detoxify H_2S in host tissues (Sun et al. 2021). Analyzing microbiomes within an evolutionary framework provides opportunities to understand how host-microbe associations arise and what role they may play in host adaptation.

In the Río Grijalva basin of southern Mexico, livebearing fishes in the *Poecilia mexicana* species complex have colonized naturally-occurring H_2S-rich springs (Tobler et al. 2018). There are several sulfide springs located along four tributaries (Pichucalco, Ixtapangajoya, Puyacatengo, and Tacotalpa) of the Río Grijalva, and evolutionarily independent lineages of the *P. mexicana* species complex have colonized several of the sulfide springs in this region (Tobler et al. 2018). H_2S is highly toxic for aerobic organisms even in small concentrations, as it disrupts oxidative phosphorylation by binding to cytochrome c oxidase (Complex IV) in the mitochondrial respiratory chain (Cooper and Brown 2008; Kelley et al. 2016; Tobler et al. 2018). In aquatic ecosystems, H_2S also causes and exacerbates hypoxia (Millero 1986; Tobler et al. 2018). Comparisons between *P. mexicana* populations in sulfidic and adjacent nonsulfidic habitats have revealed adaptations to the extreme environmental conditions, including modifications in cytochrome c oxidase (the primary toxicity target of H_2S) and the sulfide:quinone oxidoreductase pathway that is involved in detoxifying H_2S (Pfenninger et al. 2014; Tobler et al. 2018; Greenway et al. 2020). Beyond the direct responses to H_2S toxicity, the multifarious sources of selection in sulfide springs have driven changes in a variety of traits in *P.*

mexicana (Tobler et al. 2018). For example, nonsulfidic *P. mexicana* primarily feed on detritus and algae, while the sulfidic *P. mexicana* have a diet comprised mostly of bacteria and invertebrates (Tobler et al. 2015). This shift in trophic ecology has been accompanied by a shortening of intestines in sulfidic fish relative to nonsulfidic fish (Tobler et al. 2015). Additionally, although microbial mats are abundant in sulfide springs, sulfidic fish spend the majority of their time at the surface of the water to maximize their oxygen uptake (Plath et al. 2007; Lukas et al. 2021), limiting the amount of time they can spend feeding from the benthos (Tobler et al. 2009) and constraining energy availability (Tobler et al. 2006; Passow et al. 2017). H_2S toxicity, hypoxia, shift in trophic resource use, and effective energy limitation in sulfide springs are all selective pressures that may shape the microbiomes of sulfidic fish and drive the evolution of host-microbe associations. Currently nothing is known about how the microbiomes of livebearing fishes have changed upon sulfide spring colonization and whether they help mediate tolerance in adapted populations. Sulfur-oxidizing Gammaproteobacteria and Epsilonproteobacteria found in symbiotic relationships with deep-sea extremophile invertebrates are present in the environmental microbiome of sulfide springs, suggesting a potential for host-microbe interactions contributing to adaptation of livebearing fishes to sulfide spring ecosystems (Hotaling et al. 2019). As in other systems, possible adaptive symbioses might enhance nutrient uptake, provide a protective barrier to host tissues, or aid in H_2S detoxification (Vrijenhoek 2013; Colston and Jackson 2016).

To characterize gut microbiomes in sulfide spring fishes and identify potential symbionts that could mediate adaptation to extreme environments, we used 16S rRNA amplicon sequencing to address the following questions: (1) What are the commonalities and differences in gut microbiomes of *P. mexicana* from sulfidic and nonsulfidic habitats and across river drainages?

To address this question, we compared different biodiversity metrics among populations and identified the core microbiomes of populations within and across habitat types. (2) How do gut microbiomes compare to environmental microbiomes, and how do patterns of diversity vary among microbial communities that are primarily host-associated? Gut microbiomes are at least in part shaped by environmental microbiomes, and especially for benthic detritivores like *P. mexicana*, many taxa may be of dietary origin. We therefore characterized environmental microbiomes and used these data to identify microbial taxa that were primarily found in host-derived samples. Using the microbial taxa that were primarily host-associated, we reanalyzed biodiversity metrics among the populations and compared the core gut microbiomes between ecotypes and across drainages. (3) Are there any microbes that could constitute symbionts that help their hosts survive under the extreme environmental conditions? We identified both shared and unique microbiome changes in evolutionarily-independent sulfide spring populations and isolated putative symbionts as microbial taxa associated with the colonization of sulfide springs. We first identified putative symbionts as taxa that were common in the guts of sulfide spring fishes but absent or rare in the guts of fish from nonsulfidic habitats and environmental samples. Since the gut microbiomes may also include transient or diet-derived microbes from the environment, potential symbionts were separately identified from taxa that were present in the environment but absent in fish from nonsulfidic environments. (4) Is there any evidence that hosts can control the composition of their gut microbiomes? If gut microbes play a role in fish adaptation to extreme environmental conditions, hosts should be able to bias the composition of their gut microbiomes. We therefore predicted that fish from sulfidic and nonsulfidic habitats would have different microbiomes when exposed to shared environmental conditions. To test

this, we compared microbiomes of fish that have been raised in a common-garden environment in the laboratory for multiple generations.

Methods

Sample collection

Sampling occurred across four river drainages in the Río Grijalva basin in Mexico: Tacotalpa, Ixtapangajoya, Pichucalco, and Puyacatengo (Appendix Figure C-1). Each river drainage is inhabited by a pair of sulfide-adapted and sulfide-intolerant populations of *Poecilia mexicana* species complex. Individuals (N=6 per population) were collected using a seine (2 m × 5 m; 3 mm mesh size) and immediately euthanized by pithing. Each individual's intestine was removed using sterilized dissection tools and stored in 96% ethanol. *Poecilia mexicana* do not have a separate stomach and intestine, so the intestine comprises the full length of the digestive tract (Tobler et al. 2015). Between each dissection, gloves were changed, and all dissecting tools and equipment were sterilized. At each site, 3–4 water samples (up to 1-L each, except in cases where the filters became fully clogged and stopped water flow before reaching the full liter) were collected from different points along the stream at about 30 cm below the surface. The water was filtered through sterile 47-mm diameter filter papers with a pore size of 0.45 µm (Omicron Scientific) using a 47 mm magnetic filter funnel (Pall Laboratory) attached to a filtration flask and a PVac portable vacuum pump (Argos Technologies). Sediment samples were also collected from each site (2–3 per site) by scooping the stream sediment into sterile 15-mL falcon tubes with 96% ethanol. Environmental samples were collected in areas occupied by fish. Upon collection, samples were stored at room temperature for about two weeks and at –80 °C upon

returning to the laboratory until DNA was extracted. The field location coordinates and number of samples collected from each site can be found in Table 3-1.

In the laboratory, samples were collected from *P. mexicana* that were reared in shared environmental conditions and fed the same diet for several generations. Whole intestines (N=5 per population) of sulfidic and nonsulfidic *P. mexicana* from the Tacotalpa lineage were dissected in the laboratory, preserved in ethanol, and stored at –80 °C, following the same procedures as in the field.

DNA extraction, amplification, and sequencing

Genomic DNA was extracted from whole intestines from the field and laboratory using a DNeasy PowerSoil kit, following the manufacturer's protocol with the following modifications: Whole intestines were transferred from the ethanol tubes into the kit's PowerBead tubes before proceeding with the extraction protocol. A FastPrep-24 bead beater (MP Biomedicals) was used for tissue homogenization at a speed of 4 m/s for 20 s, and all centrifugation steps were performed for 1 min. After adding 50 µL of the elution buffer at the final step of the protocol, samples were left to incubate at room temperature for 5 min before centrifugation.

We then extracted DNA from the sediment and water samples using a DNeasy PowerSoil Pro kit. For the water samples, the filter paper was smashed and torn into small pieces with sterilized tweezers, and then the samples were centrifuged at 15,000 g for 10 min to remove the ethanol. The ethanol was pipetted off the pellet, and then the pellets were washed with sterile PBS and centrifuged again at 15,000 g for 10 min. The PBS was removed, and solution CD1 from the first step of the PowerSoil Pro kit (lysis buffer) was added to the pellet and transferred into the PowerBead Pro tubes. All of the shredded filter paper was then added to the PowerBead

Pro tubes. For the sediment samples, 2 mL of the sediment suspended in ethanol was added to a clean 2 mL tube, washed with PBS, and transferred in solution CD1 to the PowerSoil Pro bead tubes following the same procedure as for the water samples. During the rest of the extraction protocol, the same modifications were made as before with the gut samples, except the water and sediment samples were incubated at 65 °C for 10 min after the addition of solution CD1, and they were homogenized in the FastPrep-24 bead beater (MP Biomedicals) for two 20-s intervals at 4 m/s to increase DNA yields. DNA extracts from all of the samples were stored at –20 °C, quantified using Quant-iT PicoGreen assay kits (Life Technologies, Grand Island, NY, USA), and standardized to a concentration of 5 ng/µL before starting the PCR amplification protocol.

We amplified the V3-V4 16S rRNA gene region from the microbial DNA extracts using the primers 341F (CCTACGGGNGGCWGCAG) and 805R (GACTACHVGGGTATCTAATCC) (Tapia-Paniagua et al. 2019), following the Illumina MiSeq System 16S Metagenomic Sequencing Library Preparation protocol (Illumina). The amplicon library was sequenced on an Illumina MiSeq at the Kansas State University Integrated Genomics Facility.

Processing reads, assigning taxonomy, and filtering taxa

The QIIME2 pipeline was used to analyze the reads (Caporaso et al. 2010; Kuczynski et al. 2012). We used the DADA2 plugin for QIIME2 (Callahan et al. 2016) to merge the forward and reverse paired-end reads, quality filter the samples, and call amplicon sequence variants (ASVs). The same trimming and truncating parameters were used for both sequencing runs, and the reads from both sequencing runs were merged together before proceeding to subsequent analyses. The merging and filtering statistics for all samples can be found in Appendix Table C-1, and the

number of reads after merging and filtering can be found in Appendix Table C-2. To account for variation in sequencing depth, we rarefied the gut and environment samples from the field to 10,405 reads, and the gut samples from the laboratory were rarefied to 18,834 reads. Rarefaction curves of the samples were generated in QIIME2 to determine a rarefaction depth that maximized the number of samples retained while capturing most of the diversity within the samples (Appendix Figure C-2). In total, 45 gut samples and 41 water and sediment samples from the field remained in our analyses after rarefying, and all 10 gut samples from the laboratory were retained. A naïve Bayesian classifier trained on the V3-V4 16S rRNA gene region based on the SILVA database (version 138) was used to assign taxonomy. After taxonomy was assigned, all analyses were conducted in R (version 4.2.0). The microbiome datasets generated from QIIME2 were imported into R using the R package *phyloseq* (McMurdie and Holmes 2013). We filtered the data to remove sequences that were classified as Archaea, mitochondria, and chloroplasts. Any unidentified bacterial sequences were also removed, and we further filtered out any taxa that could not be classified to at least the family level. We finally removed any taxa that had fewer than 3 reads across all samples to filter out singletons and doubletons from the dataset. Analyses and graphing in R were conducted using the R packages *phyloseq* (McMurdie and Holmes 2013) and *ggplot2* (Wickham 2011).

Analyzing commonalities and differences in gut microbiomes between sulfidic and nonsulfidic ecotypes and across river drainages

The first aim of our study was to investigate the commonalities and differences in the gut microbiomes of *P. mexicana* from sulfidic and nonsulfidic habitats and across river drainages. Since there were four drainages, each with a sulfidic and nonsulfidic population of *P. mexicana*,

our general approach was to use different metrics of diversity to analyze the microbiome variation in two ways: 1) between sulfidic and nonsulfidic populations overall and 2) between sulfidic and nonsulfidic populations separately in each drainage, allowing us to capture potential drainage-specific differences between the population pairs. Specifically, we compared the microbial community compositions between samples using Bray-Curtis dissimilarity distances and visualized the microbiome similarity using non-metric multi-dimensional (NMDS) scaling with the *ordinate* function of the *phyloseq* package (McMurdie and Holmes 2013). We tested for significant variation among gut microbiomes associated with drainage, H_2S (present or absent), and their interaction using permutational multivariate analysis of variance (PERMANOVA) with the *adonis* function in the R package *vegan*, and *P*-values were based on 9999 permutations (Oksanen et al. 2020). To compare within-sample microbial diversity across the drainages and populations, we calculated Shannon's diversity index, Pielou evenness, Chao1 richness, and relative dominance, and we visualized patterns of these alpha diversity metrics using boxplots.

In addition to comparing patterns of diversity among the gut microbiome samples, we identified commonalities and differences among the gut microbiomes by extracting core microbial taxa in each lineage and testing how the core microbes overlapped between the populations. Microbes were considered core microbial taxa of a population if they were detected in at least half of a population's samples (O'Brien et al. 2021; Custer et al. 2023). The same core microbiome assignment threshold was used to identify the overlap in core microbial taxa for just the sulfidic populations and just the nonsulfidic populations separately, and all of the core microbiome intersections were visualized in a combined Venn diagram.

Comparing gut microbiomes to environmental microbiomes and analyzing host-specific microbial communities

The second aim of this study was to understand how *P. mexicana* gut microbiomes compare to the microbiomes of their external environments, and to identify microbes that are primarily host-associated to distinguish them from potentially transient microbes that may be of dietary origin. To compare host and environmental microbiomes, we tested for significant variation driven by sample type (gut vs. environmental), H_2S, and their interactions using a PERMANOVA based on Bray-Curtis dissimilarity, as described above. We visualized the microbiome variation of all the field samples in an NMDS plot, using the *ordinate* function in the R package *phyloseq* (McMurdie and Holmes 2013).

Primarily host-associated microbes were isolated by filtering the gut microbiomes, including only taxa that are absent or rare in the environmental samples. Specifically, we removed any taxa from the gut dataset that had fewer than 4 reads in the matching environmental samples. To investigate patterns of gut microbiome variation with the environmental microbes removed, we repeated the diversity and core microbiome analyses detailed above.

Identifying potentially adaptive host-associated microbes in sulfide springs

The third aim of this study was to identify host-associated microbes in the sulfidic fish guts that might help their hosts survive in the extreme environmental conditions. We therefore identified putative symbionts as the microbial taxa that are associated with adaptation to sulfidic environments. We conducted these analyses separately for the whole gut microbiome and the primarily host-associated microbiome (i.e., excluding environmental microbes). These analyses provide complementary insights about the origin of potential symbionts, because microbes

contributing to host adaptation may be of dietary origin or they may be specialized symbionts that are restricted to—or at least much more common in—fish intestinal tracts. To identify microbes of interest, we identified the top microbial taxa driving the differences between the sulfidic and nonsulfidic populations using the top 30 coefficients associated with H_2S from the PERMANOVA of microbiome variation described above. The top taxa associated with sulfide spring colonization were identified for the combined gut microbiome dataset (all sulfidic vs. all nonsulfidic populations) and separately for each population pair of evolutionarily independent lineages.

Testing for evidence of host regulation of the microbiome

The fourth aim of this study was to investigate whether hosts can control the composition of their gut microbial communities. If host-microbiome adaptation has occurred, we would expect that genetic divergence between host populations is accompanied by divergence in host microbiome assembly. We therefore compared the gut microbiomes between a pair of sulfidic and nonsulfidic lineages (from the Tacotalpa drainage) that were raised in a common garden environment over several generations. To test whether sulfidic and nonsulfidic *P. mexicana* have different microbiomes from each other in the laboratory, we used permutational multivariate analysis of variance (PERMANOVA) based on Bray-Curtis dissimilarity with the *adonis* function in the R package *vegan* to test for between-sample microbiome diversity driven by ecotype (Oksanen et al. 2020). We visualized the gut microbiome variation of the laboratory-reared *P. mexicana* using an NMDS plot.

Results

What are the commonalities and differences in the gut microbiomes of P. mexicana populations? After filtering, there were 764 taxa in the fish gut microbiome samples. NMDS indicated clear separation between fish microbiomes from sulfidic and nonsulfidic habitats along axis 1, and microbiomes from sulfidic fish appeared to be much more variable than those from nonsulfidic fish (Figure 3-1A). The microbiome differences between sulfidic and nonsulfidic *P. mexicana* was also reflected in a significant H_2S term in the PERMANOVA (*P*-value < 0.0001). Furthermore, we found significant differences in microbiome composition among drainages (PERMANOVA: *P*-value < 0.0001) and in the interaction between drainages and H_2S (PERMANOVA: *P*-value < 0.0001), indicating that there was variation across the drainages in how sulfidic and nonsulfidic gut microbiomes differed from each other. Based on several alpha diversity metrics, sulfidic *P. mexicana* gut microbiomes had lower diversity (Shannon's diversity index), lower richness (Chao1 index), lower evenness (Pielou's evenness index), and higher relative dominance than nonsulfidic *P. mexicana* gut microbiomes across all four drainages (Appendix Figure C-3).

Identifying the core gut microbiomes from each population pair revealed 76–111 taxa that were consistently detected within drainages (Figure 3-2A). Comparing the core microbiomes among populations, we found that 41.4% of the core microbes were shared among drainages (46 shared microbes) (Figure 3-2A). These core microbes included several taxa that have been identified as members of other fish gut microbiomes, such as *Aeromonas*, *Pseudomonas*, *Plesiomonas*, *Rhodobacter*, and *Halomonas* (Ray et al. 2012; Sullam et al. 2012; Liu et al. 2016; Belkova et al. 2017; Ruiz et al. 2024; Appendix Table C-3). Proteobacteria was the largest phylum represented by the core microbial taxa (comprising 21 taxa total), like in other fish gut

microbiomes (Sullam et al. 2012; Liu et al. 2016; Eichmiller et al. 2016). We also identified the shared core microbiomes among just the sulfidic populations and just the nonsulfidic populations independently, and we similarly found overlap in the core microbiomes of *P. mexicana* among all the drainages (Figure 3-2A; Appendix Table C-3). There were 123–160 core microbial taxa for each nonsulfidic population, and 89 taxa were shared (55.6% overlap). In each sulfidic population, there were 29–78 core microbial taxa and 20 shared microbes among all of them (25.6% overlap). The core sulfidic gut microbes included taxa involved in environmental sulfur cycling, such as *Thiothrix*, *Sulfurospirillum,* and *Desulfocapsa* (Finster et al. 2013; Kruse et al. 2018; Ravin et al. 2022), microbes associated with fish digestive tracts, including *Lawsonia*, *Cetobacterium*, *Bacteroides*, and *Paludibacter* (Liu et al. 2016; Zhou et al. 2022; Qi et al. 2023; do Vale Pereira et al. 2024), and taxa implicated in symbioses with animals in sulfur-rich environments, such as *Sulfurovum* and *Spirochaeta* 2 (Blazejak et al. 2005; Breusing et al. 2020a). Two microbial taxa—*Aeromonas* and *Sulfurovum*—were present in all of the sulfidic fish gut samples. Overall, core microbiome analyses suggest that a relatively small subset of taxa consistently appear in guts of *P. mexicana*, but that subset of taxa is relatively consistent across habitat types and drainages.

How do gut microbiomes compare to environmental microbiomes, and how do patterns of diversity vary among microbial communities that are primarily host-associated?

After filtering, there were 1,089 microbial taxa among the gut and environmental samples. To understand how *P. mexicana* gut microbiomes compared to the environmental microbiomes, we visualized the microbiome variation of all the field samples in an NMDS plot based on Bray-Curtis dissimilarity and tested for microbiome variation using PERMANOVA. We found that the

sulfidic and nonsulfidic fish not only had different gut microbiomes from each other, but they also had distinct gut microbiomes from their respective environments (PERMANOVA: *P*-value < 0.0001; Appendix Figure C-4).

After removing the environmental microbes from the gut data set, 393 microbial taxa were retained, representing ~51% of the 764 taxa that were in the gut microbiomes before environmental correction. Sulfidic and nonsulfidic fish still had distinct gut microbiomes from each other, segregating along the first NMDS axis (Figure 3-1B), and the overall patterns of variation were remarkably consistent with variation in the whole microbiomes reported above. The primarily host-associated gut microbiomes had significant variation associated with drainage (PERMANOVA: *P*-value < 0.0001), H_2S (PERMANOVA: *P*-value < 0.0001), and their interaction (PERMANOVA: *P*-value < 0.0001). We also analyzed patterns of within-sample diversity across the lineages and found a similar pattern as in whole gut microbiomes; sulfidic *P. mexicana* gut microbiomes had lower diversity (Shannon's diversity index), lower richness (Chao1 index), lower evenness (Pielou's evenness index), and higher relative dominance than nonsulfidic *P. mexicana* (Appendix Figure C-5).

Comparing core microbiomes revealed 34–48 taxa per population, and we again found shared core microbes among drainages, both when fish from different habitat types were analyzed together and separately (Figure 3-2B; Appendix Table C-3). There were 22 microbes shared when analyzing the nonsulfidic and sulfidic core microbiomes together (45.8% overlap among the drainages). Additionally, the nonsulfidic populations had 54–78 core microbes each, and they shared 40 core microbes (51.2% overlap). There were 12–40 core taxa per sulfidic population, and the sulfidic populations shared 8 microbes in total (20.0% shared). The 8 shared microbial taxa were also among those previously identified as shared core microbes in whole gut

microbiomes of the sulfidic lineages (Appendix Table C-3). As previously mentioned, several of these primarily host-associated sulfidic core microbes have been associated with hosts in other systems, such as *Enterovibrio*, Halothiobacillaceae, *Lawsonia*, and *Spirochaeta* 2 (Blazejak et al. 2005; Lavy et al. 2018; Zhou et al. 2022; Wu et al. 2022). Overall, patterns of microbiome diversity and core microbiome composition were highly consistent when we analyzed whole gut microbiomes and only taxa that were primarily host-associated.

What are potentially adaptive host-associated microbes of sulfide spring fishes?

To identify host-associated microbes that are potential candidates for adaptive symbioses in sulfide spring environments, we extracted the top 30 coefficients of taxa driving differences between sulfidic and nonsulfidic groups from the PERMANOVA. If possible, these microbial taxa were identified down to genus level. Among all of the *P. mexicana* lineages, 14 taxa were positively associated with H_2S, while 16 were negatively associated with H_2S (Figure 3-3A). The microbes positively associated with H_2S represented 13 different families (Figure 3-3A), and all of them were also previously identified in the core microbiome of sulfidic fish with the exception of four microbial taxa (Vibrionaceae, *Planktothrix* NIVA.CYA 15, *Spirulina* CCC Snake P.Y.85, and *Mycoplasma*). Repeating this analysis for each drainage separately revealed that most of the top microbes associated with H_2S were the same across the drainages, except for a few drainage-specific microbial taxa (Appendix Figure C-6). Taxa that were positively associated with H_2S in at least three drainages included an uncultured species of Halothiobacillaceae, *Desulfobulbus*, *Sulfurovum*, and *Desulfocapsa*, with the latter two emerging among the top taxa in all four drainages (Appendix Figure C-6). Several of the top H_2S-associated microbes across all drainages (Figure 3-3A) have previously been identified in host microbiome analyses. For

example, bacteria in the family Vibrionaceae are known commensals of fish guts (Gallo et al. 2020), including members of the genus *Enterovibrio* that occur in the digestive tracts of a variety of healthy fish (Thompson et al. 2002; Fidopiastis et al. 2006; Wu et al. 2022; Kim et al. 2023). *Cetobacterium* and *Aeromonas* that were among the H_2S-associated gut microbes (Figure 3-3A) are also commonly identified members of healthy fish gut microbiomes (Shi et al. 2022; Duval et al. 2022; Diwan et al. 2023). Other microbial taxa, such as *Spirochaeta*, *Desulfobulbus*, *Sulfurovum*, and Halothiobacillaceae have been found in a variety of sulfur-rich environments (Hügler et al. 2010; Barton et al. 2014; Aranda et al. 2015; Hotaling et al. 2019). More importantly, these taxa have also been found in symbioses with animals that inhabit sulfidic environments, and at least some of them have genetic hallmarks that indicate their role as symbionts in hosts such as gutless oligochaetes (*Spirochaeta*) (Blazejak et al. 2005), hydrothermal vent shrimp (*Desulfobulbus*) (Jiang et al. 2020), hydrothermal vent snails (*Sulfurovum*) (Breusing et al. 2020a), and sponges (Halothiobacillaceae) (Lavy et al. 2018).

Since transient microbes from the environment may obscure host-associated microbes that could influence host functions, we further identified microbial taxa that were the top drivers of variation between sulfidic and nonsulfidic fish gut microbiomes after environmental correction (Figure 3-3B). In this case, there were 12 microbial taxa positively associated with H_2S and 18 microbial taxa negatively associated with H_2S (Figure 3-3B). The microbes associated with sulfide spring colonization after environmental correction belong to 11 families (Figure 3-3B), and these microbial taxa also included all 8 of the shared core microbes that were previously identified from the sulfidic lineages after excluding environmental microbes (Figure 3-2B; Appendix Table C-3). Furthermore, 8 taxa overlapped with the analysis of whole gut microbiomes (Figure 3-3A). This finding suggests that divergence in microbiomes between

sulfidic and nonsulfidic habitats is not merely caused by dietary uptake of environmental microbes characteristic of sulfide springs, but at least some of the divergence is associated with primarily host-associated microbes. In addition to taxa already found in the whole microbiome set, environmental correction revealed additional taxa that are differentiated between sulfidic and nonsulfidic populations: uncharacterized taxa of the families Erysipelotrichaceae (ZOR0006) Desulfosaricinaceae (Sva0081 sediment group), Prolixibacteraceae, and Geobacteraceae. Microbial taxa that were lost after environmental correction included *Planktothrix* NIVA.CYA 15, *Desulfobulbus*, *Spirulina* CCC Snake P.Y.85, *Aeromonas*, *Cetobacterium*, and *Sulfurovum*, indicating that these taxa may be of dietary origin. Repeating this analysis for each drainage independently indicated that the top microbial taxa associated with H_2S were mostly shared among all the sulfide spring populations (Appendix Figure C-7). Taxa that were positively associated with H_2S in at least three drainages included *Enterovibrio* and another, uncharacterized member of the Vibrionaceae, *Spirochaeta* 2, *Desulfocapsa*, and an uncultured species of Halothiobacillaceae. *Desulfocapsa* and the uncultured species of Halothiobacillaceae were top taxa in all four drainages (Appendix Figure C-7).

Is there any evidence that hosts can control the composition of their gut microbiomes?

If evolution in host-microbiome interactions has occurred to facilitate host adaptation in sulfide springs, then there should be some level of host regulation of the microbiome. We tested whether hosts can bias their microbial communities by comparing sulfidic and nonsulfidic *P. mexicana* gut microbiomes in stocks that have been kept under standardized environmental conditions in the laboratory for multiple generations. Variation in gut microbiomes of sulfidic and nonsulfidic *P. mexicana* raised in a common-garden environment was visualized using an NMDS based on

Bray-Curtis dissimilarity (Figure 3-4), and the gut microbiomes appear to segregate by ecotype along the first axis of the NMDS. However, this difference between the sulfidic and nonsulfidic groups was not statistically significant (PERMANOVA: *P*-value > 0.05).

Discussion

We characterized the gut microbiomes of fish adapted to sulfidic environments and identified potentially adaptive microbial taxa associated with sulfide spring colonization in sulfide-adapted and sulfide-intolerant population pairs of *P. mexicana*. Core microbiomes of these fish contained only 76–111 taxa, but these core taxa were largely shared among populations across river drainages and habitat types. Most importantly, we found that sulfidic and nonsulfidic fish have distinct microbiomes from each other, both before and after removing environmental microbes from the gut microbiome data set. Our findings further suggest that independently evolved lineages of *P. mexicana* underwent convergent microbiome shifts upon adaptation to sulfide streams, including microbial taxa with documented symbiotic functions in other fishes and animal hosts that live in H_2S-rich environments. Hence, our results point toward microbes of interest for further investigation that may play a role in nutrition, sulfide adaptation, and hypoxia tolerance of sulfide spring fishes. Future research in this system that tests how the host genetic background shapes microbiomes and how host-associated microbial communities influence host functions, such as H_2S or hypoxia tolerance, is warranted for understanding if the microbiome plays a role in host adaptation.

Convergent shifts in microbiome composition in sulfide springs

Our finding that sulfidic fish from four independently evolved lineages share derived core microbes raises the question of whether patterns of microbiome convergence across the lineages are due to convergent evolution in host-microbiome interactions or due to similar environmental microbial communities of sulfide springs shaping the gut microbiomes. If evolution has occurred in host-microbiome interactions, we would expect divergence between sulfidic and nonsulfidic populations to correspond to differences in their microbiomes. Comparisons of diversity metrics revealed that sulfidic fish and nonsulfidic fish had distinct gut microbiomes from each other, as well as from their environments, supporting that the host digestive tract is a selective environment. Even after filtering environmental microbes from the gut microbiome dataset, sulfidic and nonsulfidic fish consistently were distinct from each other and had shared core microbes across drainages, suggesting there is an exclusively host-associated microbiome. Of course, microbes that are exclusively host-associated and follow host evolutionary patterns (in this case, reflecting host divergence along parallel ecological gradients) could also affiliate with the host because of host exposure to shared environmental conditions, rather than because of host regulation of the microbiome. Identifying host regulation of the microbiome as a reason for this pattern of host-microbiome co-diversification would require uncovering mechanisms of communication between the host microbial communities and host immune system, as has been identified in other animal extremophiles (Bang et al. 2018; Li et al. 2019; Ip et al. 2021). Our findings also support a role for the environment in these patterns, as the sulfidic and nonsulfidic environments themselves have distinct microbiomes from each other. As in other systems (Sullam et al. 2012; Small et al. 2023), it is likely that both evolutionary and environmental factors play a role in shaping the host-associated microbiomes, so disentangling them remains important for addressing questions about host-microbiome adaptation. In studies on fishes that

have aimed to disentangle evolutionary and environmental factors, the environment plays a major role in shaping host-associated microbiomes (Sullam et al. 2015; Sylvain et al. 2020; Sadeghi et al. 2023). Furthermore, fish gut microbiomes have been found to recapitulate the host's phylogeny in natural systems (Sylvain et al. 2020; Sadeghi et al. 2023), but there is limited research on host regulation of the microbiome in controlled environmental settings, although microbiome variation can be associated with the host genetic background (Sevellec et al. 2019; Riddle et al. 2023; Small et al. 2023). One approach to disentangling evolutionary and environmental factors is to control for environmental effects; hence, we investigated whether sulfidic and nonsulfidic *P. mexicana* had different microbiomes in captivity after several generations. From the NMDS visualizing *P. mexicana* gut microbiome variation in the laboratory, sulfidic and nonsulfidic lineages of *P. mexicana* appear to segregate along the first axis (Figure 3-4), but we did not find statistical significance for ecotype driving gut microbiome variation in the laboratory. It remains an open question to what extent hosts regulate their microbiomes, and additional microbiome comparisons between host lineages in a common-garden environment would help determine if sulfide-adapted and sulfide-intolerant lineages of *P. mexicana* have diverged in their recruitment of their microbiomes. At the host level, there is evidence for immune gene divergence between sulfidic and nonsulfidic populations (Greenway et al. 2023). The immune system is known to play a major role in the regulation of metazoan microbiomes (Clavel et al. 2017; Salinas and Magadán 2017), so it is possible that divergence in immune genes also results in differences in microbiomes between sulfidic and nonsulfidic populations.

Potentially adaptive host-associated microbes of sulfidic fish

Our results highlighted several microbial taxa of interest for future investigation as potentially adaptive symbionts of sulfide spring fishes. Some of these microbes have been found in other fish gut microbiomes and may play a role as nutritional symbionts in sulfidic *P. mexicana*. *Cetobacterium* and bacteria from the families Vibrionaceae (including *Enterovibrio*) and Erysipelotrichaceae (ZOR0006) were among the top microbial taxa associated with sulfide spring colonization and known to co-occur as beneficial members of gut microbiomes in other fishes (Zarrinpar et al. 2018; Gallo et al. 2020; Shi et al. 2022; Duval et al. 2022; Wu et al. 2022; Kim et al. 2023). Beneficial microbes in host digestive tracts may enhance digestion or nutrient assimilation by stabilizing the intestinal environment, functions that have previously been attributed to ZOR0006 (Zarrinpar et al. 2018; Wang et al. 2023). Gut microbes may also provide nutrients to the host or help metabolize components of the host's diet (Diwan et al. 2023), and *Cetobacterium* produces vitamin B-12 (Sugita et al. 1991; Tsuchiya et al. 2008; Qi et al. 2023) and helps with the degradation of proteins (Hao et al. 2017). Indeed, sulfide spring colonization of *P. mexicana* is associated with a trophic shift from protein-poor to protein-rich diets, with concomitant changes in gastrointestinal tract morphology (Tobler et al. 2015). Host-associated microbes in sulfidic fish may therefore improve the intestinal barrier and metabolize dietary compounds, potentially contributing to host function in the novel environment.

Other host-associated microbial taxa that were associated with sulfide spring colonization are closely related to microbes that have been implicated in H_2S adaptation and hypoxia tolerance in other animals. The sulfidic fish have gut microbes that are known to be involved in sulfur cycling, including taxa associated with hosts living in marine environments rich in H_2S. Chemoautotrophic microbes are associated with a diversity of H_2S-adapted animals, either growing within host tissues or externally in different species of nematodes, oligochaetes,

crustaceans, bivalves, and others (Dubilier et al. 1995; Blazejak et al. 2005; Bellec et al. 2020; Jiang et al. 2020; Osman and Weinnig 2022). Most of these host-microbe interactions have been described as nutritional symbioses in which the symbiont uses sulfur compounds as energy sources to fix inorganic carbon that the host can use for nutrition (Osman and Weinnig 2022). However, chemoautotrophic symbionts may also provide protection to the host from toxic effects H_2S because they rapidly oxidize H_2S to nontoxic sulfur compounds (Osman and Weinnig 2022). For example, *Spirochaeta* bacteria provide H_2S-adapted gutless oligochaetes with nutrients in deep-sea hydrothermal vents (Blazejak et al. 2005). In sulfidic fish investigated here, *Spirochaeta* was a top microbial taxon after environmental correction, and it may provide a nutritional benefit to the host as well. Furthermore, an uncultured bacterial genus in the family Halothiobacillaceae (class Gammaproteobacteria) is also prevalent in the sulfidic fish guts after environmental correction, and members of this sulfur-oxidizing family that are associated with sponges were found to have a reduced genome size compared to free-living Halothiobacillaceae, suggesting that likely play a role as an obligate symbiont (Lavy et al. 2018).

We also found microbes implicated in host-microbiome adaptation to sulfidic environments that were dominant members of the sulfidic fish microbiomes but were also detected in the environment, such as *Desulfobulbus* and *Sulfurovum*. Host-associated *Desulfobulbus* strains with reduced genomes have been identified in H_2S-adapted hydrothermal vent shrimp, with unique metabolic features differentiating them from the most closely related free-living *Desulfobulbus* and supporting their dependence on a host (Jiang et al. 2020). In addition, *Desulfobulbus* bacteria have been identified as part of the diverse microbial communities of the deep-sea tubeworm *Riftia pachyptila* (López-García et al. 2002) as well as snails, polychaetes, shrimp, and other crustaceans (Giovannelli et al. 2016; Lee et al. 2021). In

deep-sea hydrothermal vent snails, *Sulfurovum* is a gut symbiont that can be genetically and morphologically distinguished from free-living *Sulfurovum* (Breusing et al. 2020a). The *Sulfurovum* symbiont genome has genes involved in biofilm formation and the transport of carbohydrates and amino acids, indicating its potential involvement in maintaining stability of the host-microbe symbiotic relationship (via maintenance of the biofilm under stressful environmental conditions) and facilitating transfer of nutrients to the host (Li et al. 2020; Breusing et al. 2020a). In other animals from deep-sea hydrothermal vents, *Sulfurovum* bacteria are associated with the gut (shrimp) (Qi et al. 2022), body wall (polychaetes) (Giovannelli et al. 2016), and surface of the gills (shrimp and other crustaceans) (Tokuda et al. 2008; Chiu et al. 2022), and they are typically thought to be acquired from the environment (Wang et al. 2022). In addition to their potential ability to provide nutrients and energy to the host, *Sulfurovum* in these systems have been suggested to protect the host from H_2S toxicity by oxidizing sulfur at the tissues (Tokuda et al. 2008; Lee et al. 2021), but future studies are necessary for determining how these sulfur-oxidizing bacteria are providing protective or detoxification benefits to the host.

In addition to H_2S toxicity, hypoxia in sulfide streams is a major selective pressure that *P. mexicana* experience. As previously mentioned for its role in fish nutrition, *Cetobacterium* is a dominant member of the sulfidic fish gut microbiomes, and it is abundant in the microbiomes of other fish that cope with hypoxic conditions, including arapaima (Ramírez et al. 2018) and desert pupfish (Bhute et al. 2020). Exposure to chronic hypoxia causes a variety of physiological changes in aquatic animals, such as increased susceptibility to infections (Abdel-Tawwab et al. 2019), decreased protein digestibility (Zambonino-Infante et al. 2017), and disruption of the gut microbiome (Sun et al. 2020; Ma et al. 2021). Maintaining gut microbiome stability may therefore be especially important for the utilization of dietary protein and overall health of

sulfidic fish living in chronic hypoxia. In zebrafish, *Cetobacterium* has been shown to inhibit pathogen infection by stabilizing the gut microbiome through its vitamin B-12 production (Qi et al. 2023). *Poecilia mexicana* encounter pathogens in sulfide springs (Tobler et al. 2014; Riesch et al. 2020), and it is possible that microbiome-stabilizing members of the host microbial communities, such as *Cetobacterium*, could protect the host from infections in the intestine, especially since sulfidic fish may have increased susceptibility to pathogen infections due to their exposure to hypoxia and their lower gut microbiome diversity than nonsulfidic fish (Xiong et al. 2019). Approaches that can investigate the functional potential of host-associated microbes in sulfidic fish, such as shotgun metagenomics, and experiments testing the role of the microbiome in host functions will shed further light on whether host-associated microbial communities of sulfide spring fishes can be adaptive for the host.

Conclusions

Lineages of *P. mexicana* have adapted in several ways to tolerate the extreme conditions of sulfide springs, including in their physiology, behavior, and morphology (Tobler et al. 2009, 2015; Greenway et al. 2020). These adaptations are driven by physiochemical and ecological stressors, yet a previously unexplored component of the ecology of *P. mexicana* is their associated microbiomes and their exposure to novel microbial communities upon colonizing sulfide streams. In fishes and other vertebrates, host-associated microbiomes may play a role in the evolution of host traits because they can influence host functions (Diwan et al. 2023), and microbial communities respond more quickly (structurally, genetically, and functionally) than the host to environmental pressures, potentially facilitating host tolerance of stressors in novel environments (Bang et al. 2018). Our findings that *P. mexicana* have different microbiomes

between sulfidic and nonsulfidic populations and share commonalities in their microbes associated with sulfide spring colonization provide a foundation for addressing whether these microbiome changes are linked to the host's evolution in an extreme environment.

Tables

Table 3-1. Coordinate locations of all the sampling sites, as well as the number of samples (gut, water, and sediment) collected at each site. Numbers in parentheses indicate the final number of samples included in the analyses after rarefying the data.

Site	Coordinates (Latitude, longitude)	Hydrogen sulfide	*N* individuals sampled for gut microbiome analyses	*N* water samples	*N* sediment samples
Pichucalco river drainage					
Arroyo Rosita	17.485, -93.104	-	6	3	2
La Gloria	17.532, -93.015	+	6 (5)	3	3
Ixtapangajoya river drainage					
Tributary to Ixtapangajoya	17.510, -92.980	-	6	3	2 (1)
La Esperanza	17.511, -92.983	+	6 (5)	3	2
Puyacatengo river drainage					
Puyacatengo downstream	17.564, -92.912	-	6	3	2
La Lluvia	17.464, -92.896	+	6	3	2
Tacotalpa river drainage					
Arroyo Bonita	17.427, -92.757	-	6 (5)	4	2
El Azufre I	17.442, -92.775	+	6	3	2

Figures

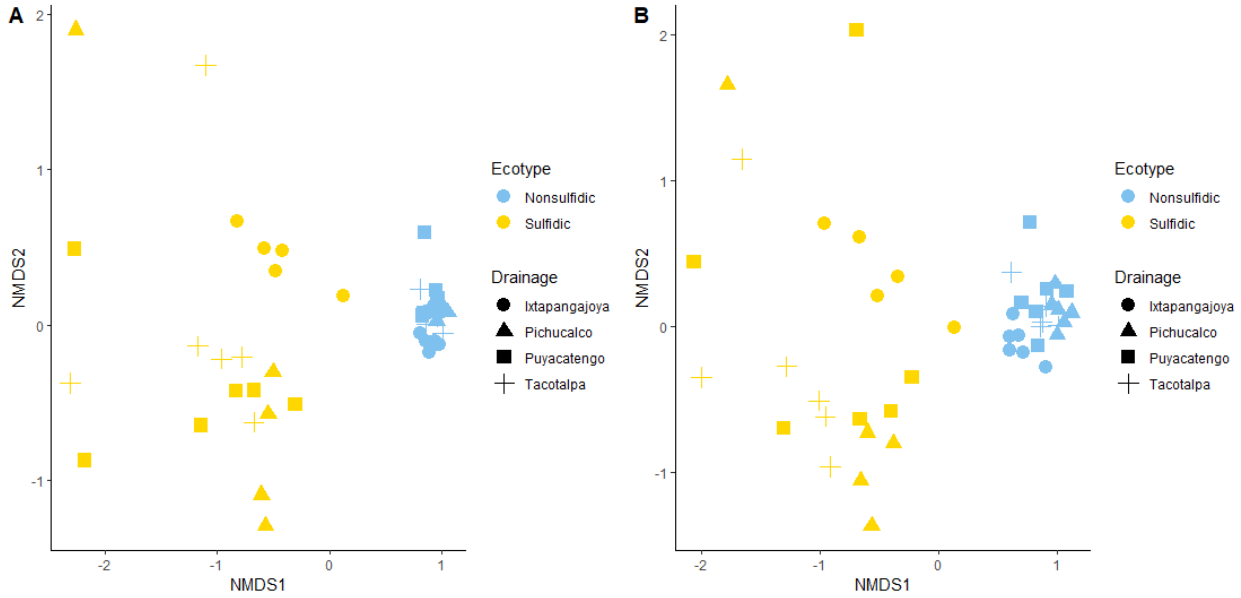


Figure 3-1. NMDS plot based on Bray-Curtis dissimilarity of sulfidic and nonsulfidic *P. mexicana* gut microbial communities A) before environmental correction and B) after environmental correction.

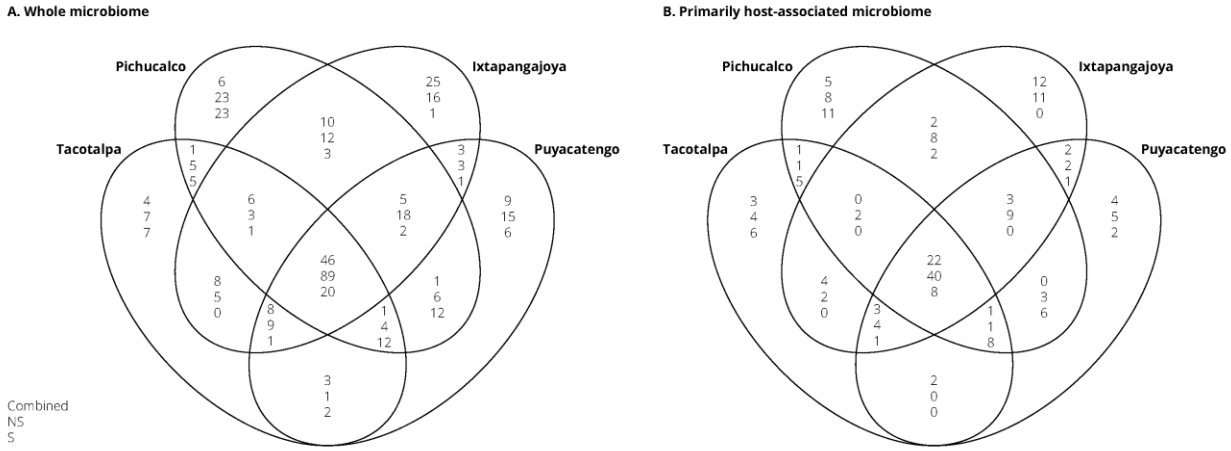


Figure 3-2. Venn diagram showing the shared and unique core microbes among all four drainages when comparing A) the whole microbiome (before environmental correction) and B) the primarily host-associated microbiome (after environmental correction). The numbers of core microbial taxa from the nonsulfidic and sulfidic microbiomes combined are the top numbers in the Venn diagram, while the numbers of core microbial taxa from just the nonsulfidic populations are the middle numbers, and the numbers of core microbial taxa from just the sulfidic populations are the bottom numbers.

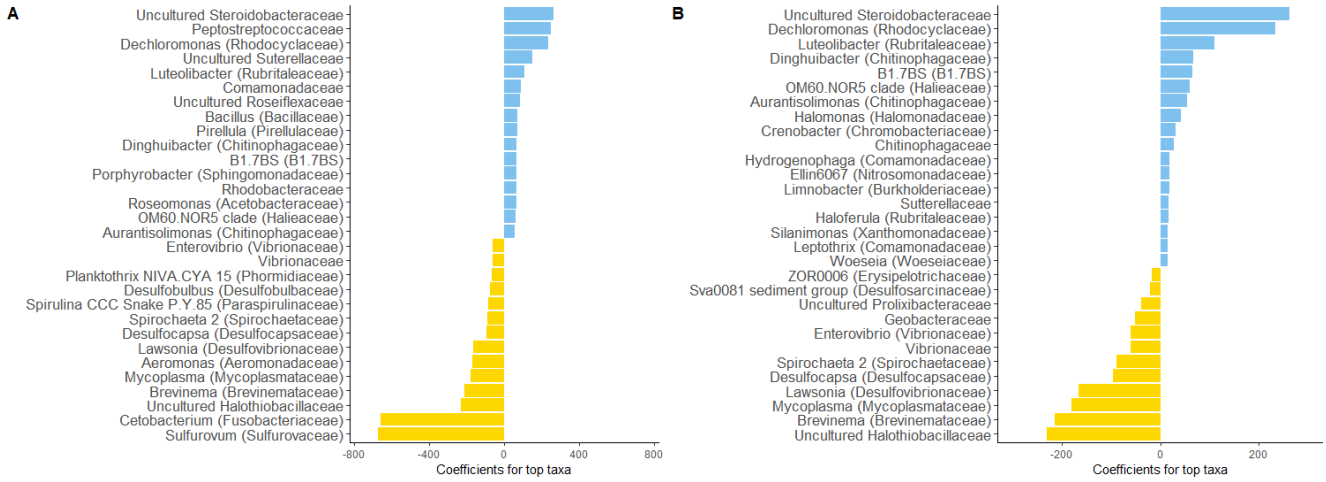


Figure 3-3. The top microbial taxa driving differences between sulfidic and nonsulfidic *P. mexicana* gut microbial community composition A) before environmental correction and B) after environmental correction. Taxa in nonsulfidic lineages are indicated by blue bars, and taxa in sulfidic lineages are indicated by yellow bars. Taxa are labeled to the genus level or the lowest identifiable taxonomic level, and family names are listed in parentheses next to each genus.

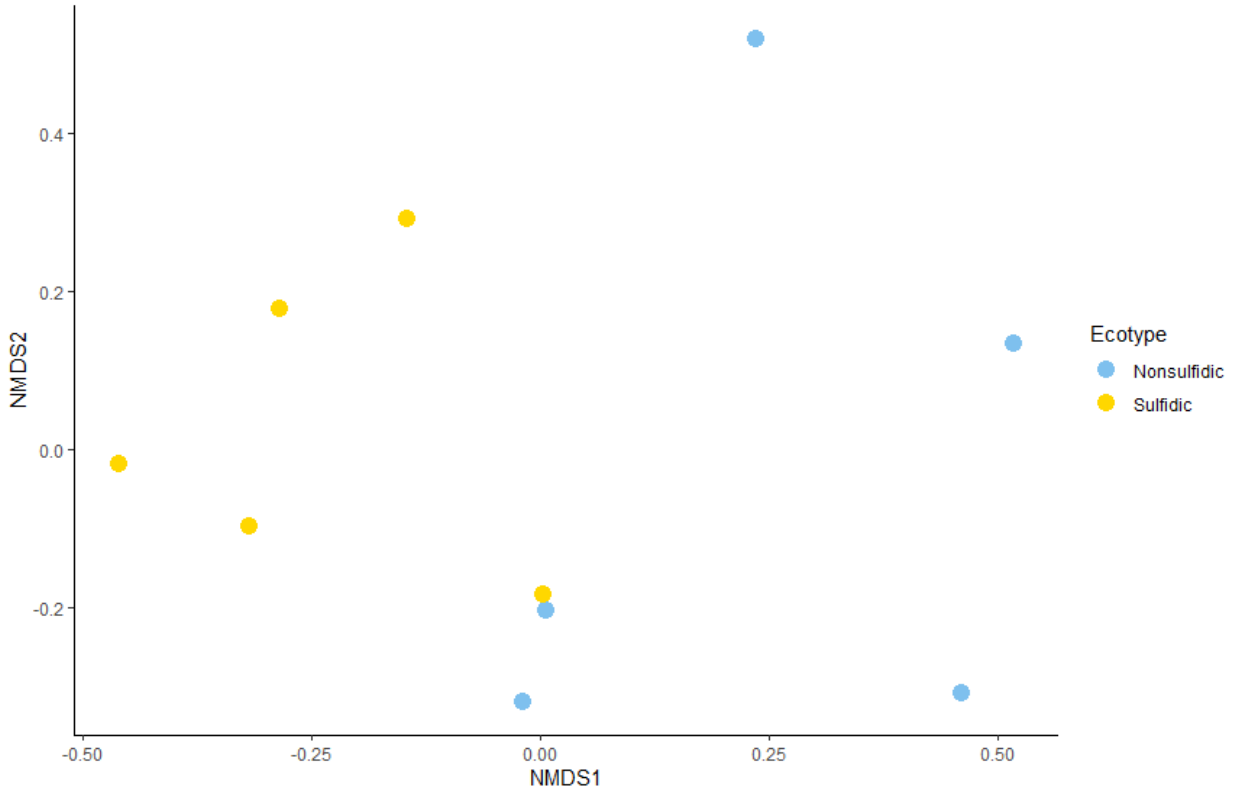


Figure 3-4. NMDS plot showing gut microbiome variation of sulfidic and nonsulfidic ecotypes of *P. mexicana* raised in a common-garden environment.

Chapter 4 - Comparative phylogenetic analysis of Poeciliid microbiomes: effects of evolutionary history and the environment

Elizabeth J. Wilson, Sonny T.M. Lee, Lydia H. Zeglin, Lenin Arias-Rodriguez, Michael Tobler

Abstract

Investigating how host-associated microbiomes are shaped by host evolutionary and environmental factors is important for understanding how hosts recruit their microbial communities and for understanding mechanisms of host-microbiome evolution. We investigated host-microbiome diversification among 23 populations (16 species) of livebearing fishes (Poeciliidae) throughout southern Mexico by characterizing gut and environmental microbiomes using 16S rRNA sequencing and then testing for correlations between the gut microbiome dissimilarity, host phylogenetic relationships, environmental microbiome dissimilarity, and environmental distance based on bioclimatic variables. We found that host-associated microbiomes were most strongly correlated with their environmental microbial communities, but there was not a significant correlation with environmental distance. We also found a significant correlation with the host phylogeny, indicating that the gut microbiome variation in poeciliids reflects the host evolutionary relationships. Furthermore, some of the poeciliid lineages in this region have colonized toxic environments rich in hydrogen sulfide, providing an opportunity to explore whether shared physiochemical and ecological selective forces drive parallel shifts in gut microbiome composition. Using phylogenetic comparative methods, we found evidence for strong convergence in the microbiomes of sulfidic fishes, and we also identified microbial taxa that exhibited convergent abundance shifts.

Introduction

Animal microbiomes influence the physiology, behavior, ecology, and life history of the host (McFall-Ngai et al. 2013). Due to their impact on host function, host-associated microbes can also assist the host in adapting to novel environmental conditions (Vrijenhoek 2013; Petersen et al. 2023). It has historically been challenging to investigate the role of the microbiome in host adaptation because selection in a novel environment acts on both the host and its microbiome, and the host and the microbiome mutually influence each other (Alberdi et al. 2016). Hence, disentangling how changes in the host genotype, its microbes, and their interactions impact host fitness and ultimately its evolution under different environmental conditions remains a challenge for the field (Alberdi et al. 2016; Petersen et al. 2023). Assembly of animal microbiomes is shaped by many factors, including the host's genetic background (which is a result of the host's evolutionary history) and the microbes present in the environment, both of which have been particularly studied for their role in the formation of the gut microbiome (Youngblut et al. 2019). There is evidence for long-term co-diversification of hosts and their microbiota or symbionts, especially in systems with strong vertical transmission, supporting the role of host evolutionary history in shaping the microbiome (Moran 2001; Gupta and Nair 2020; Li et al. 2022). In contrast, there is also evidence for strong environmental influences on the host-associated microbiomes, and host microbiomes can be reflective of the environmental pool of microbes (Cuellar-Gempeler and Leibold 2019; Zeng et al. 2020). Both host evolutionary and environmental factors likely play roles in shaping host-associated microbiomes, and for many systems it remains to be tested to what degree host microbiome assembly can be predicted from the host phylogenetic relationships, and to what degree host microbiome assembly can be predicted based on the pool of microbes available in the host's environment? Among fishes, gut

microbiomes were primarily thought to be influenced by environmental factors, such as diet and water microbial communities (Tarnecki et al. 2017; Kim et al. 2021), but more recent analyses suggest that the host's genetic background and evolutionary history can also contribute to gut microbiome composition (Smith et al. 2015; Riddle et al. 2023; Sadeghi et al. 2023). To what degree fish microbiome composition can be attributed to different ecological and evolutionary factors remains an important question, especially because comparative analyses among taxa remain scarce. Ultimately, insight into the eco-evolutionary dynamics shaping host-associated microbiomes is crucial for understanding the mechanisms involved in host-microbiome evolution and the role of the microbiome in host adaptation.

The livebearing fishes of the family Poeciliidae are an ecologically diverse group distributed throughout the Neotropics, and well-resolved phylogenies make this group an excellent model for comparative analyses (Culumber and Tobler 2017; Furness et al. 2019). Several species of this group inhabit springs, streams, and rivers of the Río de la Sierra region in southern Mexico (Tobler et al. 2018, 2021; Pease et al. 2023), and they vary along multiple axes of ecological function, including macro- and microhabitat use, foraging behavior and diet, as well as aspects of their life history (Miller et al. 2005). Most importantly, this assemblage includes distantly related species that occupy the same biotope and closely related species and even populations of the same species that occupy starkly different habitat types. The latter case is best exemplified in some springs with hypoxia and high levels of naturally occurring hydrogen sulfide (H_2S), a respiratory toxicant that causes high levels of mortality even at low concentrations (Tobler et al. 2016, 2018; Kelley et al. 2016). Poeciliid lineages in the genera *Gambusia*, *Poecilia*, *Pseudoxiphophorus*, and *Xiphophorus* have independently colonized and adapted to these extreme environments, frequently exhibiting convergent evolution of traits that

mediate survival in the presence of H_2S and hypoxia (Greenway et al. 2014, 2023). Sulfide spring occupants have profoundly different ecologies compared to conspecifics in adjacent nonsulfidic habitats—for example, in terms of their diet, microhabitat use, and exposure to natural enemies (Tobler et al. 2014, 2015; Lukas et al. 2021). In addition, sulfide springs have distinct environmental microbiomes, and at least part of their primary production comes from chemoautotrophic bacteria that are abundant in the water and sediment (Hotaling et al. 2019).

The close geographic distribution of different habitat types and multiple host species in this system is useful not only for studying patterns of host-microbiome diversification in an explicit ecological and evolutionary context, but the repeated invasions of host lineages into extreme environments allows for testing whether exposure to shared ecological stressors result in convergent microbiome shifts. The explicit selective forces—H_2S toxicity, hypoxia, and the dietary shift—found in sulfide springs could all mediate changes of host-associated microbes in sulfide spring fish, and beneficial symbioses in other systems point to the potential for an adaptive role of the microbiome in sulfide spring colonization. For example, fishes have microbial communities that benefit hosts through nutrient assimilation and support of immune function (Legrand et al. 2020E; Sehnal et al. 2021). Benefits provided by the gut microbiota may also improve host responses to hypoxia, as supported by other vertebrate microbiome studies, such as by mitigating the effects of microbiome disruption and pathogen susceptibility caused by hypoxia (Sundarraman et al. 2020; Qi et al. 2023; Morshed and Lee 2023) or by producing amino acids and vitamins involved in host hypoxia resistance pathways (Zhao et al. 2023). In other H_2S-rich environments, diverse host-microbe symbioses have been described, including chemoautotrophic bacteria that can provide nutrition or H_2S detoxification to their invertebrate hosts (Powell and Somero 1986; Vrijenhoek 2013). Understanding whether patterns of

microbiome convergence emerge from repeated colonization of extreme environments therefore also provides an opportunity to identify candidate microbes for future investigation as symbionts.

In this study, we analyzed host-associated microbiomes within explicit ecological and phylogenetic frameworks to understand factors underlying microbiome assembly. We sampled 23 populations of livebearing fishes (in 16 species) across 13 sites throughout the Río de la Sierra region, and we characterized their intestinal microbiomes using 16S rRNA amplicon sequencing. Along with the host-associated microbiomes, we sampled water and sediment microbiomes from each site to address the following questions: 1) How are host-associated microbiomes shaped by the host's evolutionary history, by environmental factors, and by the environmental microbiomes? We investigated microbiome variation among the host and environmental samples, and we tested whether host microbiome dissimilarity between lineages was correlated with host phylogenetic distance, environmental distance (based on bioclimatic variables), and environmental microbiome dissimilarity. 2) Is there evidence for microbiome convergence among extremophile lineages of livebearing fishes? We used phylogenetic comparative methods to test for evidence of convergence among the microbiomes of all the sulfidic lineages, and we identified microbes showing convergent abundance shifts among the sulfidic populations relative to the nonsulfidic populations.

Methods

Sample collection

We sampled the intestinal microbiomes of 23 populations of livebearing fishes, including 16 species in 10 genera, throughout the Río de la Sierra region and the Río Grijalva floodplains in the states of Chiapas and Tabasco, Mexico (Figure 4-1) (Pease et al. 2023). Of the 23

populations sampled, 8 represented lineages inhabiting sulfide springs. The fish were caught by seining (2 m × 5 m seine size; 3 mm mesh size), and depending on availability, we collected up to 7 individuals from each population, with most populations having 6 individuals (N=132 individuals total) (Table 4-1). Each fish was euthanized by pithing and then dissected to remove the whole intestine, which was preserved in 96% ethanol. The dissecting tools were sterilized, and gloves were changed between individuals. At each site, we also collected sediment and water samples to characterize the environmental microbial communities. At each site, 2–3 sediment samples were scooped into sterile 15-mL conical tubes filled with 96% ethanol (N=27 sediment samples total), and 3–4 water samples were collected by filtering 1 L of water through sterile filter papers (Omicron Scientific; 47-mm diameter; 0.45 μm pore size) and stored in 96% ethanol (N=40 water samples total). The water filtration system consisted of a PVac portable vacuum pump (Argos Technologies), a 1-L filtration flask, and a 47 mm magnetic filter funnel (Pall Laboratory). The water samples were collected from each stream at about 30 cm below the surface of the water. In cases when the filter papers clogged before passing the full 1-L sample through the filtration system, the pump was stopped, and the filter paper was collected. During sampling in the field, the samples were kept at room temperature for about two weeks and then were stored at –80 °C in the laboratory.

DNA extraction and 16S rRNA amplicon sequencing

We extracted genomic DNA from the intestine samples using a DNeasy PowerSoil kit, with a few modifications to the protocol that was provided by the manufacturer. Specifically, we began the extractions by moving the whole intestines from ethanol to the PowerBead tubes provided in the kit. At the tissue homogenization step, we used a FastPrep-24 bead beater (MP Biomedicals)

at 4 m/s for 20 s. Throughout the protocol, all centrifugation intervals were run for 1 min. At the final elution step, 50 µL of the elution buffer was added to each spin column and allowed to incubate for 5 min at room temperature before centrifuging the samples and collecting the DNA.

We similarly made modifications to the DNA extraction protocol for water and sediment samples to increase DNA yields. DNA was extracted from the water and sediment samples using a DNeasy PowerSoil Pro kit. We used sterilized tweezers to finely tear and smash the filter paper, and the filter paper was centrifuged at 15,000 x g for 10 minutes to form a pellet. The ethanol was removed, and the filter paper was washed once with sterile PBS and centrifuged again with the same centrifugation parameters to remove the PBS. Lysis buffer (solution CD1 from the extraction kit) was added to the filter paper as indicated in the manufacturer's protocol, and the entire sample with lysis buffer was moved to the PowerBead Pro tube before continuing with the rest of the extraction. 2 mL of the sediment preserved in ethanol was also washed with sterile PBS as before with the water samples, and solution CD1 was added to the sediment in a PowerBead Pro tube for each sample. We further implemented the same modifications to the water and sediment extraction protocol as we used for the intestine DNA extractions, with the addition of an incubation step (65 °C for 10 min) to the samples suspended in solution CD1, and we also homogenized the samples twice for 20 s each at 4 m/s in a FastPrep-24 bead beater (MP Biomedicals) after the incubation step. The final DNA extracts were stored at –20 °C. To quantify DNA yields, a Quant-iT PicoGreen assay kit (Life Technologies, Grand Island, NY, USA) was used, and all samples were diluted to a concentration of 5 ng/µL.

To amplify the bacterial DNA, we used the primers 341F CCTACGGGNGGCWGCAG and 805R GACTACHVGGGTATCTAATCC (Tapia-Paniagua et al. 2019), which target the V3-V4 region of the 16S rRNA gene. We followed the Illumina MiSeq System 16S Metagenomic

Sequencing Library Preparation protocol (Illumina) for the PCR amplification, and we sequenced the sample library on an Illumina MiSeq at the Kansas State University Integrated Genomics Facility. There were three sequencing runs in total—two runs for the gut microbiome samples, and one run for the environmental microbiome samples.

Processing the sequencing reads

We used QIIME 2 to process the sequencing reads obtained from Illumina sequencing (Caporaso et al. 2010; Kuczynski et al. 2012; Bolyen et al. 2019). Amplicon sequence variants were generated by trimming and merging the forward and reverse reads and quality filtering the samples using the DADA2 pipeline (Callahan et al. 2016), and all of the sequencing runs were pooled together. The merging and filtering statistics are in Appendix Table D-1, and the read counts after trimming and merging are in Appendix Table D-2. We rarefied all of the sequences to 4,500 reads, based on rarefaction curves generated in QIIME2 (Appendix Figure D-1). Four samples had fewer than 4,500 reads and were not included, leaving 195 total samples in the analysis (Table 4-1). We assigned taxonomy using a naïve Bayesian classifier trained on the V3-V4 16S rRNA gene region with the SILVA database (version 138). After generating ASV and taxonomy tables in QIIME2, all of the proceeding analyses were conducted in R (version 4.2.0). We used the R package *phyloseq* to import the data and conduct microbiome analyses (McMurdie and Holmes 2013). Archaeal, mitochondrial, chloroplast, and unidentified sequences were removed from the dataset. We finally filtered the dataset to only include microbial taxa that had at least three reads in at least 5% of the samples.

Analyzing variation in microbiomes

To better understand factors that shape host microbiome assembly in poeciliid guts, we first explored how microbiome variation is influenced by sample type (gut vs. environment) and habitat type (nonsulfidic vs. sulfidic). Microbiome variation of all the samples was visualized using an NMDS plot based on Bray-Curtis dissimilarity, and we used permutational multivariate analysis of variance (PERMANOVA) with the *adonis* function in the R package *vegan* (Oksanen et al. 2020) to test how sample type, habitat type, and their interaction impacted microbiome composition.

We also analyzed variation in gut microbiome composition in an explicit host phylogenetic context by first calculating the mean abundance of microbial taxa (mean number of rarified reads) for each of the 23 lineages in our dataset. Gut microbiome variation was visualized with an NMDS plot based on Bray-Curtis dissimilarity, and we superimposed the host phylogeny connecting points in the NMDS in a matter that reflects host evolutionary relationships using the R package *ggphylomorpho*.

Testing for effects of evolutionary history and environment

To disentangle how the host evolutionary history, variation in bioclimatic environmental factors, and the composition of environmental microbiomes shaped host-associated microbiomes, we used a distance-based approach that identified correlations in the similarity among these metrics across all lineage pairs in our data set. To generate a matrix of dependent variables, we calculated pairwise Bray-Curtis dissimilarity of gut microbiome composition between all lineages using the *vegdist* function in the R package *vegan* (Dixon 2003). In addition, we calculated matrices for three independent variables: (1) Pairwise phylogenetic distances between host lineages was calculated using previously-established maximum likelihood tree of lineages in

the family Poeciliidae, which was based on the consensus sequences of 167 genes (Barts et al. 2018). The evolutionary relationships predicted by the maximum likelihood tree aligned with previously-established phylogenies of the family Poeciliidae (Hrbek et al. 2007; Pollux et al. 2014; Palacios et al. 2016). Phylogenetic distances between all possible pairs of lineages were calculated using the maximum likelihood tree with the *cophenetic* function in the R base package *stats*. (2) Pairwise dissimilarity in bioclimatic environmental factors was calculated based on climate data associated with the coordinate locations of the sampling sites (Dixon 2003). To do so, we extracted multiple bioclimatic variables, encompassing seasonal and long-term temperature and precipitation measurements from the WorldClim database using the *getData* function in the R package *raster* at a resolution of 0.5 minutes of a degree (Fick and Hijmans 2017; Hijmans et al. 2023). The *vegidst* function was then used to calculate Euclidean environmental distances between all lineage pairs. (3) Pairwise dissimilarity in environmental microbiomes was calculated in the same manner as for the corresponding gut microbiomes.

Applying these distance-based approaches resulted in four matrices, and we used partial Mantel tests (with the *mantel* function of the *ecodist* R package; Goslee and Urban 2007) to test for correlations between the host microbiome dissimilarity and the three predictor variables (host phylogenetic distance, dissimilarity in bioclimatic environmental factors, and dissimilarity in environmental microbiomes). Data were visualized with scatterplots using *ggplot2* (Wickham 2011).

Whole-microbiomes vs. primarily host-associated microbiomes

Fish gut microbiomes contain both transient and resident microbes, so a portion of the gut microbiome is expected to come from the external environment. Especially generalist

detritivores, like many species included in this study, may contain large amounts of microbes that are ingested as part of their diet. To ascertain that patterns of variation in gut microbiome composition were not primarily driven by microbes that are common in the environment and frequently ingested by the fish, we contrasted patterns of variation in whole gut microbiomes with patterns of variation in primarily host-associated microbiomes. These primarily host-associated microbiomes were approximated by removing any taxa that were detected in the environment from the gut microbiome data set. Specifically, only taxa in the gut microbiomes that had fewer than 2 reads in the environmental microbiomes were retained. Filtered primarily host-associated microbiome data was then used to repeat analyzes described above (NMDS and partial Mantel tests).

Testing for convergence in host-associated microbiomes using phylogenetic comparative methods

Multiple lineages of the livebearing fishes that we sampled have colonized toxic springs rich in H_2S, raising questions about whether repeated invasions into extreme environments drive convergent changes in host-microbiome associations. We therefore used a phylogenetic comparative approach to quantifying multiple metrics of convergence in multivariate traits (Stayton 2015), as implemented in the R package *convevol* (Stayton 2015; Brightly and Stayton 2023). NMDS scores along two axes were used for the calculation of all convergence metrics. We first quantified convergence by counting the number of times that lineages have invaded a region of multivariate space that is defined by lineages hypothesized to be convergent. To do so, we used the convnum function and designated lineages collected in sulfide springs as hypothesized convergent tips. We assessed the significance of convergence using 10,000

simulations of the convnum metric. In addition, we calculated convergence metrics C_1-C_4 using the calcConv function. C_1 is a relative estimate of the degree of convergence, ranging from 0 (no convergence) to 1 (complete convergence where lineages are indistinguishable). C_2 is an absolute estimate of the degree of convergence, with larger values representing higher degrees of convergence. C_3 is the proportion of evolution in convergent lineages that is attributable to convergence, and C_4 is the proportion of convergence within clades that contain the convergent taxa (Stayton 2015). Significance was calculated using 10,000 simulations with the convSig function.

To identify microbial taxa that show convergent abundance shifts upon host colonization of sulfide springs, we used individual-level microbial abundance data as dependent variables in Expression Variance and Evolution (EVE) models (Rohlfs et al. 2014; Rohlfs and Nielsen 2015). Although originally developed for the analysis of gene expression data, EVE can be used to analyze the abundance of other metrics (including microbial taxa) for which replicated data within lineages is available. EVE models implement an extended Ornstein-Uhlenbeck process that incorporates within-species abundance variance to test for branch-specific shifts in abundance by comparing likelihoods that an abundance parameter (θ_i) for a given taxon is shared between two groups of lineages versus θ_i for that taxon being significantly different between the groups (Rohlfs and Nielsen 2015). We designated branches associated with lineages from sulfidic habitats as one group and those associated with nonsulfidic lineages as another group and then contrasted θ_i for each taxon between these two groups. For each microbial taxon, we employed a likelihood ratio test (LRT_θ) contrasting the null hypothesis ($\theta_i^{sulfidic} = \theta_i^{nonsulfidic}$) to the alternative hypothesis ($\theta_i^{sulfidic} \neq \theta_i^{nonsulfidic}$) using a X^2 distribution to assess statistical significance (Rohlfs and Nielsen 2015). To account for multiple testing, we calculated FDR

adjusted *P*-values using the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995). For both convergence tests (*convevol* and EVE analyses), we again repeated these analyses with the filtered primarily host-associated microbiome data.

Results

Variation in environmental and host-associated microbiomes

Prior to filtering, there were 1,857 microbial taxa in the gut and environmental microbiomes. After filtering, we retained 389 taxa in the gut microbiome samples and 414 taxa in the environmental microbiome samples. Poeciliid gut microbiomes were clearly distinct from the environmental microbiomes of their habitats, segregating along the first NMDS axis (Figure 4-2). In addition, both environmental and host-associated microbiomes differed between nonsulfidic and sulfidic habitats, segregating along the second NMDS axis (Figure 4-2). PERMANOVA confirmed these patterns, with habitat type (presence or absence of H_2S; $P < 0.0001$), sample type (environment vs. gut; $P < 0.0001$), and their interaction ($P < 0.0001$) all explaining significant variation in microbiome composition.

To analyze variation in gut microbiome composition among species in a phylogenetic context, we averaged relative abundances of different microbe taxa for each host species. Superimposing the host evolutionary relationships onto the pattern of gut microbiome variation revealed several interesting patterns: (1) While there is some signal of more closely related species clustering in the NMDS plot, branches of the phylogeny frequently crisscross throughout microbiome space, indicating that gut microbiome similarity among host species is not only shaped by phylogenetic relatedness, but also homoplasy driven by either convergence or reversals. (2) Microbiome variation along NMDS axis 1 appeared to be correlated with variation

in host diet, ranging from species with protein-rich diets (more negative axis scores) to species with more protein-poor diets (more positive axis scores). For example, *Belonesox belizanus* is the only piscivorous species in the family Poeciliidae (Ferry-Graham et al. 2010), and species of the genera *Gambusia*, *Heterphallus*, *Priapella*, and *Pseudoxiphophorus* include substantial amounts of invertebrate prey into their diet (Gibb et al. 2008; Riesch et al. 2011, 2012). In contrast, species of the genera *Poecilia* and *Xiphophorus* primarily feed on algae and detritus (Winemiller 1993; Kramer and Bryant 1995; Gibb et al. 2008; Tobler et al. 2015). (3) Species collected from sulfide spring habitats clustered separately (high scores along axes 1 and 2) from species in nonsulfidic habitats. While the lack of detailed dietary data for all populations restricted formal analyses about the role of trophic ecology shaping gut microbiomes, we used phylogenetic comparative analyses to test for the effects of evolutionary relatedness and environmental factors in gut microbiome composition, and we tested whether clustering of populations from sulfidic habitats represents a case of convergent evolution.

Effects of evolutionary history and environment

To quantitatively analyze the roles of evolutionary history and ecological factors on host microbiome composition, we calculated pair-wise dissimilarity of microbiomes among all host populations and asked whether it can be predicted based on host phylogenetic distance, environmental dissimilarity, or the dissimilarity of environmental microbiomes. Partial Mantel tests revealed that the strongest predictor of host microbiome dissimilarity was the dissimilarity of the corresponding environmental microbiomes (Mantel correlation: 0.271; $P = 0.0008$; Figure 4-4A). Hence, species that were exposed to similar microbes in their environments also tended to have more similar microbial communities in their guts. In addition, gut microbiome dissimilarity

was positively correlated with the host phylogenetic distance (Mantel correlation: 0.238; P = 0.0017; Figure 4-4B), indicating that more closely related populations also had more similar microbiomes. This correlation may arise because more closely related lineages share a more similar genetic background, or because they are more ecologically similar. However, we did not find a significant correlation between host microbiome dissimilarity and environmental distance based on bioclimatic variables (Mantel correlation: 0.033; P = 0.5959; Appendix Figure D-2).

Honing in on primarily host-associated microbes

The correlation between dissimilarities in gut and environmental microbiomes might suggest gut microbiomes are a simple reflection of the environmental pool from which microbes can be recruited. Consequently, we isolated primarily host-associated microbes by removing taxa that were detected in the environment from the gut microbiome data and reanalyzed patterns of community composition and similarities across species. After filtering out the environmental microbial taxa, there were 216 microbial taxa in the primarily host-associated microbiomes, which is ~56% of the 389 microbial taxa in the guts before environmental filtering. If similarities across species are a simple function of update of microbes that are common in the environment, we predicted that the correlation between similarities in gut and environmental microbiomes should disappear. However, we found that results were largely congruent with those from whole-microbiome analyses. Host species not only segregated in similar ways along the first two NMDS axes when primarily host-associated microbes were analyzed (Appendix Figure D-3), but the correlations between pairwise host microbiome dissimilarity with environmental microbiome dissimilarity and host phylogenetic distance persisted: the Mantel correlation with the environmental microbiome dissimilarity was 0.358 ($P < 0.0001$), and the Mantel correlation with

the host phylogenetic distance was 0.204 (P = 0.0065; Appendix Figure D-4). As before, there was not a significant correlation between the filtered gut microbiome dissimilarity and the environmental distance based on bioclimatic variables (Mantel correlation: –0.028; P = 0.7201). Consequently, the similarity between host and environmental microbiomes is not merely caused by fish ingesting microbes that are common in the environment.

Microbiome convergence among lineages of extremophile livebearing fishes

PERMANOVA revealed significant differences between sulfidic and nonsulfidic gut microbiomes, but is there evidence for convergence in microbiome composition among lineages that have independently colonized sulfide springs? Using a phylogenetic comparative approach to quantify convergence (Stayton 2015), we found strong evidence for convergence in the microbiomes of the sulfidic lineages. Six lineages have independently crossed into the region of space occupied by the convergent taxa, significantly more than expected by chance (P = 0.0032). On average, 43.8% of microbiome variation in sulfide spring lineages was attributable to convergence (C_1; P = 0.0003), and absolute measures of convergence were also significantly different than expected by change (C_2; P < 0.0001). However, only 18.2% of evolution within the sulfidic lineages was attributable to convergence (C_3; P = 0.0036), and there was not as much within-clade convergence (C_4), as only about 0.25% of the within-clade evolution was attributable to convergence (P = 0.0405). Overall, these analyses indicate that there is significant and strong signal of convergence in sulfide spring fish microbiomes, even though nonconvergent changes are also evident in sulfidic and nonsulfidic lineages. These nonconvergent changes could be attributable to other aspects of the ecology of host species (e.g., diet) or the environmental microbe pool that different lineages are exposed to. We found similar C_1-C_4

values when repeating the analysis on the primarily host-associated gut microbiome data set (Appendix Table D-3).

Upon finding evidence for convergence among the sulfidic lineages, we asked whether there were specific microbial taxa that exhibited convergent abundance shifts upon sulfide spring colonization, using expression variance and evolution (EVE) models that account for phylogenetic relationships among host lineages and variation in taxon abundance within lineages (Table 4-2). There were eight microbial taxa with convergent abundance shifts in whole-microbiomes (FDR < 0.05), representing 5 different classes (Table 4-2). Three taxa of Gammaproteobacterial bacteria—*Propionivibrio*, Chromatiaceae, and *Klebsiella*—are generally found in aquatic ecosystems (Podschun et al. 2001; Hotaling et al. 2019; Xie et al. 2023), and members of *Klebsiella* have been identified in fish gut microbiomes and can be pathogenic (Tyagi et al. 2019). There were two microbes in the family Geobacteraceae (identified as Geobacteraceae and *Trichlorobacter*), and these bacteria have been studied for their contributions to nutrient cycling in anaerobic environments (Sorokin et al. 2023). Like Geobacteraceae, Methylacidiphilaceae contribute to several nutrient cycles, but they are acidophilic methane-oxidizers with diverse metabolic capabilities found in extreme environments such as hot springs (Kruse et al. 2019; Dedysh et al. 2021). Prolixibacteraceae, like other Bacteroidia, likely degrade polysaccharides and proteins (Fernández-Gómez et al. 2013; Hamann et al. 2017), and members of this family have been found in the environment (Huang et al. 2014; Zhou et al. 2019) and in host-associated microbiomes of animals (Hinsu et al. 2021; Wang et al. 2022; Marangon et al. 2023). Finally, we identified one convergent microbial taxon, Babeliales, in the phylum Dependentiae. Dependentiae is not well described, but the only known members of Dependentiae are obligate intracellular symbionts and are found in a variety of

aquatic environments, including extreme environments, likely because of their association with protist hosts (Weisse et al. 2023).

Rerunning EVE models on primarily host-associated microbiomes (i.e., excluding taxa that are common in the environment) resulted in a smaller set of microbes with evidence for convergent abundance shifts. Five additional microbial taxa—*Terrimicrobium*, LKC2.127.25, *Paludibacter*, *Annamia* HOs24, and uncultured Phormidiaceae—emerged as convergent among the sulfidic lineages (Table 4-2). Several of these have also been identified in sediments and microbial mats of other environments, and some have been found in host-associated microbiomes as well. LKC2.127.25 is a microbe that has only been identified in a sulfidic cave (Engel et al. 2004). *Annamia* HOs24 and Phormidiaceae are cyanobacteria that have been found in other extreme environments such as high altitude saline lakes (Corona Ramírez et al. 2022) and sulfide-rich mangroves (Guidi-Rontani et al. 2014). While also found in environmental samples (Qiu et al. 2014), *Terrimicrobium* has additionally been identified in host associated microbiomes of animals, including in fish gut microbiomes (Xia et al. 2022; Deng et al. 2022; Huang et al. 2023). *Paludibacter* has also been found in the gut microbiomes of fish (Li et al. 2014; do Vale Pereira et al. 2024), and it is known to break down bacterial biomass (Mahmoud and Magdy 2021).

Discussion

A major question in microbial ecology is to what degree we are able to predict host-associated microbiome composition based on host phylogenetic background and the ecological conditions and environmental microbe pool that hosts are regularly exposed to. Our study of livebearing fishes inhabiting a variety of freshwater habitats indicated host-associated microbiomes were

distinct from the water and sediment microbiomes of their surrounding environments. However, different host species had more similar microbiomes if the microbial communities of their habitats were more alike, supporting past studies that found strong environmental effects on the composition of fish gut microbiomes (Kim et al. 2021; Sadeghi et al. 2023). In addition, we also found significant effects of host phylogenetic relationships, with more closely related species exhibiting more similar microbiomes. This effect may be a consequence of more closely related hosts sharing a more similar genetic background (if hosts exert any control over microbiome assembly) or more similar ecologies that expose them to the same environmental sources of microbes. Since correlation between host and environmental microbiomes was stronger than the correlation between the host microbiomes and host phylogenetic distance, the environmental microbial pool of host habitat seems to play a larger role than host phylogenetic relatedness in explaining microbiome variation in this system. In contrast, we found no evidence that variation in bioclimatic conditions impacts host microbiome composition, but colonization of extreme environments in the form of toxic springs rich in H_2S led to significant signatures of convergence among extremophile lineages. Microbial taxa that showed convergent abundance shifts among the extremophile populations included microbes that have been identified in host-associated microbiomes of other animals and may be of interest for future research testing the role of the microbiome in host adaptation. Overall, our findings shed light on the ecological and evolutionary drivers of host-microbiome diversification, and they show how selection can drive convergent shifts in host-associated microbial communities.

How are host-associated microbiomes shaped by the host's evolutionary history and environment?

In alignment with other research supporting a strong role for host habitat shaping fish microbiomes (Kim et al. 2021; Sadeghi et al. 2023), we observed that the fish gut microbiomes were most highly correlated with the environmental microbiomes, but the host-associated microbiomes were also reflective of the host's evolutionary history. In fishes, microbes from the host's environment that colonize the gut are especially derived from the surrounding water and the host's diet (Wong and Rawls 2012; Giatsis et al. 2015; Kashinskaya et al. 2018; Kim et al. 2021). Interestingly, even when we filtered microbes from the surrounding environment out of the gut microbiome data set, we still found significant correlations with the environmental microbiomes, suggesting that other environmental factors may contribute to the correlation between the environmental microbial communities and the gut microbial communities. We did not find a significant correlation with environmental distance based on local bioclimatic variables, but other environmental factors, such as physiochemical characteristics, have been shown to influence fish gut microbiome assembly and likely also play a role in shaping both the environmental and host-associated microbiomes in this system (Wong and Rawls 2012; Sylvain et al. 2016, 2019; Ramoneda et al. 2023). From our findings visualized in the NMDS with the overlaid host phylogeny (Figure 4-3), the pattern of host-associated microbiome variation may also be reflective of diet, although we could not formally test for microbiome correlations with host diet due to a lack of available diet data for several populations. Along NMDS axis 1, the microbiomes tended to align with what is known about the protein content of the host diets: host species with high amounts of protein in their diets had more negative axis scores, while host species with low amounts of protein in their diets had more positive axis scores. Protein-rich diets include those of *Belonesox belizanus*, which is a piscivorous species (Ferry-Graham et al. 2010), as well as species of the genera *Gambusia*, *Heterophallus*, *Priapella*, and

Pseudoxiphophorus, which consume invertebrates (Gibb et al. 2008; Riesch et al. 2011, 2012). Species with protein-poor diets include those of the genera *Poecilia* and *Xiphophorus*, which consume detritus and algae (Winemiller 1993; Kramer and Bryant 1995; Gibb et al. 2008; Tobler et al. 2015). In fishes, diet is one of the major factors that influence the gut microbiome composition, and fish trophic ecology itself is shaped by host evolution and the environment, making it challenging to assess its relative importance in gut microbiome assembly (Kim et al. 2021). Dietary shifts likely also contributed to microbiome variation in sulfide spring lineages, which transitioned from feeding on detritus and algae to feeding on microbial mats and invertebrates upon colonizing sulfide streams (Tobler et al. 2015). Even after excluding environmental microbes from the gut microbiome dataset, microbial taxa such as Chromatiaceae and Phormidiaceae were still present in the sulfidic fish guts. Bacteria in these families have been identified in microbial mats of other environments, suggesting that there may be some diet-derived microbes influencing variation that were also unaccounted for by water and sediment sampling (Zaar et al. 2003; Guidi-Rontani et al. 2014).

A major question about host-microbiome diversification is if microbiomes recapitulate the host evolutionary patterns. We found that the host-associated microbial communities were correlated with the host phylogeny. Alignment between host phylogenetic relatedness and host-associated microbiomes—a pattern referred to as "phylosymbiosis"—would be expected if co-evolution between hosts and their microbiota has occurred, but co-diversification of hosts and microbes can arise for reasons other than co-evolution (Mazel et al. 2018). For example, more closely related lineages often share similarities in ecology and physiology, which can lead to patterns of microbial recruitment that mirror the host phylogeny (Mazel et al. 2018; Mallott and Amato 2021). Detecting patterns of phylosymbiosis establishes the possibility that host

divergence has been accompanied by divergence in host regulation of the microbiome, but future research investigating how the host's genetic background shapes the host-associated microbiomes in a controlled environment and experiments testing for host regulation of the microbiome will be necessary for understanding how this pattern of host-microbiome co-diversification arose. For example, comparing the gut microbiomes between genetically distinct populations housed in the same environmental conditions for multiple generations has revealed microbiome differences in other fishes, supporting divergence in host recruitment of microbes while excluding the possibility of variation in environmental microbial pools influencing these microbiome differences (Riddle et al. 2023; Small et al. 2023). To further understand if microbiome variation associated with host genetic variation are due to selective recruitment of microbes and not a result of correlated host traits, future studies linking host immune evolution to microbial uptake will be important (Van den Abbeele et al. 2011).

Is there evidence of convergence in the microbiomes of hosts adapted to extreme environments? Identifying patterns of host-microbiome diversification also requires understanding how selection on both the host and microbiome drives changes in host-associated microbial communities. Our finding that multiple lineages of fishes showed convergence in their gut microbiomes upon colonization of sulfide springs suggests that similar environmental conditions can select for convergent shifts in host-associated microbial communities. In other fishes, microbiome convergence has not always been found, but it has overall not been well-studied. For example, Trinidadian guppies (also in the family Poeciliidae) did not show microbiome convergence among low-predation ecotypes relative to high-predation ecotypes, which are ecotypes with divergence in traits such as host physiology, trophic ecology, and gut length

(Sullam et al. 2015). Dwarf whitefish pairs similarly did not show convergence in their microbiomes relative to normal whitefish, despite their convergence in other traits, including their trophic ecology (Laporte et al. 2015; Dalziel et al. 2017; Sevellec et al. 2018). In contrast, African cichlids showed convergence in their microbiomes based on their trophic level (convergence among herbivores) (Baldo et al. 2017), and two species of Amazon fishes exhibited structural microbiome convergence associated with water chemistry (Sylvain et al. 2019). There are different potential explanations for the microbiome convergence we found among extremophile poeciliids. In our system, it is well-established that sulfidic fish have adaptations that allow them to tolerate high levels of H_2S and hypoxia (Barts et al. 2018; Greenway et al. 2020), and they also have convergent changes in gut morphology (a shortened gut length) due to the shift toward a more protein-rich diet through feeding on bacteria and invertebrates in sulfide springs (Tobler et al. 2015). The microbiome convergence we found may occur because sulfidic fish have shared changes in phenotypic traits and environmental microbial communities, but not necessarily because they have evolved to recruit adaptive microbes. For example, changes in host physiology and morphology in sulfide streams may affect which microbes colonize the host (Sullam et al. 2015). Certain bacteria in the external environment may be better equipped to colonize the gut environment, without any selective uptake of microbes from the host. Convergence is also likely driven in part by hosts selecting from a similar pool of microbes, as evidenced by our finding that the environmental microbiomes of sulfide springs are similar to each other. Alternatively, selection on host-microbe interactions in sulfidic environments may drive the evolution of adaptive symbioses in which the host's microbiota provide benefits to the host, such as H_2S detoxification, protein digestion, or hypoxia tolerance. In other H_2S-rich environments, sulfur-oxidizing bacteria form symbioses with invertebrates,

providing nutritional or detoxification benefits to their hosts (Dubilier et al. 1995; Beinart et al. 2019). While most of the microbes exhibiting convergent abundance shifts among the sulfidic lineages have not been well-described, some of the microbial taxa may include bacteria with sulfur-oxidizing capabilities, such as bacteria in the genus *Klebsiella* (Mawad et al. 2021), and potential sulfur-oxidizers in sulfidic fish gut microbiomes could be important to investigate further for possible nutritional or detoxification benefits to the host. Bacteria in the genus *Klebsiella* are also found in fish gut microbiomes, along with *Paludibacter* and *Terrimicrobium* (Li et al. 2014; Tyagi et al. 2019; Deng et al. 2022; do Vale Pereira et al. 2024). In fish digestive tracts, *Paludibacter* has been associated with the degradation of protein derived from bacterial biomass (do Vale Pereira et al. 2024). Products that may be beneficial to the host are produced at various steps of bacterial biomass degradation by *Paludibacter*, including short-chain fatty acids such as propionate, butyrate, and acetate (Mahmoud and Magdy 2021; do Vale Pereira et al. 2024). Bacteria in the family Prolixibacteraceae, which also was a convergent taxon among the sulfidic lineages, have similarly been identified in host-associated microbiomes of animals and are thought to break down protein and polysaccharides (Hinsu et al. 2021). Consequently, microbes like *Paludibacter* and Prolixibacteraceae that may help degrade protein from the bacteria-rich diet of sulfidic fish could provide nutritional benefits to the host. Protein degradation is expected to be important for sulfidic fish, but sulfide-adapted lineages also must cope with chronic hypoxia, which has negative effects on protein digestibility and can disrupt gut microbiome stability in aquatic organisms, increasing susceptibility to pathogens (Zambonino-Infante et al. 2017; Xiong et al. 2019; Abdel-Tawwab et al. 2019). An important role of the resident gut microbiota in fish is to protect the host against infection by maintaining stability of the gut microbial community (Gallo et al. 2020; Pérez-Pascual et al. 2021), and microbes from

our analyses that also are consistently found in other fish species may be stable members of healthy fish gut microbiomes. Host-associated microbes that maintain gut microbiome stability and facilitate protein digestion may especially be important for sulfidic fish to rely on a protein-rich diet while tolerating hypoxia.

Conclusion

Identifying the factors shaping patterns of diversification of hosts and their associated microbial communities is important for understanding host-microbiome evolution. Our findings that host-associated microbial communities of livebearing fishes are shaped by the external environmental microbial communities and by the host's evolutionary history provide insight into fish microbiome assembly, and they establish a basis for investigating the mechanisms underlying patterns of host-microbiome diversification. We also found convergence in the gut microbiomes of lineages of fishes that have independently colonized toxic sulfide springs. In sulfide springs, extreme environmental pressures give rise to predictable host-microbiome associations, which provides a foundation for studying the outcomes of these repeated shifts in host-microbiome assembly. For example, does natural selection drive changes in microbiome function, and can selection on host-microbiome interactions give rise to adaptive symbioses? Future research investigating the role of the microbiome in host functions will be important for addressing these questions, such as microbiome reciprocal transplant experiments to understand the interaction between the microbiome and host physiology, coupled with shotgun metagenomics approaches to investigate the functional potential of host-associated extremophile microbes in nature.

Tables

Table 4-1. Coordinates of the sampling sites and the number of water, sediment, and intestine samples collected from all 13 sites. In total, there were 199 samples collected, but 195 samples remained in the analysis after removing 4 samples with fewer than 4,500 reads. In cases where not all of the collected samples were included in the analysis, the number of samples included in the analysis is listed in parentheses. There were 128 total intestine samples and 66 environmental samples (40 water samples and 26 sediment samples) included in the analysis.

Site	Coordinates	H_2S	*N* water	*N* sediment	Species	*N*
Arroyo Rosita	17.485, -93.104	-	3	2	*Belonesox belizanus*	6 (5)
					Gambusia yucatana	2
					Poecilia mexicana	6
La Gloria	17.532, -93.015	+	3	3	*Poecilia sulphuraria*	6
					Pseudoxiphophorus bimaculatus	7
					Xiphophorus hellerii	6
Arroyo Caracol	17.558, -93.043	-	3	2 (1)	*Heterandria bimaculata*	6
					Xiphophorus hellerii	6
Baños del Azufre	17.552, -92.999	+	3	2	*Gambusia eurystoma*	7
Tributary to Ixtapangajoya	17.510, -92.980	-	3	2	*Poecilia mexicana*	6
La Esperanza	17.511, -92.983	+	3	2	*Poecilia thermalis*	6
Puyacatengo downstream	17.564, -92.912	-	3	2	*Poecilia mexicana*	6
La Lluvia	17.464, -92.896	+	3	2	*Poecilia mexicana*	6
Arroyo Bonita	17.427, -92.757	-	4	2	*Heterophallus milleri*	6
					Priapella chamulae	6
					Poecilia mexicana	6 (5)
El Azufre I	17.442, -92.775	+	3	2	*Poecilia mexicana*	6
Laguna near Teapa	17.590,-92.981	-	3	2	*Carlhubbsia kidderi*	6
Laguna near Simón Sarlat	18.341, -92.792	-	4	2	*Gambusia sexradiata*	6
					Phallichthys fairweatheri	6
					Poecilia kykesis	1
					Xiphophorus maculatus	6
Cueva del Azufre	17.442, -92.775	+	2	2	*Poecilia mexicana*	6 (5)

Table 4-2. Microbial taxa showing convergent shifts among the sulfidic lineages. The convergent microbes were identified both before and after filtering the gut microbiome dataset to exclude microbes that were detected in the environment. Each microbe in the table was listed down to its lowest known taxonomic level. After environmental correction, five additional taxa showed convergent abundance shifts: *Terrimicrobium*, LKC2.127.25, *Paludibacter*, *Annamia* HOs24, and uncultured Phormidiaceae.

Convergent microbial taxa before environmental correction					
Phylum	**Class**	**Order**	**Family**	**Genus**	**FDR**
Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	*Propionivibrio*	0.0119
Proteobacteria	Gammaproteobacteria	Chromatiales	Chromatiaceae		0.0044
Proteobacteria	Gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Klebsiella*	0.0007
Verrucomicrobiota	Verrucomicrobiae	Methylacidiphilales	Methylacidiphilaceae	uncultured	0.0053
Bacteroidota	Bacteroidia	Bacteroidales	Prolixibacteraceae	uncultured	0.0008
Dependentiae	Babeliae	Babeliales			0.0040
Desulfobacterota	Desulfuromonadia	Geobacterales	Geobacteraceae		0.0095
Desulfobacterota	Desulfuromonadia	Geobacterales	Geobacteraceae	*Trichlorobacter*	0.0049
Convergent microbial taxa after environmental correction					
Proteobacteria	Gammaproteobacteria	Chromatiales	Chromatiaceae		0.0024
Verrucomicrobiota	Verrucomicrobiae	Chthoniobacterales	Terrimicrobiaceae	*Terrimicrobium*	0.0200
Bacteroidota	Bacteroidia	Bacteroidales	LKC2.127.25	LKC2.127.25	0.0021
Bacteroidota	Bacteroidia	Bacteroidales	Paludibacteraceae	*Paludibacter*	0.0221
Cyanobacteria	Cyanobacteriia	Cyanobacteriales	Cyanobacteriaceae	*Annamia* HOs24	0.0360
Cyanobacteria	Cyanobacteriia	Cyanobacteriales	Phormidiaceae	uncultured	0.0329

Figures

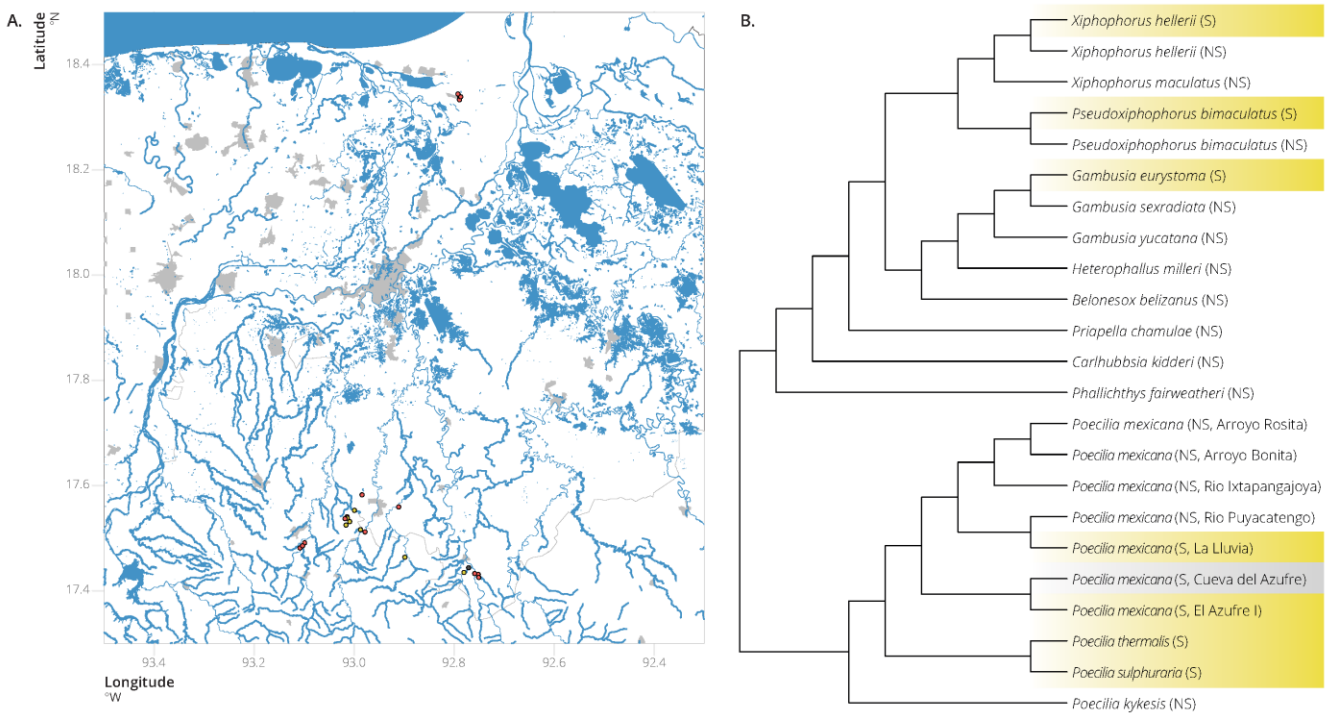


Figure 4-1. A) Geographic region of all the sampling sites. Orange points represent nonsulfidic sites, and yellow points represent sulfidic sites. The sulfidic cave site is represented by a gray point. B) Phylogenetic tree showing the evolutionary relatedness between host lineages. Sulfidic lineages are highlighted in yellow, and the sulfidic and cave lineage is highlighted in gray.

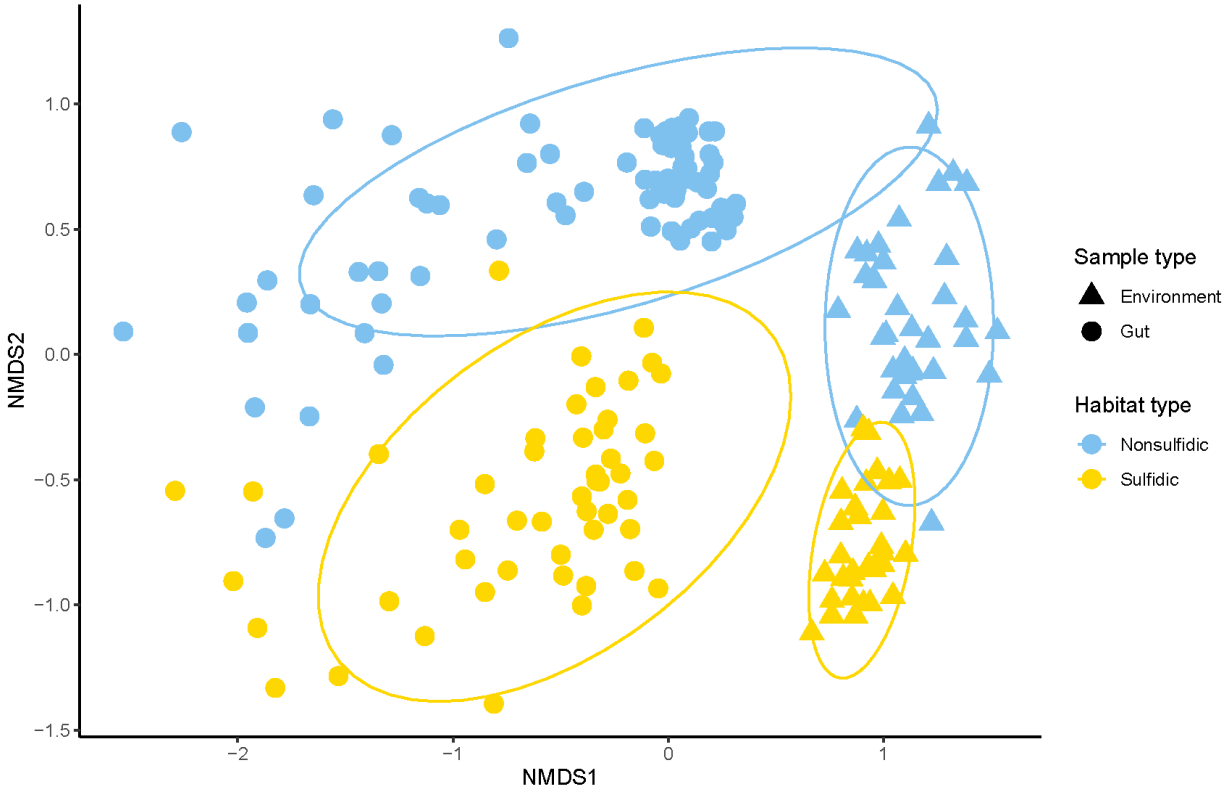


Figure 4-2. NMDS based on Bray-Curtis dissimilarity showing gut and environmental microbiome variation. Samples from sulfidic environments are marked as yellow points, and samples from nonsulfidic environments are marked as blue points. Gut samples are represented by circles, and environmental samples are represented by triangles.

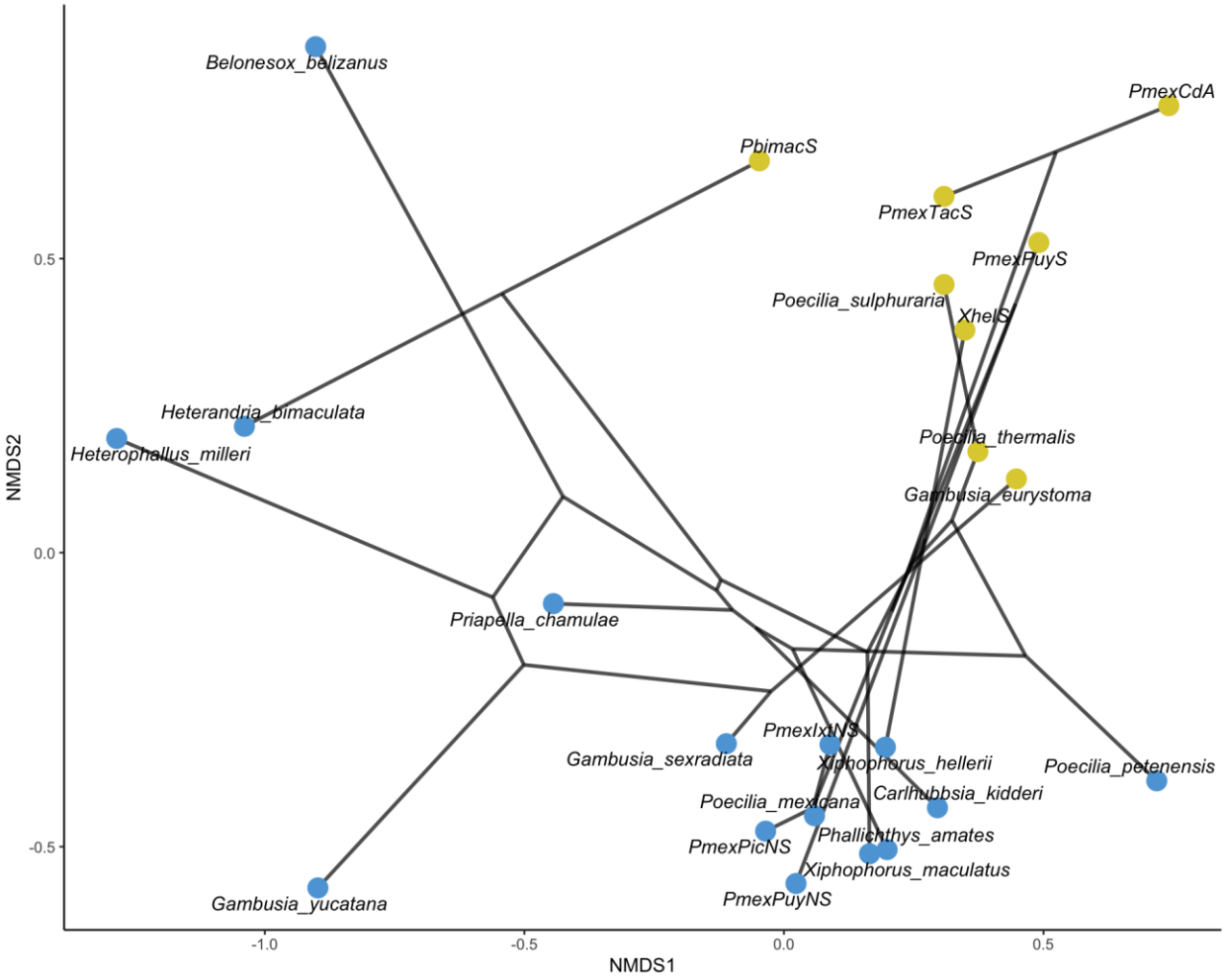


Figure 4-3. NMDS based on Bray-Curtis dissimilarity showing the gut microbiome variation of all the lineages. The points are NMDS coordinate averages of each lineage's microbiome composition, and the lines connecting the points reflect the host phylogenetic relationships.

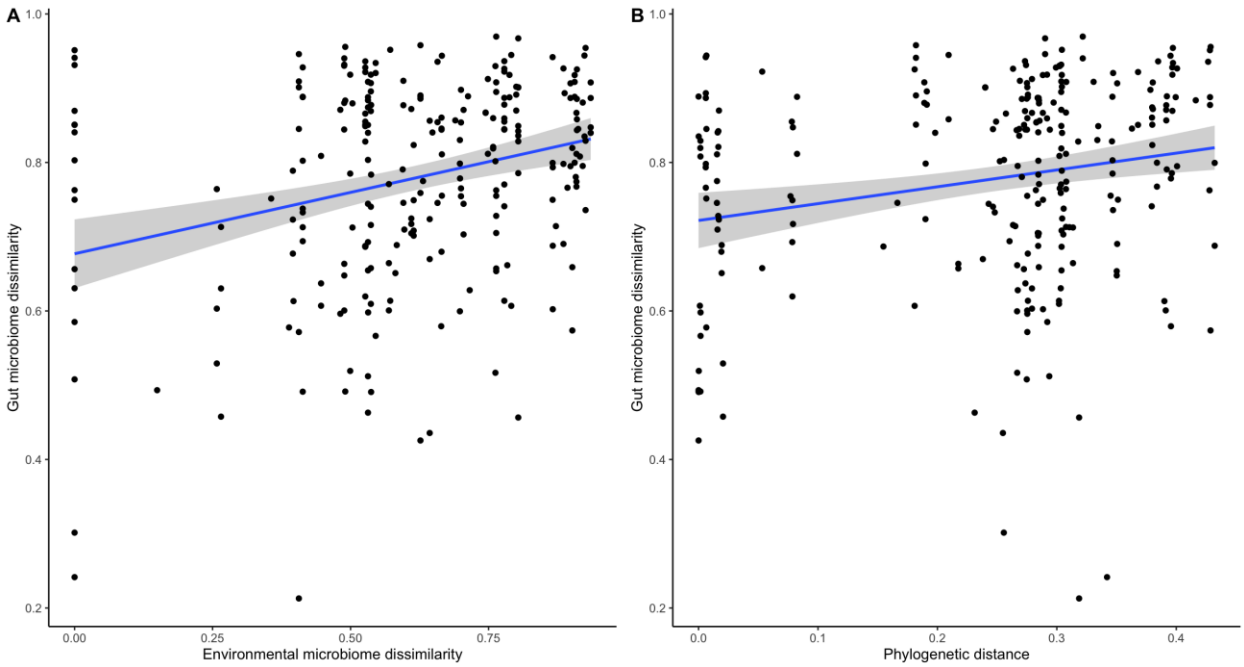


Figure 4-4. Correlations between gut microbiome dissimilarity and A) environmental microbiome dissimilarity (Mantel correlation: 0.271; *P* = 0.0008) and B) host phylogenetic distance (Mantel correlation: 0.238; *P* = 0.0017).

References

Abdel-Tawwab M, Monier MN, et al (2019) Fish response to hypoxia stress: growth, physiological, and immunological biomarkers. Fish Physiology and Biochemistry 45:997–1013. https://doi.org/10.1007/s10695-019-00614-9

Alberdi A, Aizpurua O, Bohmann K, et al (2016) Do Vertebrate Gut Metagenomes Confer Rapid Ecological Adaptation? Trends in Ecology & Evolution 31:689–699. https://doi.org/10.1016/j.tree.2016.06.008

Aranda CP, Valenzuela C, Matamala Y, et al (2015) Sulphur-cycling bacteria and ciliated protozoans in a *Beggiatoaceae* mat covering organically enriched sediments beneath a salmon farm in a southern Chilean fjord. Marine Pollution Bulletin 100:270–278. https://doi.org/10.1016/j.marpolbul.2015.08.040

Aspiras AC, Rohner N, Martineau B, et al (2015) Melanocortin 4 receptor mutations contribute to the adaptation of cavefish to nutrient-poor conditions. PNAS 112:9668–9673. https://doi.org/10.1073/pnas.1510802112

Avise JC, Selander RK (1972) Evolutionary Genetics of Cave-Dwelling Fishes of the Genus Astyanax. Evolution 26:1–19. https://doi.org/10.2307/2406978

Baldo L, Pretus JL, Riera JL, et al (2017) Convergence of gut microbiotas in the adaptive radiations of African cichlid fishes. ISME J 11:1975–1987. https://doi.org/10.1038/ismej.2017.62

Bang C, Dagan T, Deines P, et al (2018) Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? Zoology 127:1–19. https://doi.org/10.1016/j.zool.2018.02.004

Barton LL, Fardeau M-L, Fauque GD (2014) Hydrogen Sulfide: A Toxic Gas Produced by Dissimilatory Sulfate and Sulfur Reduction and Consumed by Microbial Oxidation. In: Kroneck PMH, Torres MES (eds) The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment. Springer Netherlands, Dordrecht, pp 237–277.

Barts N, Greenway R, Passow CN, et al (2018) Molecular evolution and expression of oxygen transport genes in livebearing fishes (Poeciliidae) from hydrogen sulfide rich springs. Genome 61:273–286. https://doi.org/10.1139/gen-2017-0051

Bates D, Mächler M, Bolker B, Walker S (2014) Fitting Linear Mixed-Effects Models using lme4. Journal of Statistical Software 67:1–48

Behrmann-Godel J, Nolte AW, Kreiselmaier J, et al (2017) The first European cave fish. Current Biology 27:R257–R258. https://doi.org/10.1016/j.cub.2017.02.048

Beinart RA, Luo C, Konstantinidis KT, et al (2019) The Bacterial Symbionts of Closely Related Hydrothermal Vent Snails With Distinct Geochemical Habitats Show Broad Similarity in

Chemoautotrophic Gene Content. Frontiers in Microbiology 10. https://doi.org/10.3389/fmicb.2019.01818

Beladjal L, Vandekerckhove TTM, Muyssen B, et al (2002) B-chromosomes and male-biased sex ratio with paternal inheritance in the fairy shrimp *Branchipus schaefferi* (Crustacea, Anostraca). Heredity 88:356–360. https://doi.org/10.1038/sj.hdy.6800061

Belkova NL, Sidorova TV, Glyzina OY, et al (2017) Gut microbiome of juvenile coregonid fishes: comparison of sympatric species and their F1 hybrids. Fundamental and Applied Limnology 279–290. https://doi.org/10.1127/fal/2016/0804

Bellec L, Cambon-Bonavita M-A, Durand L, et al (2020) Microbial Communities of the Shallow-Water Hydrothermal Vent Near Naples, Italy, and Chemosynthetic Symbionts Associated With a Free-Living Marine Nematode. Frontiers in Microbiology 11. https://doi.org/10.3389/fmicb.2020.02023

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bhute SS, Escobedo B, Haider M, et al (2020) The gut microbiome and its potential role in paradoxical anaerobism in pupfishes of the Mojave Desert. Animal Microbiome 2:20. https://doi.org/10.1186/s42523-020-00037-5

Blanckenhorn WU (2000) The Evolution of Body Size: What Keeps Organisms Small? The Quarterly Review of Biology 75:385–407. https://doi.org/10.1086/393620

Blazejak A, Erséus C, Amann R, Dubilier N (2005) Coexistence of Bacterial Sulfide Oxidizers, Sulfate Reducers, and Spirochetes in a Gutless Worm (Oligochaeta) from the Peru Margin. Applied and Environmental Microbiology 71:1553–1561. https://doi.org/10.1128/AEM.71.3.1553-1561.2005

Bolyen E, Rideout JR, Dillon MR, et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology 37:852–857. https://doi.org/10.1038/s41587-019-0209-9

Borowsky R (2008) Determining the Sex of Adult *Astyanax mexicanus*. Cold Spring Harbor Protocols 2008:pdb.prot5090. https://doi.org/10.1101/pdb.prot5090

Bradic M, Beerli P, García-de León FJ, et al (2012) Gene flow and population structure in the Mexican blind cavefish complex (*Astyanax mexicanus*). BMC Evolutionary Biology 12:9. https://doi.org/10.1186/1471-2148-12-9

Breusing C, Johnson SB, Tunnicliffe V, et al (2020a) Allopatric and Sympatric Drivers of Speciation in Alviniconcha Hydrothermal Vent Snails. Molecular Biology and Evolution 37:3469–3484. https://doi.org/10.1093/molbev/msaa177

Breusing C, Mitchell J, Delaney J, et al (2020b) Physiological dynamics of chemosynthetic symbionts in hydrothermal vent snails. ISME J 14:2568–2579. https://doi.org/10.1038/s41396-020-0707-2

Brightly WH, Stayton CT (2023) convevol: Analysis of Convergent Evolution. https://CRAN.R-project.org/package=convevol

Burnham KP, Anderson DR (2001) Kullback-Leibler information as a basis for strong inference in ecological studies. Wildlife Research 28:111–119. https://doi.org/10.1071/wr99107

Callahan BJ, McMurdie PJ, Rosen MJ, et al (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13:581–583. https://doi.org/10.1038/nmeth.3869

Caporaso JG, Kuczynski J, Stombaugh J, et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303

Cavanaugh CM, McKiness ZP, Newton ILG, Stewart FJ (2006) Marine Chemosynthetic Symbioses. In: Dworkin M, Falkow S, Rosenberg E, et al. (eds) The Prokaryotes. Springer New York, New York, NY, pp 475–507.

Chiu L, Wang M-C, Tseng K-Y, et al (2022) Shallow-water hydrothermal vent system as an extreme proxy for discovery of microbiome significance in a crustacean holobiont. Frontiers in Marine Science 9. https://doi.org/10.3389/fmars.2022.976255

Clark FE, Kocher TD (2019) Changing sex for selfish gain: B chromosomes of Lake Malawi cichlid fish. Scientific Reports 9:20213. https://doi.org/10.1038/s41598-019-55774-8

Clavel T, Neto JCG, Lagkouvardos I, Ramer-Tait AE (2017) Deciphering interactions between the gut microbiota and the immune system via microbial cultivation and minimal microbiomes. Immunological Reviews 279:8–22. https://doi.org/10.1111/imr.12578

Colston TJ, Jackson CR (2016) Microbiome evolution along divergent branches of the vertebrate tree of life: what is known and unknown. Molecular Ecology 25:3776–3800. https://doi.org/10.1111/mec.13730

Cooper CE, Brown GC (2008) The inhibition of mitochondrial cytochrome oxidase by the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical mechanism and physiological significance. Journal of Bioenergetics and Biomembranes 40:533–539. https://doi.org/10.1007/s10863-008-9166-6

Cornelio D, Castro JP, Santos MH, et al (2017) Hermaphroditism can compensate for the sex ratio in the *Astyanax scabripinnis* species complex (Teleostei: Characidae): expanding the B chromosome study model. Reviews in Fish Biology and Fisheries 27:681–689. https://doi.org/10.1007/s11160-017-9488-8

Corona Ramírez A, Cailleau G, Fatton M, et al (2022) Diversity of Lysis-Resistant Bacteria and Archaea in the Polyextreme Environment of Salar de Huasco. Frontiers in Microbiology 13:. https://doi.org/10.3389/fmicb.2022.826117

Cuellar-Gempeler C, Leibold MA (2019) Key colonist pools and habitat filters mediate the composition of fiddler crab–associated bacterial communities. Ecology 100:e02628. https://doi.org/10.1002/ecy.2628

Culumber ZW, Tobler M (2017) Sex-specific evolution during the diversification of live-bearing fishes. Nature Ecology & Evolution 1:1185–1191. https://doi.org/10.1038/s41559-017-0233-4

Culver DC, Holsinger JR (1969) Preliminary observations on sex ratios in the subterranean amphipod genus *Stygonectes* (Gammaridae). The American Midland Naturalist 82:631–633. https://doi.org/10.2307/2423810

Culver D, Kane T, Fong D (1995). Adaptation and Natural Selection in Caves: The Evolution of *Gammarus minus*. Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674419070

Custer GF, Gans M, van Diepen LTA, et al (2023) Comparative Analysis of Core Microbiome Assignments: Implications for Ecological Synthesis. mSystems 8:e01066-22. https://doi.org/10.1128/msystems.01066-22

Dalziel AC, Laporte M, Rougeux C, et al (2017) Convergence in organ size but not energy metabolism enzyme activities among wild Lake Whitefish (*Coregonus clupeaformis*) species pairs. Molecular Ecology 26:225–244. https://doi.org/10.1111/mec.13847

Dedysh SN, Beletsky AV, Ivanova AA, et al (2021) Peat-Inhabiting Verrucomicrobia of the Order Methylacidiphilales Do Not Possess Methanotrophic Capabilities. Microorganisms 9:2566. https://doi.org/10.3390/microorganisms9122566

Deng Y, Borewicz K, van Loo J, et al (2022) In-Situ Biofloc Affects the Core Prokaryotes Community Composition in Gut and Enhances Growth of Nile Tilapia (*Oreochromis niloticus*). Microbial Ecology 84:879–892. https://doi.org/10.1007/s00248-021-01880-y

Dick GJ (2019) The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. Nature Reviews Microbiology 17:271–283. https://doi.org/10.1038/s41579-019-0160-2

Diwan AD, Harke SN, Panche AN (2023) Host-microbiome interaction in fish and shellfish: An overview. Fish and Shellfish Immunology Reports 4:100091. https://doi.org/10.1016/j.fsirep.2023.100091

Dixon P (2003) VEGAN, a package of R functions for community ecology. Journal of Vegetation Science 14:927–930. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

do Vale Pereira G, Teixeira C, Couto J, et al (2024) Dietary Protein Quality Affects the Interplay between Gut Microbiota and Host Performance in Nile Tilapia. Animals 14:714. https://doi.org/10.3390/ani14050714

Dubilier N, Giere O, Distel DL, Cavanaugh CM (1995) Characterization of chemoautotrophic bacterial symbionts in a gutless marine worm (Oligochaeta, Annelida) by phylogenetic 16S rRNA sequence analysis and in situ hybridization. Applied and Environmental Microbiology 61:2346–2350. https://doi.org/10.1128/aem.61.6.2346-2350.1995

Duval C, Marie B, Foucault P, Duperron S (2022) Establishment of the Bacterial Microbiota in a Lab-Reared Model Teleost Fish, the Medaka *Oryzias latipes*. Microorganisms 10:2280. https://doi.org/10.3390/microorganisms10112280

Eichmiller JJ, Hamilton MJ, Staley C, et al (2016) Environment shapes the fecal microbiome of invasive carp species. Microbiome 4:44. https://doi.org/10.1186/s40168-016-0190-1

Engel AS, Porter ML, Stern LA, et al (2004) Bacterial diversity and ecosystem function of filamentous microbial mats from aphotic (cave) sulfidic springs dominated by chemolithoautotrophic "Epsilonproteobacteria." FEMS Microbiology Ecology 51:31–53. https://doi.org/10.1016/j.femsec.2004.07.004

Enyidi U, Kiljunen M, Jones RI, et al (2013) Nutrient Assimilation by First-Feeding African Catfish, *Clarias gariepinus*, Assessed Using Stable Isotope Analysis. Journal of the World Aquaculture Society 44:161–172. https://doi.org/10.1111/jwas.12016

Espinasa L, Bibliowicz J, Jeffery WR, Rétaux S (2014) Enhanced prey capture skills in *Astyanax* cavefish larvae are independent from eye loss. EvoDevo 5:35. https://doi.org/10.1186/2041-9139-5-35

Espinasa L, Bonaroti N, Wong J, et al (2017) Contrasting feeding habits of post-larval and adult *Astyanax* cavefish. Subterranean Biology 21:1–17. https://doi.org/10.3897/subtbiol.21.11046

Espinasa L, Borowsky RB (2001) Origins and Relationship of Cave Populations of the Blind Mexican Tetra, *Astyanax fasciatus*, in the Sierra de El Abra. Environmental Biology of Fishes 62:233–237. https://doi.org/10.1023/A:1011881921023

Espinasa L, Legendre L, Fumey J, et al (2018) A new cave locality for *Astyanax* cavefish in Sierra de El Abra, Mexico. Subterranean Biology 26:39–53. https://doi.org/10.3897/subtbiol.26.26643

Espino del Castillo A, Castaño G, Davila-Montes M, et al (2009) Seasonal distribution and circadian activity in the troglophile long-footed robber frog, *Eleutherodactylus longipes* (Anura: Brachycephalidae) at Los Riscos Cave, Querétaro, Mexico: Field and laboratory studies. Journal of Cave and Karst Studies 71: 24-31.

Fenolio DB, Graening GO, Collier BA, Stout JF (2006) Coprophagy in a cave-adapted salamander; the importance of bat guano examined through nutritional and stable isotope

analyses. Proceedings of the Royal Society B 273:439–443. https://doi.org/10.1098/rspb.2005.3341

Fernández-Gómez B, Richter M, Schüler M, et al (2013) Ecology of marine Bacteroidetes: a comparative genomics approach. ISME J 7:1026–1037. https://doi.org/10.1038/ismej.2012.169

Ferry-Graham LA, Hernandez LP, Gibb AC, Pace C (2010) Unusual kinematics and jaw morphology associated with piscivory in the poeciliid, *Belonesox belizanus*. Zoology 113:140–147. https://doi.org/10.1016/j.zool.2009.09.001

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37:4302–4315. https://doi.org/10.1002/joc.5086

Fidopiastis PM, Bezdek DJ, Horn MH, Kandel JS (2006) Characterizing the resident, fermentative microbial consortium in the hindgut of the temperate-zone herbivorous fish, *Hermosilla azurea* (Teleostei: Kyphosidae). Marine Biology 148:631–642. https://doi.org/10.1007/s00227-005-0106-2

Finster KW, Kjeldsen KU, Kube M, et al (2013) Complete genome sequence of *Desulfocapsa sulfexigens*, a marine deltaproteobacterium specialized in disproportionating inorganic sulfur compounds. Standards in Genomic Sciences 8:58–68. https://doi.org/10.4056/sigs.3777412

Fisher MA, Oleksiak MF (2007) Convergence and divergence in gene expression among natural populations exposed to pollution. BMC Genomics 8:108. https://doi.org/10.1186/1471-2164-8-108

Fontaine SS, Mineo PM, Kohl KD (2022) Experimental manipulation of microbiota reduces host thermal tolerance and fitness under heat stress in a vertebrate ectotherm. Nature Ecology & Evolution 6:405–417. https://doi.org/10.1038/s41559-022-01686-2

Frank SA (1990) Sex Allocation Theory for Birds and Mammals. Annual Review of Ecology and Systematics 21:13–55. https://doi.org/10.1146/annurev.es.21.110190.000305

Franz-Odendaal TA, Hall BK (2006) Modularity and sense organs in the blind cavefish, *Astyanax mexicanus*. Evolution & Development 8:94–100. https://doi.org/10.1111/j.1525-142X.2006.05078.x

Frøland Steindal IA, Beale AD, Yamamoto Y, Whitmore D (2018) Development of the Astyanax mexicanus circadian clock and non-visual light responses. Developmental Biology 441:345–354. https://doi.org/10.1016/j.ydbio.2018.06.008

Fryxell DC, Arnett HA, Apgar TM, et al (2015) Sex ratio variation shapes the ecological effects of a globally introduced freshwater fish. Proceedings of the Royal Society B 282:20151970. https://doi.org/10.1098/rspb.2015.1970

Furness AI, Pollux BJA, Meredith RW, et al (2019) How conflict shapes evolution in poeciliid fishes. Nature Communications 10:3335. https://doi.org/10.1038/s41467-019-11307-5

Gallo BD, Farrell JM, Leydet BF (2020) Fish Gut Microbiome: A Primer to an Emerging Discipline in the Fisheries Sciences. Fisheries 45:271–282. https://doi.org/10.1002/fsh.10379

Giatsis C, Sipkema D, Smidt H, et al (2015) The impact of rearing environment on the development of gut microbiota in tilapia larvae. Scientific Reports 5:18206. https://doi.org/10.1038/srep18206

Gibb A, Ferry-Graham LA, Hernandez LP, et al (2008) Functional significance of intramandibular bending in Poeciliid fishes. Environmental Biology of Fishes 83:507–519. https://doi.org/10.1007/s10641-008-9369-z

Giovannelli D, Chung M, Staley J, et al (2016) *Sulfurovum riftiae* sp. nov., a mesophilic, thiosulfate-oxidizing, nitrate-reducing chemolithoautotrophic epsilonproteobacterium isolated from the tube of the deep-sea hydrothermal vent polychaete *Riftia pachyptila*. International Journal of Systematic and Evolutionary Microbiology 66:2697–2701. https://doi.org/10.1099/ijsem.0.001106

Global climate and weather data — WorldClim 1 documentation. https://worldclim.org/data/index.html.

Goslee SC, Urban DL (2007) The ecodist Package for Dissimilarity-based Analysis of Ecological Data. Journal of Statistical Software 22:1–19. https://doi.org/10.18637/jss.v022.i07

Gould SJ (1989) Wonderful Life: The Burgess Shale and the Nature of History. W. W. Norton & Company.

Greenway R, Arias-Rodriguez L, Diaz P, Tobler M (2014) Patterns of Macroinvertebrate and Fish Diversity in Freshwater Sulphide Springs. Diversity 6:597–632. https://doi.org/10.3390/d6030597

Greenway R, Barts N, Henpita C, et al (2020) Convergent evolution of conserved mitochondrial pathways underlies repeated adaptation to extreme environments. Proceedings of the National Academy of Sciences 117:16424–16430. https://doi.org/10.1073/pnas.2004223117

Greenway R, De-Kayne R, Brown AP, et al (2023) Integrative analyses of convergent adaptation in sympatric extremophile fishes. 2021.06.28.450104. bioRxiv. https://doi.org/10.1101/2021.06.28.450104

Gross JB (2012) The complex origin of *Astyanax* cavefish. BMC Evolutionary Biology 12:105. https://doi.org/10.1186/1471-2148-12-105

Guidi-Rontani C, Jean MRN, Gonzalez-Rizzo S, et al (2014) Description of new filamentous toxic Cyanobacteria (Oscillatoriales) colonizing the sulfidic periphyton mat in marine mangroves. FEMS Microbiology Letters 359:173–181. https://doi.org/10.1111/1574-6968.12551

Gupta A, Nair S (2020) Dynamics of Insect–Microbiome Interaction Influence Host and Microbial Symbiont. Frontiers in Microbiology 11:. https://doi.org/10.3389/fmicb.2020.01357

Hamann E, Tegetmeyer HE, Riedel D, et al (2017) Syntrophic linkage between predatory *Carpediemonas* and specific prokaryotic populations. ISME J 11:1205–1217. https://doi.org/10.1038/ismej.2016.197

Hao YT, Wu SG, Xiong F, et al (2017) Succession and Fermentation Products of Grass Carp (*Ctenopharyngodon idellus*) Hindgut Microbiota in Response to an Extreme Dietary Shift. Frontiers in Microbiology 8. https://doi.org/10.3389/fmicb.2017.01585

Härer A, Mauro AA, Laurentino TG, et al (2023) Gut microbiota parallelism and divergence associated with colonisation of novel habitats. Molecular Ecology 32:5661–5672. https://doi.org/10.1111/mec.17135

Hervant F (2012) Starvation in Subterranean Species Versus Surface-Dwelling Species: Crustaceans, Fish, and Salamanders. In: McCue MD (ed) Comparative Physiology of Fasting, Starvation, and Food Limitation. Springer, Berlin, Heidelberg, pp 91–102.

Hervant F, Renault D (2002) Long-term fasting and realimentation in hypogean and epigean isopods: a proposed adaptive strategy for groundwater organisms. Journal of Experimental Biology 205:2079–2087. https://doi.org/10.1242/jeb.205.14.2079

Hijmans RJ, Etten J van, Sumner M, et al (2023) raster: Geographic Data Analysis and Modeling. https://CRAN.R-project.org/package=raster

Hinaux H, Devos L, Blin M, et al (2016) Sensory evolution in blind cavefish is driven by early embryonic events during gastrulation and neurulation. Development 143:4521–4532. https://doi.org/10.1242/dev.141291

Hinsu AT, Tulsani NJ, Panchal KJ, et al (2021) Characterizing rumen microbiota and CAZyme profile of Indian dromedary camel (*Camelus dromedarius*) in response to different roughages. Scientific Reports 11:9400. https://doi.org/10.1038/s41598-021-88943-9

Hoang KL, Gerardo NM, Morran LT (2021) Association with a novel protective microbe facilitates host adaptation to a stressful environment. Evolution Letters 5:118–129. https://doi.org/10.1002/evl3.223

Hollister JW (2021). elevatr: access elevation data from various APIs. R package version 0.4.1. https://CRAN.R-project.org/package=elevatr/

Hotaling S, Quackenbush CR, Bennett-Ponsford J, et al (2019) Bacterial Diversity in Replicated Hydrogen Sulfide-Rich Streams. Microbial Ecology 77:559–573. https://doi.org/10.1007/s00248-018-1237-6

Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Molecular Phylogenetics and Evolution 43:986–998. https://doi.org/10.1016/j.ympev.2006.06.009

Huang Q, Shan H-W, Chen J-P, Wu W (2023) Diversity and Dynamics of Bacterial Communities in the Digestive and Excretory Systems across the Life Cycle of Leafhopper, *Recilia dorsalis*. Insects 14:545. https://doi.org/10.3390/insects14060545

Huang X-F, Liu YJ, Dong J-D, et al (2014) Mangrovibacterium diazotrophicum gen. nov., sp. nov., a nitrogen-fixing bacterium isolated from a mangrove sediment, and proposal of Prolixibacteraceae fam. nov. International Journal of Systematic and Evolutionary Microbiology 64:875–881. https://doi.org/10.1099/ijs.0.052779-0

Hughes LC, Somoza GM, Nguyen BN, et al (2017) Transcriptomic differentiation underlying marine-to-freshwater transitions in the South American silversides *Odontesthes argentinensis* and *O. bonariensis* (Atheriniformes). Ecology and Evolution 7:5258–5268. https://doi.org/10.1002/ece3.3133

Hügler M, Gärtner A, Imhoff JF (2010) Functional genes as markers for sulfur cycling and CO2 fixation in microbial communities of hydrothermal vents of the Logatchev field. FEMS Microbiology Ecology 73:526–537. https://doi.org/10.1111/j.1574-6941.2010.00919.x

Hüppop K (1985) The role of metabolism in the evolution of cave animals. Natl Speleol Soc Bull 47:136–146.

Hüppop K (1987) Food-finding ability in cave fish (Astyanax fasciatus). International Journal of Speleology 16. http://dx.doi.org/10.5038/1827-806X.16.1.4

Hüppop K (1986) Oxygen consumption of *Astyanax fasciatus* (Characidae, Pisces): a comparison of epigean and hypogean populations. Environmental Biology of Fishes 17:299–308. https://doi.org/10.1007/BF00001496

Imarazene B, Beille S, Jouanno E, et al (2020) Primordial Germ Cell Migration and Histological and Molecular Characterization of Gonadal Differentiation in Pachón Cavefish *Astyanax mexicanus*. SXD 14:80–98. https://doi.org/10.1159/000513378

Imarazene B, Du K, Beille S, et al (2021) A supernumerary “B-sex” chromosome drives male sex determination in the Pachón cavefish, *Astyanax mexicanus*. Current Biology. https://doi.org/10.1016/j.cub.2021.08.030

Ip JC-H, Xu T, Sun J, et al (2021) Host–Endosymbiont Genome Integration in a Deep-Sea Chemosymbiotic Clam. Molecular Biology and Evolution 38:502–518. https://doi.org/10.1093/molbev/msaa241

Jeffery WR (2009) Regressive Evolution in *Astyanax* Cavefish. Annual Review of Genetics 43:25–47. https://doi.org/10.1146/annurev-genet-102108-134216

Jiang L, Liu X, Dong C, et al (2020) "Candidatus *Desulfobulbus rimicarensis*," an Uncultivated Deltaproteobacterial Epibiont from the Deep-Sea Hydrothermal Vent Shrimp *Rimicaris exoculata*. Applied and Environmental Microbiology 86:e02549-19. https://doi.org/10.1128/AEM.02549-19

Johnson JB, Omland KS (2004) Model selection in ecology and evolution. Trends in Ecology & Evolution 19:101–108. https://doi.org/10.1016/j.tree.2003.10.013

Jorquera MA, Maruyama F, Ogram AV, et al (2016) Rhizobacterial Community Structures Associated with Native Plants Grown in Chilean Extreme Environments. Microbial Ecology 72:633–646. https://doi.org/10.1007/s00248-016-0813-x

Kaeuffer R, Peichel CL, Bolnick DI, Hendry AP (2012) Parallel and Nonparallel Aspects of Ecological, Phenotypic, and Genetic Divergence Across Replicate Population Pairs of Lake and Stream Stickleback. Evolution 66:402–418. https://doi.org/10.1111/j.1558-5646.2011.01440.x

Kashinskaya EN, Simonov EP, Kabilov MR, et al (2018) Diet and other environmental factors shape the bacterial communities of fish gut in an eutrophic lake. Journal of Applied Microbiology 125:1626–1641. https://doi.org/10.1111/jam.14064

Kasumyan AO, Marusov EA (2015) Chemoorientation in the feeding behavior of the blind Mexican cavefish *Astyanax fasciatus* (Characidae, Teleostei). Russian Journal of Ecology 46:559–563. https://doi.org/10.1134/S1067413615060053

Keene A, Yoshizawa M, McGaugh SE (2015) Biology and Evolution of the Mexican Cavefish. Academic Press.

Kelley JL, Arias-Rodriguez L, Patacsil Martin D, et al (2016) Mechanisms Underlying Adaptation to Life in Hydrogen Sulfide–Rich Environments. Molecular Biology and Evolution 33:1419–1434. https://doi.org/10.1093/molbev/msw020

Kim PS, Shin N-R, Lee J-B, et al (2021) Host habitat is the major determinant of the gut microbiome of fish. Microbiome 9:166. https://doi.org/10.1186/s40168-021-01113-x

Kim Y-O, Park I-S, Kim D-G, et al (2023) Enterovibrio paralichthyis sp. nov., isolated from the gut of an olive flounder Paralichthys olivaceus. International Journal of Systematic and Evolutionary Microbiology 72:005593. https://doi.org/10.1099/ijsem.0.005593

Konec M, Prevorčnik S, Sarbu SM, et al (2015) Parallels between two geographically and ecologically disparate cave invasions by the same species, *Asellus aquaticus* (Isopoda, Crustacea). Journal of Evolutionary Biology 28:864–875. https://doi.org/10.1111/jeb.12610

Kowalko J (2020) Utilizing the blind cavefish *Astyanax mexicanus* to understand the genetic basis of behavioral evolution. Journal of Experimental Biology 223:. https://doi.org/10.1242/jeb.208835

Kowalko JE, Rohner N, Linden TA, et al (2013) Convergence in feeding posture occurs through different genetic loci in independently evolved cave populations of *Astyanax mexicanus*. PNAS 110:16933–16938. https://doi.org/10.1073/pnas.1317192110

Kramer DL, Bryant MJ (1995) Intestine length in the fishes of a tropical stream: 2. Relationships to diet — the long and short of a convoluted issue. Environmental Biology of Fishes 42:129–141. https://doi.org/10.1007/BF00001991

Kruse S, Goris T, Westermann M, et al (2018) Hydrogen production by *Sulfurospirillum* species enables syntrophic interactions of Epsilonproteobacteria. Nature Communications 9:4872. https://doi.org/10.1038/s41467-018-07342-3

Kruse T, Ratnadevi CM, Erikstad H-A, Birkeland N-K (2019) Complete genome sequence analysis of the thermoacidophilic verrucomicrobial methanotroph "Candidatus Methylacidiphilum kamchatkense" strain Kam1 and comparison with its closest relatives. BMC Genomics 20:642. https://doi.org/10.1186/s12864-019-5995-4

Kuczynski J, Stombaugh J, Walters WA, et al (2012) Using QIIME to Analyze 16S rRNA Gene Sequences from Microbial Communities. Current Protocols in Microbiology 0 1:Unit-1E.5. https://doi.org/10.1002/9780471729259.mc01e05s27

Lachapelle J, Reid J, Colegrave N (2015) Repeatability of adaptation in experimental populations of different sizes. Proceedings of the Royal Society B: Biological Sciences 282:20143033. https://doi.org/10.1098/rspb.2014.3033

Laporte M, Rogers SM, Dion-Côté A-M, et al (2015) RAD-QTL Mapping Reveals Both Genome-Level Parallelism and Different Genetic Architecture Underlying the Evolution of Body Shape in Lake Whitefish (*Coregonus clupeaformis*) Species Pairs. G3 Genes|Genomes|Genetics 5:1481–1491. https://doi.org/10.1534/g3.115.019067

Lavy A, Keren R, Yu K, et al (2018) A novel Chromatiales bacterium is a potential sulfide oxidizer in multiple orders of marine sponges. Environmental Microbiology 20:800–814. https://doi.org/10.1111/1462-2920.14013

Lee W-K, Juniper SK, Perez M, et al (2021) Diversity and characterization of bacterial communities of five co-occurring species at a hydrothermal vent on the Tonga Arc. Ecology and Evolution 11:4481–4493. https://doi.org/10.1002/ece3.7343

Legrand TPRA, Wynne JW, Weyrich LS, Oxley APA (2020) A microbial sea of possibilities: current knowledge and prospects for an improved understanding of the fish microbiome. Reviews in Aquaculture 12:1101–1134. https://doi.org/10.1111/raq.12375

Li J, Wei X, Huang D, Xiao J (2022) The Phylosymbiosis Pattern Between the Fig Wasps of the Same Genus and Their Associated Microbiota. Frontiers in Microbiology 12. https://doi.org/10.3389/fmicb.2021.800190

Li L, Wang M, Li L, et al (2020) Endosymbionts of Metazoans Dwelling in the PACManus Hydrothermal Vent: Diversity and Potential Adaptive Features Revealed by Genome Analysis. Applied and Environmental Microbiology 86:e00815-20. https://doi.org/10.1128/AEM.00815-20

Li S-J, Hua Z-S, Huang L-N, et al (2014a) Microbial communities evolve faster in extreme environments. Scientific Reports 4:6205. https://doi.org/10.1038/srep06205

Li XM, Zhu YJ, Yan QY, et al (2014b) Do the intestinal microbiotas differ between paddlefish (*Polyodon spathala*) and bighead carp (*Aristichthys nobilis*) reared in the same pond? Journal of Applied Microbiology 117:1245–1252. https://doi.org/10.1111/jam.12626

Li Y, Tassia MG, Waits DS, et al (2019) Genomic adaptations to chemosymbiosis in the deep-sea seep-dwelling tubeworm *Lamellibrachia luymesi*. BMC Biology 17:91. https://doi.org/10.1186/s12915-019-0713-x

Lin H-J, Kao W-Y, Wang Y-T (2007) Analyses of stomach contents and stable isotopes reveal food sources of estuarine detritivorous fish in tropical/subtropical Taiwan. Estuarine, Coastal and Shelf Science 73:527–537. https://doi.org/10.1016/j.ecss.2007.02.013

Liu H, Guo X, Gooneratne R, et al (2016) The gut microbiome and degradation enzyme activity of wild freshwater fishes influenced by their trophic levels. Scientific Reports 6:24340. https://doi.org/10.1038/srep24340

López-García P, Gaill F, Moreira D (2002) Wide bacterial diversity associated with tubes of the vent worm *Riftia pachyptila*. Environmental Microbiology 4:204–215. https://doi.org/10.1046/j.1462-2920.2002.00286.x

Lukas J, Auer F, Goldhammer T, et al (2021) Diurnal Changes in Hypoxia Shape Predator-Prey Interaction in a Bird-Fish System. Frontiers in Ecology and Evolution 9. https://doi.org/10.3389/fevo.2021.619193

Ma Q, Wang X, Li L-Y, et al (2021) High protein intake promotes the adaptation to chronic hypoxia in zebrafish (*Danio rerio*). Aquaculture 535:736356. https://doi.org/10.1016/j.aquaculture.2021.736356

Mahler DL, Weber MG, Wagner CE, Ingram T (2017) Pattern and Process in the Comparative Study of Convergent Evolution. The American Naturalist 190:S13–S28. https://doi.org/10.1086/692648

Mahmoud MAA, Magdy M (2021) Metabarcoding profiling of microbial diversity associated with trout fish farming. Scientific Reports 11:421. https://doi.org/10.1038/s41598-020-80236-x

Mallott EK, Amato KR (2021) Host specificity of the gut microbiome. Nature Reviews in Microbiology 19:639–653. https://doi.org/10.1038/s41579-021-00562-3

Marangon E, Uthicke S, Patel F, et al (2023) Life-stage specificity and cross-generational climate effects on the microbiome of a tropical sea urchin (Echinodermata: Echinoidea). Molecular Ecology 32:5645–5660. https://doi.org/10.1111/mec.17124

Mawad AMM, Hassanein M, Aldaby ES, Yousef N (2021) Desulphurisation kinetics of thiophenic compound by sulphur oxidizing *Klebsiella oxytoca* SOB-1. Journal of Applied Microbiology 130:1181–1191. https://doi.org/10.1111/jam.14829

Mazel F, Davis KM, Loudon A, et al (2018) Is Host Filtering the Main Driver of Phylosymbiosis across the Tree of Life? mSystems 3:10.1128/msystems.00097-18. https://doi.org/10.1128/msystems.00097-18

McArdle BH, Anderson ML (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82:290–297. https://doi.org/10.1890/0012-9658(2001)082[0290:FMMTCD]2.0.CO;2

McFall-Ngai M, Hadfield MG, Bosch TCG, et al (2013) Animals in a bacterial world, a new imperative for the life sciences. Proceedings of the National Academy of Sciences 110:3229–3236. https://doi.org/10.1073/pnas.1218525110

McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS One 8:e61217. https://doi.org/10.1371/journal.pone.0061217

Miller RR, Minckley WL, Norris SM (2005) Freshwater fishes of Mexico. University of Chicago Press, Chicago, IL, USA.

Millero FJ (1986) The thermodynamics and kinetics of the hydrogen sulfide system in natural waters. Marine Chemistry 18:121–147. https://doi.org/10.1016/0304-4203(86)90003-4

Mitchell RW, Elliott WR, Russell WH (1977) Mexican eyeless characin fishes, genus *Astyanax*: environment, distribution, and evolution. Texas Tech Press 12:1–89.

Moore JC, Berlow EL, Coleman DC, et al (2004) Detritus, trophic dynamics and biodiversity. Ecology Letters 7:584–600. https://doi.org/10.1111/j.1461-0248.2004.00606.x

Moran D, Softley R, Warrant EJ (2014) Eyeless Mexican Cavefish Save Energy by Eliminating the Circadian Rhythm in Metabolism. PLOS One 9. https://doi.org/10.1371/journal.pone.0107877

Moran NA (2001) The Coevolution of Bacterial Endosymbionts and Phloem-Feeding Insects. Annals of the Missouri Botanical Garden 88:35–44. https://doi.org/10.2307/2666130

Morshed SM, Lee T-H (2023) The role of the microbiome on fish mucosal immunity under changing environments. Fish & Shellfish Immunology 139:108877. https://doi.org/10.1016/j.fsi.2023.108877

Natri HM, Merilä J, Shikano T (2019) The evolution of sex determination associated with a chromosomal inversion. Nature Communications 10:145. https://doi.org/10.1038/s41467-018-08014-y

O'Brien PA, Andreakis N, Tan S, et al (2021) Testing cophylogeny between coral reef invertebrates and their bacterial and archaeal symbionts. Molecular Ecology 30:3768–3782. https://doi.org/10.1111/mec.16006

Oksanen J, Blanchet FG, Friendly M, et al (2020) vegan: Community Ecology Package. https://cran.r-project.org/package=vegan

Olsen L, Levy M, Medley JK, et al (2023) Metabolic reprogramming underlies cavefish muscular endurance despite loss of muscle mass and contractility. Proceedings of the National Academy of Sciences 120:e2204427120. https://doi.org/10.1073/pnas.2204427120

Osman EO, Weinnig AM (2022) Microbiomes and Obligate Symbiosis of Deep-Sea Animals. Annual Review of Animal Biosciences 10:151–176. https://doi.org/10.1146/annurev-animal-081621-112021

Palacios M, Voelker G, Arias Rodriguez L, et al (2016) Phylogenetic analyses of the subgenus Mollienesia (*Poecilia*, Poeciliidae, Teleostei) reveal taxonomic inconsistencies, cryptic biodiversity, and spatio-temporal aspects of diversification in Middle America. Molecular Phylogenetics and Evolution 103:230–244. https://doi.org/10.1016/j.ympev.2016.07.025

Passow CN, Arias-Rodriguez L, Tobler M (2017) Convergent evolution of reduced energy demands in extremophile fish. PLOS One 12:e0186935. https://doi.org/10.1371/journal.pone.0186935

Passow CN, Greenway R, Arias-Rodriguez L, et al (2015) Reduction of Energetic Demands through Modification of Body Size and Routine Metabolic Rates in Extremophile Fish. Physiological and Biochemical Zoology 88:371–383. https://doi.org/10.1086/681053

Pease AA, Capps KA, Castillo MM, et al (2023) Chapter 23 - Rivers of Mexico. In: Delong MD, Jardine TD, Benke AC, Cushing CE (eds) Rivers of North America (Second Edition). Academic Press, San Diego, pp 974–1024.

Pérez-Pascual D, Vendrell-Fernández S, Audrain B, et al (2021) Gnotobiotic rainbow trout (*Oncorhynchus mykiss*) model reveals endogenous bacteria that protect against Flavobacterium columnare infection. PLOS Pathogens 17:e1009302. https://doi.org/10.1371/journal.ppat.1009302

Pérez-Rodríguez R, Esquivel-Bobadilla S, Orozco-Ruíz AM, Olivas-Hernández JL, García-De-León FJ (2021) Genetic structure and historical and contemporary gene flow of *Astyanax*

mexicanus in the Gulf of Mexico slope: a microsatellite-based analysis. PeerJ 9:e10784. https://doi.org/10.7717/peerj.10784

Petersen C, Hamerich IK, Adair KL, et al (2023) Host and microbiome jointly contribute to environmental adaptation. The ISME Journal 17:1953–1965. https://doi.org/10.1038/s41396-023-01507-9

Petersen JM, Ramette A, Lott C, et al (2010) Dual symbiosis of the vent shrimp Rimicaris exoculata with filamentous gamma- and epsilonproteobacteria at four Mid-Atlantic Ridge hydrothermal vent fields. Environmental Microbiology 12:2204–2218. https://doi.org/10.1111/j.1462-2920.2009.02129.x

Pfenninger M, Lerp H, Tobler M, et al (2014) Parallel evolution of cox genes in H2S-tolerant fish as key adaptation to a toxic environment. Nature Communications 5:3873. https://doi.org/10.1038/ncomms4873

Plath M, Tobler M, Riesch R, et al (2007) Survival in an extreme habitat: the roles of behaviour and energy limitation. Naturwissenschaften 94:991–996. https://doi.org/10.1007/s00114-007-0279-2

Podschun R, Pietsch S, Höller C, Ullmann U (2001) Incidence of *Klebsiella* Species in Surface Waters and Their Expression of Virulence Factors. Applied and Environmental Microbiology 67:3325–3327. https://doi.org/10.1128/AEM.67.7.3325-3327.2001

Pollux BJA, Meredith RW, Springer MS, et al (2014) The evolution of the placenta drives a shift in sexual selection in livebearing fish. Nature 513:233–6. https://doi.org/10.1038/nature13451

Porter ML, Crandall KA (2003) Lost along the way: the significance of evolution in reverse. Trends in Ecology & Evolution 18:541–547. https://doi.org/10.1016/S0169-5347(03)00244-1

Powell MA, Somero GN (1986) Adaptations to sulfide by hydrothermal vent animals: sites and mechanisms of detoxification and metabolism. The Biological Bulletin 171:274–290. https://doi.org/10.2307/1541923

Powell R (2007) Is convergence more than an analogy? Homoplasy and its implications for macroevolutionary predictability. Biology & Philosophy 22:565–578. https://doi.org/10.1007/s10539-006-9057-3

Premate E, Borko Š, Kralj-Fišer S, et al (2021) No room for males in caves: Female-biased sex ratio in subterranean amphipods of the genus *Niphargus*. Journal of Evolutionary Biology 34: 1653–1661. https://doi.org/10.1111/jeb.13917

Qi L, Lian C-A, Zhu F-C, et al (2022) Comparative Analysis of Intestinal Microflora Between Two Developmental Stages of *Rimicaris kairei*, a Hydrothermal Shrimp From the Central Indian Ridge. Frontiers in Microbiology 12:802888. https://doi.org/10.3389/fmicb.2021.802888

Qi X, Zhang Y, Zhang Y, et al (2023) Vitamin B12 produced by *Cetobacterium somerae* improves host resistance against pathogen infection through strengthening the interactions within gut microbiota. Microbiome 11:135. https://doi.org/10.1186/s40168-023-01574-2

Qiu Y-L, Kuang X, Shi X, et al (2014) *Terrimicrobium sacchariphilum* gen. nov., sp. nov., an anaerobic bacterium of the class 'Spartobacteria' in the phylum Verrucomicrobia, isolated from a rice paddy field. International Journal of Systematic and Evolutionary Microbiology 64:1718–1723. https://doi.org/10.1099/ijs.0.060244-0

R Core Team 2020 R: The R Project for Statistical Computing. https://www.r-project.org/.

Ramírez C, Coronado J, Silva A, Romero J (2018) *Cetobacterium* Is a Major Component of the Microbiome of Giant Amazonian Fish (*Arapaima gigas*) in Ecuador. Animals 8:189. https://doi.org/10.3390/ani8110189

Ramoneda J, Stallard-Olivera E, Hoffert M, et al (2023) Building a genome-based understanding of bacterial pH preferences. Science Advances 9:eadf8998. https://doi.org/10.1126/sciadv.adf8998

Rampelotto PH (2013) Extremophiles and Extreme Environments. Life 3:482–485. https://doi.org/10.3390/life3030482

Ravin NV, Rossetti S, Beletsky AV, et al (2022) Two New Species of Filamentous Sulfur Bacteria of the Genus *Thiothrix*, *Thiothrix winogradskyi* sp. nov. and 'Candidatus *Thiothrix sulfatifontis*' sp. nov. Microorganisms 10:1300. https://doi.org/10.3390/microorganisms10071300

Ray A k., Ghosh K, Ringø E (2012) Enzyme-producing bacteria isolated from fish gut: a review. Aquaculture Nutrition 18:465–492. https://doi.org/10.1111/j.1365-2095.2012.00943.x

Reichard M, Polačik M, Blažek R, Vrtílek M (2014) Female bias in the adult sex ratio of African annual fishes: interspecific differences, seasonal trends and environmental predictors. Evolutionary Ecology 28:1105–1120. https://doi.org/10.1007/s10682-014-9732-9

Rétaux S, Casane D (2013) Evolution of eye development in the darkness of caves: adaptation, drift, or both? EvoDevo 4:26. https://doi.org/10.1186/2041-9139-4-26

Riddle M, Nguyen NK, Nave M, et al (2023) Evolution of the Mexican Cavefish Gut Microbiome. Current Biology. http://dx.doi.org/10.2139/ssrn.4520069

Riddle MR, Aspiras AC, Gaudenz K, et al (2018) Insulin resistance in cavefish as an adaptation to a nutrient-limited environment. Nature 555:647–651. https://doi.org/10.1038/nature26136

Riddle MR, Damen F, Aspiras A, et al (2019) Evolution of the gastrointestinal tract morphology and plasticity in cave-adapted Mexican tetra, *Astyanax mexicanus*. bioRxiv 852814. https://doi.org/10.1101/852814

Riesch R, Colston TJ, Joachim BL, Schlupp I (2011) Natural history and life history of the Grijalva gambusia *Heterophallus milleri* Radda, 1987 (Teleostei: Poeciliidae). aqua: International Journal of Ichthyology 17:95–103

Riesch R, Martin RA, Bierbach D, et al (2012) Natural history, life history, and diet of *Priapella chamulae* Schartl, Meyer & Wilde 2006 (Teleostei: Poeciliidae). aqua: International Journal of Ichthyology 18:95–103.

Riesch R, Morley NJ, Jourdan J, et al (2020) Sulphide-toxic habitats are not refuges from parasite infections in an extremophile fish. Acta Oecologica 106:103602. https://doi.org/10.1016/j.actao.2020.103602

Roach KA, Tobler M, Winemiller KO (2011) Hydrogen sulfide, bacteria, and fish: a unique, subterranean food chain. Ecology 92:2056–2062. https://doi.org/10.1890/11-0276.1

Rohlfs RV, Harrigan P, Nielsen R (2014) Modeling Gene Expression Evolution with an Extended Ornstein–Uhlenbeck Process Accounting for Within-Species Variation. Molecular Biology and Evolution 31:201–211. https://doi.org/10.1093/molbev/mst190

Rohlfs RV, Nielsen R (2015) Phylogenetic ANOVA: The Expression Variance and Evolution Model for Quantitative Trait Evolution. Systematic Biology 64:695–708. https://doi.org/10.1093/sysbio/syv042

Rosenblum EB, Parent CE, Brandt EE (2014) The Molecular Basis of Phenotypic Convergence. Annual Review of Ecology, Evolution and Systematics 45:203–226. https://doi.org/10.1146/annurev-ecolsys-120213-091851

Rothschild LJ, Mancinelli RL (2001) Life in extreme environments. Nature 409:1092–1101. https://doi.org/10.1038/35059215

Ruiz A, Gisbert E, Andree KB (2024) Impact of the diet in the gut microbiota after an inter-species microbial transplantation in fish. Scientific Reports 14:4007. https://doi.org/10.1038/s41598-024-54519-6

Sadeghi J, Chaganti SR, Johnson TB, Heath DD (2023) Host species and habitat shape fish-associated bacterial communities: phylosymbiosis between fish and their microbiome. Microbiome 11:258. https://doi.org/10.1186/s40168-023-01697-6

Salin K, Voituron Y, Mourin J, Hervant F (2010) Cave colonization without fasting capacities: An example with the fish *Astyanax fasciatus mexicanus*. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology 156:451–457. https://doi.org/10.1016/j.cbpa.2010.03.030

Salinas I, Magadán S (2017) Omics in fish mucosal immunity. Developmental & Comparative Immunology 75:99–108. https://doi.org/10.1016/j.dci.2017.02.010

Santini F, Nguyen MTT, Sorenson L, et al (2013) Do habitat shifts drive diversification in teleost fishes? An example from the pufferfishes (Tetraodontidae). Journal of Evolutionary Biology 26:1003–1018. https://doi.org/10.1111/jeb.12112

Schluter D (2000) The Ecology of Adaptive Radiation. OUP Oxford.

Secor SM (2009) Specific dynamic action: a review of the postprandial metabolic response. Journal of Comparative Physiology B 179:1–56. https://doi.org/10.1007/s00360-008-0283-7

Sehnal L, Brammer-Robbins E, Wormington AM, et al (2021) Microbiome Composition and Function in Aquatic Vertebrates: Small Organisms Making Big Impacts on Aquatic Animal Health. Frontiers in Microbiology 12. https://doi.org/10.3389/fmicb.2021.567408

Sevellec M, Derome N, Bernatchez L (2018) Holobionts and ecological speciation: the intestinal microbiota of lake whitefish species pairs. Microbiome 6:47. https://doi.org/10.1186/s40168-018-0427-2

Sevellec M, Laporte M, Bernatchez A, et al (2019) Evidence for host effect on the intestinal microbiota of whitefish (*Coregonus* sp.) species pairs and their hybrids. Ecology and Evolution 9:11762–11774. https://doi.org/10.1002/ece3.5676

Shi F, Huang Y, Yang M, et al (2022) Antibiotic-induced alternations in gut microflora are associated with the suppression of immune-related pathways in grass carp (Ctenopharyngodon idellus). Frontiers in Immunology 13. https://doi.org/10.3389/fimmu.2022.970125

Shu W-S, Huang L-N (2022) Microbial diversity in extreme environments. Nature Reviews in Microbiology 20:219–235. https://doi.org/10.1038/s41579-021-00648-y

Sierra MA, Ryon KA, Tierney BT, et al (2022) Microbiome and metagenomic analysis of Lake Hillier Australia reveals pigment-rich polyextremophiles and wide-ranging metabolic adaptations. Environmental Microbiome 17:60. https://doi.org/10.1186/s40793-022-00455-9

Simon V, Elleboode R, Mahé K, et al (2017) Comparing growth in surface and cave morphs of the species *Astyanax mexicanus*: insights from scales. EvoDevo 8:23. https://doi.org/10.1186/s13227-017-0086-6

Singh P, Jain K, Desai C, et al (2019) Chapter 18 - Microbial Community Dynamics of Extremophiles/Extreme Environment. In: Das S, Dash HR (eds) Microbial Diversity in the Genomic Era. Academic Press, pp 323–332.

Small CM, Beck EA, Currey MC, et al (2023) Host genomic variation shapes gut microbiome diversity in threespine stickleback fish. mBio 14:e00219-23. https://doi.org/10.1128/mbio.00219-23

Smith CC, Snowberg LK, Gregory Caporaso J, et al (2015) Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. ISME J 9:2515–2526. https://doi.org/10.1038/ismej.2015.64

Smith P, Willemsen D, Popkes M, et al (2017) Regulation of life span by the gut microbiota in the short-lived African turquoise killifish. eLife 6:e27014. https://doi.org/10.7554/eLife.27014

Soares D, Niemiller ML (2013) Sensory Adaptations of Fishes to Subterranean Environments. BioScience 63:274–283. https://doi.org/10.1525/bio.2013.63.4.7

Sorokin DY, Tikhonova TV, Koch H, et al (2023) *Trichlorobacter ammonificans*, a dedicated acetate-dependent ammonifier with a novel module for dissimilatory nitrate reduction to ammonia. The ISME Journal 17:1639–1648. https://doi.org/10.1038/s41396-023-01473-2

South A (2017). rnaturalearth: World map data from Natural Earth. R package version 0.1.0. https://CRAN.R-project.org/package=rnaturalearth

Stayton CT (2015) The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. Evolution 69:2140–2153. https://doi.org/10.1111/evo.12729

Sugita H, Miyajima C, Deguchi Y (1991) The vitamin B12-producing ability of the intestinal microflora of freshwater fish. Aquaculture 92:267–276. https://doi.org/10.1016/0044-8486(91)90028-6

Sullam KE, Essinger SD, Lozupone CA, et al (2012) Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. Molecular Ecology 21:3363–3378. https://doi.org/10.1111/j.1365-294X.2012.05552.x

Sullam KE, Rubin BE, Dalton CM, et al (2015) Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. ISME J 9:1508–1522. https://doi.org/10.1038/ismej.2014.231

Sun S, Yang M, Fu H, et al (2020) Altered intestinal microbiota induced by chronic hypoxia drives the effects on lipid metabolism and the immune response of oriental river prawn *Macrobrachium nipponense*. Aquaculture 526:735431. https://doi.org/10.1016/j.aquaculture.2020.735431

Sun Y, Wang M, Zhong Z, et al (2021) Adaption to hydrogen sulfide-rich environments: Strategies for active detoxification in deep-sea symbiotic mussels, *Gigantidas platifrons*. Science of The Total Environment 150054. https://doi.org/10.1016/j.scitotenv.2021.150054

Sundarraman D, Hay EA, Martins DM, et al (2020) Higher-Order Interactions Dampen Pairwise Competition in the Zebrafish Gut Microbiome. mBio 11:10.1128/mbio.01667-20. https://doi.org/10.1128/mbio.01667-20

Suzuki TA (2017) Links between Natural Variation in the Microbiome and Host Fitness in Wild Mammals. Integrative and Comparative Biology 57:756–769. https://doi.org/10.1093/icb/icx104

Sylvain F-É, Cheaib B, Llewellyn M, et al (2016) pH drop impacts differentially skin and gut microbiota of the Amazonian fish tambaqui (*Colossoma macropomum*). Scientific Reports 6:32032. https://doi.org/10.1038/srep32032

Sylvain F-É, Holland A, Audet-Gilbert É, et al (2019) Amazon fish bacterial communities show structural convergence along widespread hydrochemical gradients. Molecular Ecology 28:3612–3626. https://doi.org/10.1111/mec.15184

Sylvain F-É, Holland A, Bouslama S, et al (2020) Fish Skin and Gut Microbiomes Show Contrasting Signatures of Host Species and Habitat. Applied Environmental Microbiology 86. https://doi.org/10.1128/AEM.00789-20

Tapia-Paniagua ST, Fumanal M, Anguís V, et al (2019) Modulation of Intestinal Microbiota in *Solea senegalensis* Fed Low Dietary Level of *Ulva ohnoi*. Frontiers in Microbiology 10. https://doi.org/10.3389/fmicb.2019.00171

Tarnecki A m., Burgos F a., Ray C l., Arias C r. (2017) Fish intestinal microbiome: diversity and symbiosis unravelled by metagenomics. Journal of Applied Microbiology 123:2–17. https://doi.org/10.1111/jam.13415

Thompson FL, Hoste B, Thompson CC, et al (2002) *Enterovibrio norvegicus* gen. nov., sp. nov., isolated from the gut of turbot (*Scophthalmus maximus*) larvae: a new member of the family Vibrionaceae. International Journal of Systematic and Evolutionary Microbiology 52:2015–2022. https://doi.org/10.1099/00207713-52-6-2015

Tobler M (2008) Divergence in trophic ecology characterizes colonization of extreme habitats. Biological Journal of the Linnean Society 95:517–528. https://doi.org/10.1111/j.1095-8312.2008.01063.x

Tobler M, Greenway R, Kelley JL (2021) Ecology drives the degree of convergence in the gene expression of extremophile fishes. bioxRiv 2021.12.13.472416. https://doi.org/10.1101/2021.12.13.472416

Tobler M, Kelley JL, Plath M, Riesch R (2018) Extreme environments and the origins of biodiversity: Adaptation and speciation in sulphide spring fishes. Molecular Ecology 27:843–859. https://doi.org/10.1111/mec.14497

Tobler M, Passow CN, Greenway R, et al (2016) The Evolutionary Ecology of Animals Inhabiting Hydrogen Sulfide–Rich Environments. Annual Review of Ecology, Evolution, and Systematics 47:239–262. https://doi.org/10.1146/annurev-ecolsys-121415-032418

Tobler M, Plath M, Riesch R, et al (2014) Selection from parasites favours immunogenetic diversity but not divergence among locally adapted host populations. Journal of Evolutionary Biology 27:960–974. https://doi.org/10.1111/jeb.12370

Tobler M, Riesch RW, Tobler CM, Plath M (2009) Compensatory behaviour in response to sulphide-induced hypoxia affects time budgets, feeding efficiency, and predation risk. Evolutionary Ecology Research 11:935–948

Tobler M, Scharnweber K, Greenway R, et al (2015) Convergent changes in the trophic ecology of extremophile fish along replicated environmental gradients. Freshwater Biology 60:768–780. https://doi.org/10.1111/fwb.12530

Tobler M, Schlupp I, Heubel KU, et al (2006) Life on the edge: hydrogen sulfide and the fish communities of a Mexican cave and surrounding waters. Extremophiles 10:577–585. https://doi.org/10.1007/s00792-006-0531-2

Tokuda G, Yamada A, Nakano K, et al (2008) Colonization of *Sulfurovum* sp. on the gill surfaces of *Alvinocaris longirostris*, a deep-sea hydrothermal vent shrimp. Marine Ecology 29:106–114. https://doi.org/10.1111/j.1439-0485.2007.00211.x

Trajano E (2001) Ecology of Subterranean Fishes: An Overview. Environmental Biology of Fishes 62:133–160. https://doi.org/10.1023/A:1011841913569

Tsuchiya C, Sakata T, Sugita H (2008) Novel ecological niche of *Cetobacterium somerae*, an anaerobic bacterium in the intestinal tracts of freshwater fish. Letters in Applied Microbiology 46:43–48. https://doi.org/10.1111/j.1472-765X.2007.02258.x

Tyagi A, Singh B, et al (2019) Shotgun metagenomics offers novel insights into taxonomic compositions, metabolic pathways and antibiotic resistance genes in fish gut microbiome. Archives of Microbiology 201:295–303. http://dx.doi.org/10.1007/s00203-018-1615-y

Van den Abbeele P, Van de Wiele T, Verstraete W, Possemiers S (2011) The host selects mucosal and luminal associations of coevolved gut microorganisms: a novel concept. FEMS Microbiology Reviews 35:681–704. https://doi.org/10.1111/j.1574-6976.2011.00270.x

Velotta JP, McCormick SD, Whitehead A, et al (2022) Repeated Genetic Targets of Natural Selection Underlying Adaptation of Fishes to Changing Salinity. Integrative and Comparative Biology icac072. https://doi.org/10.1093/icb/icac072

Vrijenhoek RC (2013) On the instability and evolutionary age of deep-sea chemosynthetic communities. Deep Sea Research Part II: Topical Studies in Oceanography 92:189–200. https://doi.org/10.1016/j.dsr2.2012.12.004

Wang C, Li P, Guo L, et al (2022a) A new potential risk: The impacts of *Klebsiella pneumoniae* infection on the histopathology, transcriptome and metagenome of *Chinese mitten crab* (*Eriocheir sinensis*). Fish & Shellfish Immunology 131:918–928. https://doi.org/10.1016/j.fsi.2022.11.010

Wang J, Ye J, Zhang Z, et al (2023) Comparison of the nutrient value, nonspecific immunity, and intestinal microflora of red swamp crayfish (*Procambarus clarkii*) in different culture modes. Aquaculture Reports 31:101683. https://doi.org/10.1016/j.aqrep.2023.101683

Wang Y, Bi H-Y, Chen H-G, et al (2022b) Metagenomics Reveals Dominant Unusual Sulfur Oxidizers Inhabiting Active Hydrothermal Chimneys From the Southwest Indian Ridge. Frontiers in Microbiology 13:861795. https://doi.org/10.3389/fmicb.2022.861795

Wang Y, Guo B (2019) Adaption to extreme environments: a perspective from fish genomics. Reviews in Fish Biology and Fisheries 29:735–747. https://doi.org/10.1007/s11160-019-09577-9

Weisse L, Héchard Y, Moumen B, Delafont V (2023) Here, there and everywhere: Ecology and biology of the Dependentiae phylum. Environmental Microbiology 25:597–605. https://doi.org/10.1111/1462-2920.16307

Whelan CJ, Brown JS (2005) Optimal foraging and gut constraints: reconciling two schools of thought. Oikos 110:481–496. https://doi.org/10.1111/j.0030-1299.2005.13387.x

Wickham H (2011) ggplot2. WIREs Computational Statistics 3:180–185. https://doi.org/10.1002/wics.147

Wilkens H, Burns RJ (1972) A new *Anoptichthys* cave population (Characidae, Pisces). Ann Spéléol 27:263–270.

Winemiller KO (1993) Seasonality of Reproduction by Livebearing Fishes in Tropical Rainforest Streams. Oecologia 95:266–276.

Wong S, Rawls JF (2012) Intestinal microbiota composition in fishes is influenced by host ecology and environment. Molecular Ecology 21:3100–3102. https://doi.org/10.1111/j.1365-294X.2012.05646.x

Wu Z, Zhang Q, Yang J, et al (2022) Significant alterations of intestinal symbiotic microbiota induced by intraperitoneal vaccination mediate changes in intestinal metabolism of NEW Genetically Improved Farmed Tilapia (NEW GIFT, Oreochromis niloticus). Microbiome 10:221. https://doi.org/10.1186/s40168-022-01409-6

Xia Y, Wang G, Yu E, et al (2022) Addition of berberine to formulated feed changes the glucose utilisation, intestinal microbiota and serum metabolites of Largemouth bass (*Micropterus salmoides*). Aquaculture Reports 23:101018. https://doi.org/10.1016/j.aqrep.2022.101018

Xie C-J, Tang R, Yang S, et al (2023) A novel nitrogen-fixing bacterium, *Propionivibrio soli* sp. nov. isolated from paddy soil. Archives of Microbiology 205:68. https://doi.org/10.1007/s00203-023-03413-2

Xiong J-B, Nie L, Chen J (2019) Current understanding on the roles of gut microbiota in fish disease and immunity. Zoological Research 40:70–76. https://doi.org/10.24272/j.issn.2095-8137.2018.069

Xiong S, Krishnan J, Peuß R, Rohner N (2018) Early adipogenesis contributes to excess fat accumulation in cave populations of *Astyanax mexicanus*. Developmental Biology 441:297–304. https://doi.org/10.1016/j.ydbio.2018.06.003

Xu J, Li Q, Xu L, et al (2013) Gene expression changes leading extreme alkaline tolerance in Amur ide (*Leuciscus waleckii*) inhabiting soda lake. BMC Genomics 14:682. https://doi.org/10.1186/1471-2164-14-682

Yoshida K, Terai Y, Mizoiri S, et al (2011) B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. PLOS Genetics 7:e1002203. https://doi.org/10.1371/journal.pgen.1002203

Yoshizawa M, Gorički Š, Soares D, Jeffery WR (2010) Evolution of a Behavioral Shift Mediated by Superficial Neuromasts Helps Cavefish Find Food in Darkness. Current Biology 20:1631–1636. https://doi.org/10.1016/j.cub.2010.07.017

Youngblut ND, Reischer GH, Walters W, et al (2019) Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. Nature Communications 10:2200. https://doi.org/10.1038/s41467-019-10191-3

Zaar A, Fuchs G, Golecki JR, Overmann J (2003) A new purple sulfur bacterium isolated from a littoral microbial mat, *Thiorhodococcus drewsii* sp. nov. Archives of Microbiology 179:174–183. https://doi.org/10.1007/s00203-002-0514-3

Zambonino-Infante JL, Mazurais D, Dubuc A, et al (2017) An early life hypoxia event has a long-term impact on protein digestion and growth in juvenile European sea bass. Journal of Experimental Biology 220:1846–1851. https://doi.org/10.1242/jeb.154922

Zarrinpar A, Chaix A, Xu ZZ, et al (2018) Antibiotic-induced microbiome depletion alters metabolic homeostasis by affecting gut signaling and colonic metabolism. Nature Communications 9:2872. https://doi.org/10.1038/s41467-018-05336-9

Zbinden M, Marqué L, Gaudron SM, et al (2015) Epsilonproteobacteria as gill epibionts of the hydrothermal vent gastropod Cyathermia naticoides (North East-Pacific Rise). Marine Biology 162:435–448. https://doi.org/10.1007/s00227-014-2591-7

Zeng A, Tan K, Gong P, et al (2020) Correlation of microbiota in the gut of fish species and water. 3 Biotech 10:472. https://doi.org/10.1007/s13205-020-02461-5

Zhao F, Yang L, Zhang T, et al (2023) Gut microbiome signatures of extreme environment adaption in Tibetan pig. npj Biofilms Microbiomes 9:1–10. https://doi.org/10.1038/s41522-023-00395-3

Zhou C, Yang S, Gao P, Long R (2022) Association of Gut Microbiota With Metabolism in Rainbow Trout Under Acute Heat Stress. Front Microbiol 13:. https://doi.org/10.3389/fmicb.2022.846336

Zhou L-Y, Yu Z-L, Xu W, et al (2019) *Maribellus luteus* gen. nov., sp. nov., a marine bacterium in the family Prolixibacteraceae isolated from coastal seawater. International Journal of Systematic and Evolutionary Microbiology 69:2388–2394. https://doi.org/10.1099/ijsem.0.003495

Appendix A - Gene expression signatures of salinity transitions in *Limia perugiae* (Poeciliidae), with comparisons to other teleosts

Appendix A Tables

Appendix Table A-1. Mapping statistics for all four population pairs included in the analysis. Reads for *Limia perugiae* were mapped against the *Poecilia mexicana* genome, for *Leuciscus waleckii* against the *Cyprinus carpio* genome, for *Odontesthes argentinensis* and *Odontesthes bonariensis* against the *Menidia menidia* transcriptome, and for *Gasterosteus aculeatus* against the *Gasterosteus aculeatus* genome. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix Table A-2. Results of weighted gene co-expression network analysis for all 18,659 genes identified in our RNA-seq analysis of *Limia perugiae*. Each gene's corresponding module is included in the moduleColor column. The correlation coefficient between each gene's expression and salinity is also included (GS.enviro), as well as the *P*-value for each correlation coefficient (p.GS.enviro). The remaining columns contain the correlation coefficients between each gene and each module, as well as the P-values for each of those correlations. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix Table A-3. GO terms for the turquoise module from the WGCNA results. The GO processes highlighted in yellow are unique to the WGCNA results and were not associated with up-regulated genes in the hypersaline ecotype. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix Table A-4. GO terms for the royalblue module from the WGCNA results. The GO processes highlighted in yellow are unique to the WGCNA results and were not associated with down-regulated genes in the hypersaline ecotype. For comparison with other WGCNA module results, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix Table A-5. GO terms for the blue module from the WGCNA results. The GO processes highlighted in yellow are unique to the WGCNA results and were not associated with down-regulated genes in the hypersaline ecotype. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix B - Natural history and trophic ecology of three populations of the Mexican cavefish, *Astyanax mexicanus*

Appendix B Tables

Appendix Table B-1. Gut contents per fish at sampling times between February 2001 and March 2002 for the Pachón, Sabinos, and Tinaja caves. Absence of diet item categories are indicated by 0's, and presence of diet item categories are indicated by 1's. For the "Full" category, 1=full and 0=empty. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix C - Host-microbiome associations in livebearing fishes adapted to toxic environments rich in hydrogen sulfide

Appendix C Tables

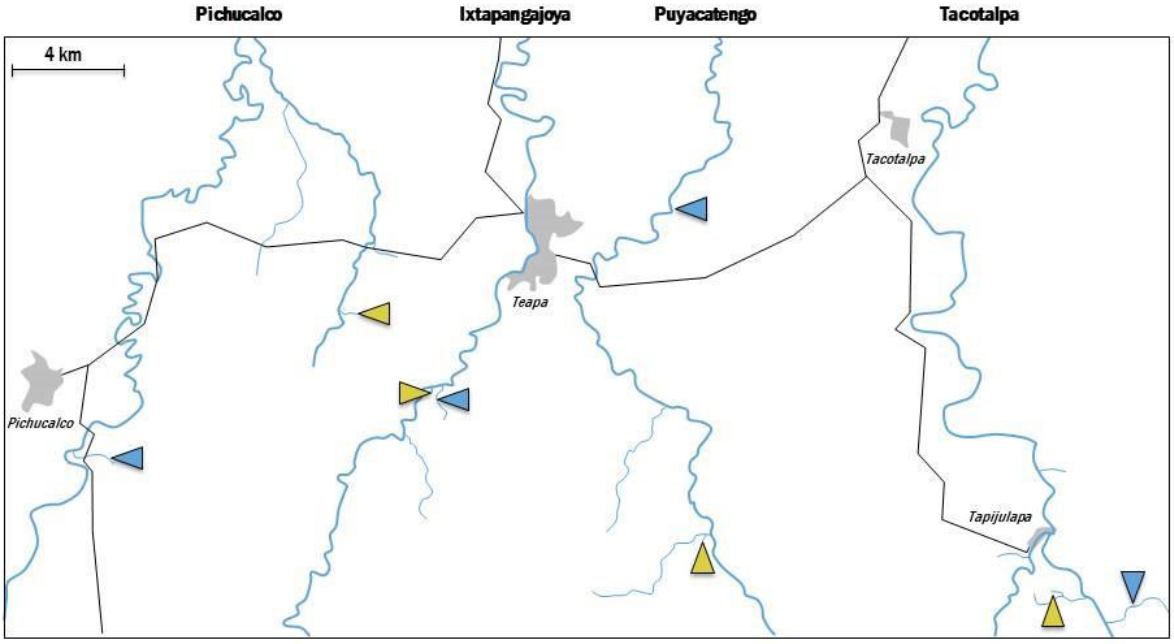
Appendix Table C-1. Merging and filtering statistics for all samples. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

Appendix Table C-2. Read counts after trimming and merging for all of the field and laboratory samples.

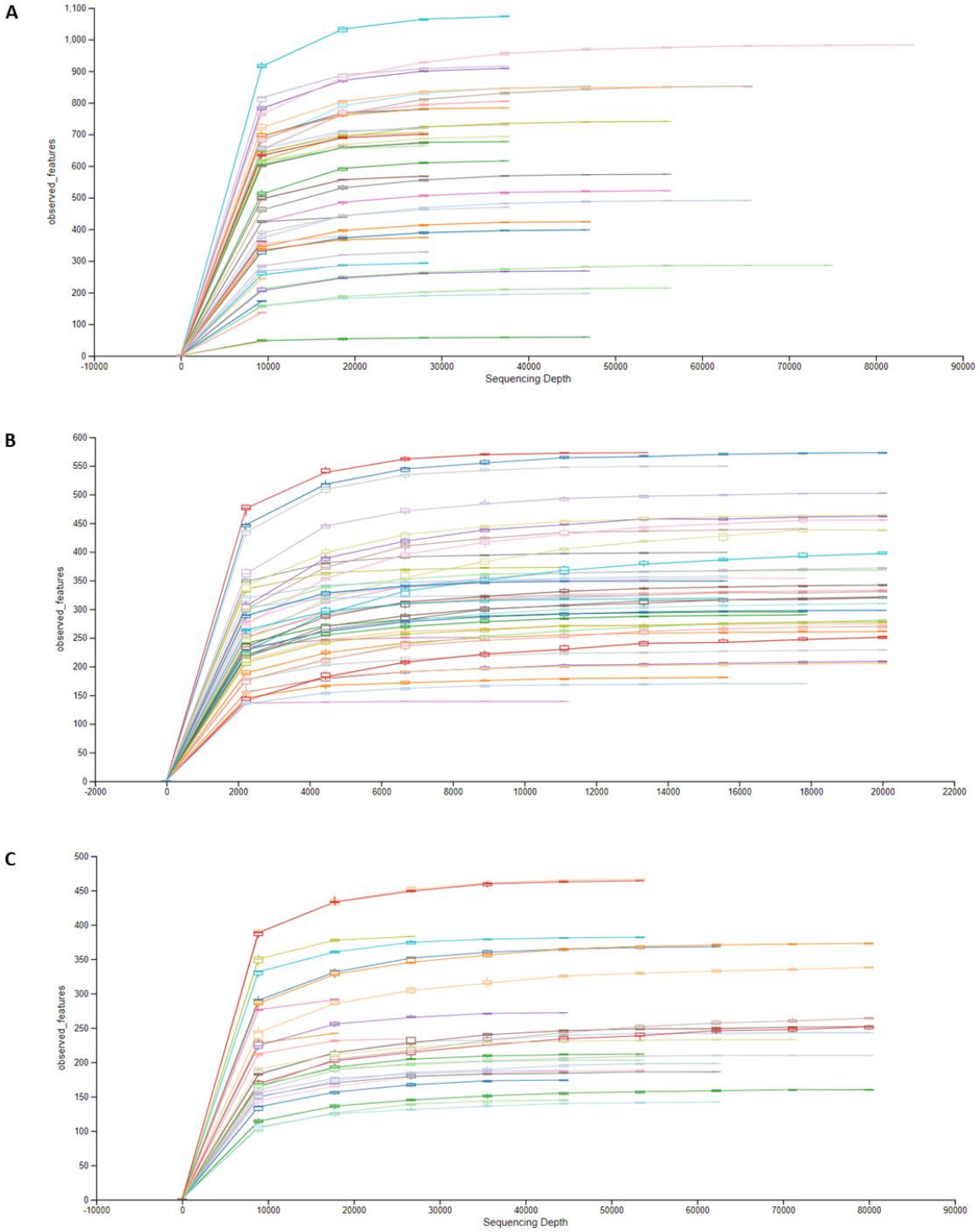
Sample type	Number of samples	Total reads	Minimum reads per sample	Maximum reads per sample
Intestine	48	1,798,794	4,110	83,911
Environment	42	1,086,399	9,400	53,553
Laboratory	10	714,394	52,127	119,815

Appendix Table C-3. Core microbes shared among all four drainages before and after environmental filtering. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

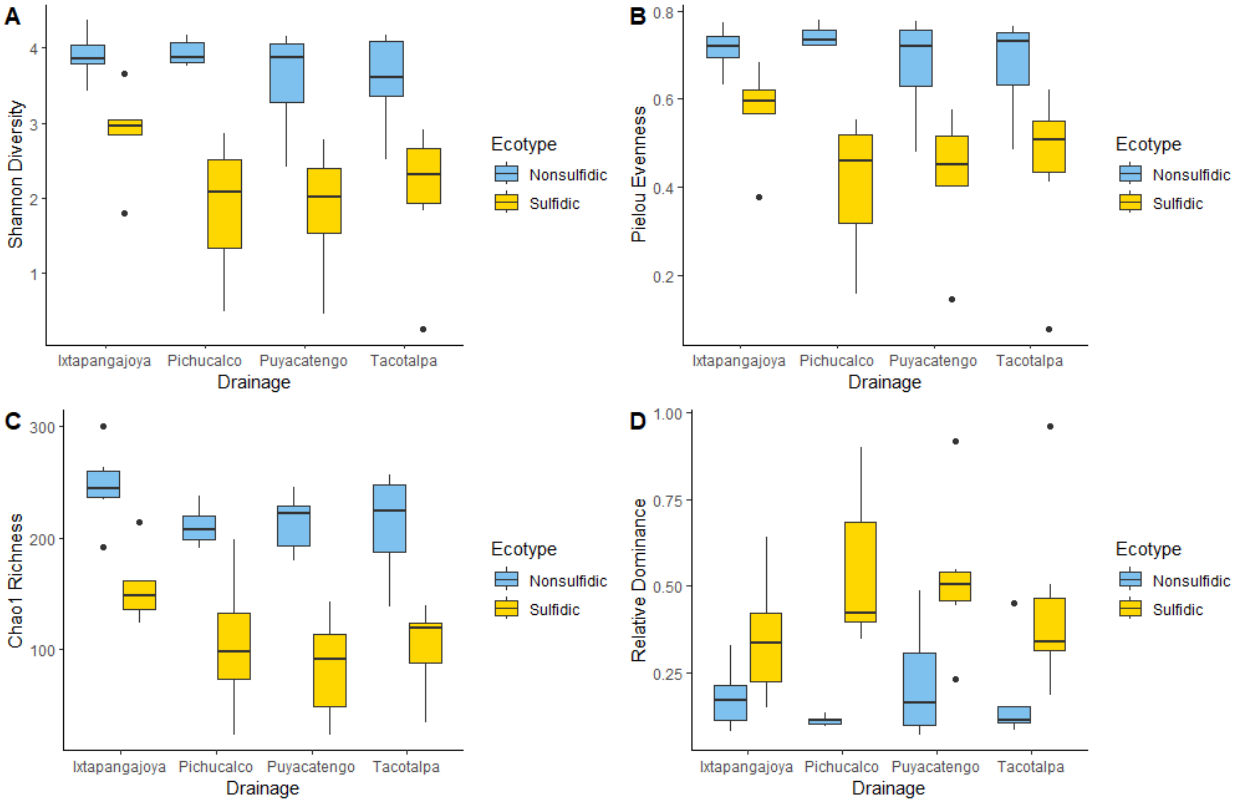
Appendix C Figures



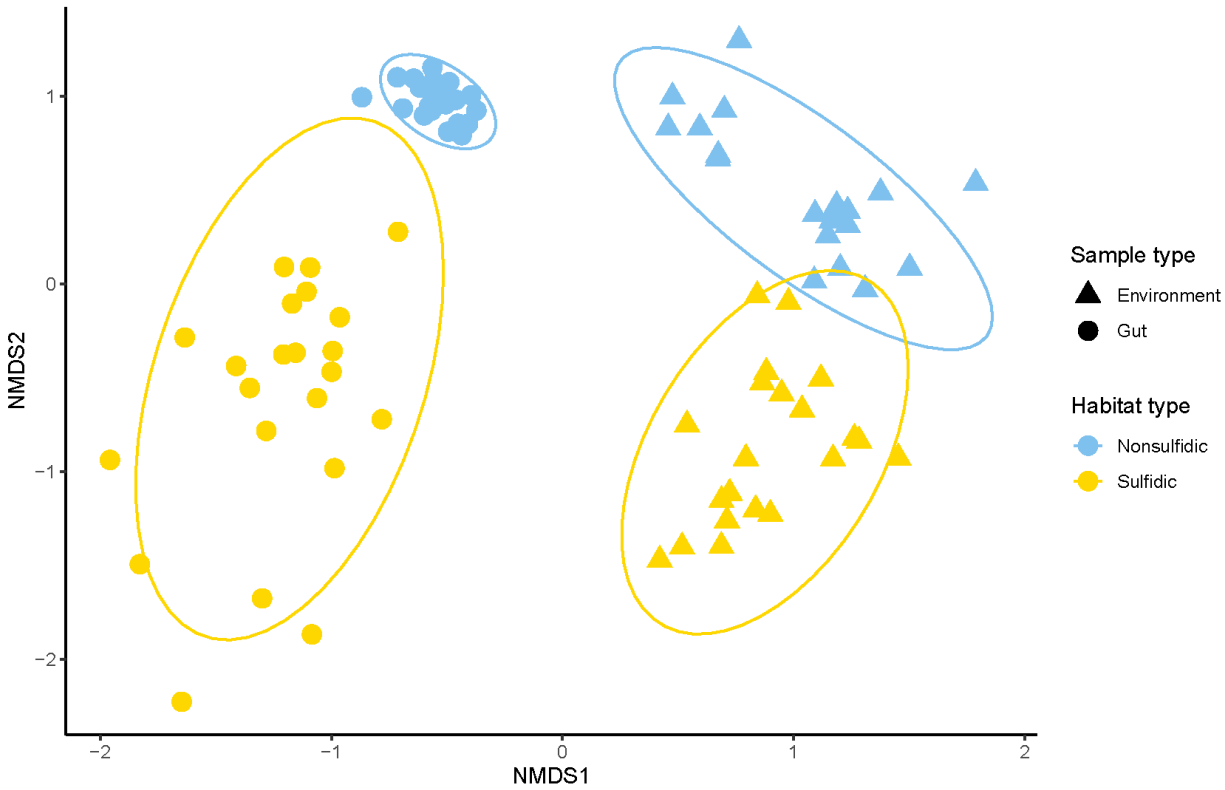
Appendix Figure C-1. Map of the sampling sites. Sampling occurred along the Pichucalco, Ixtapangajoya, Puyacatengo, and Tacotalpa river drainages in the Río Grijalva basin. Water, sediment, and fish gut microbiome samples were taken from a pair of sulfidic and nonsulfidic streams in each river drainage. Sulfidic sampling sites are indicated by yellow triangles, and nonsulfidic sampling sites are indicated by blue triangles.



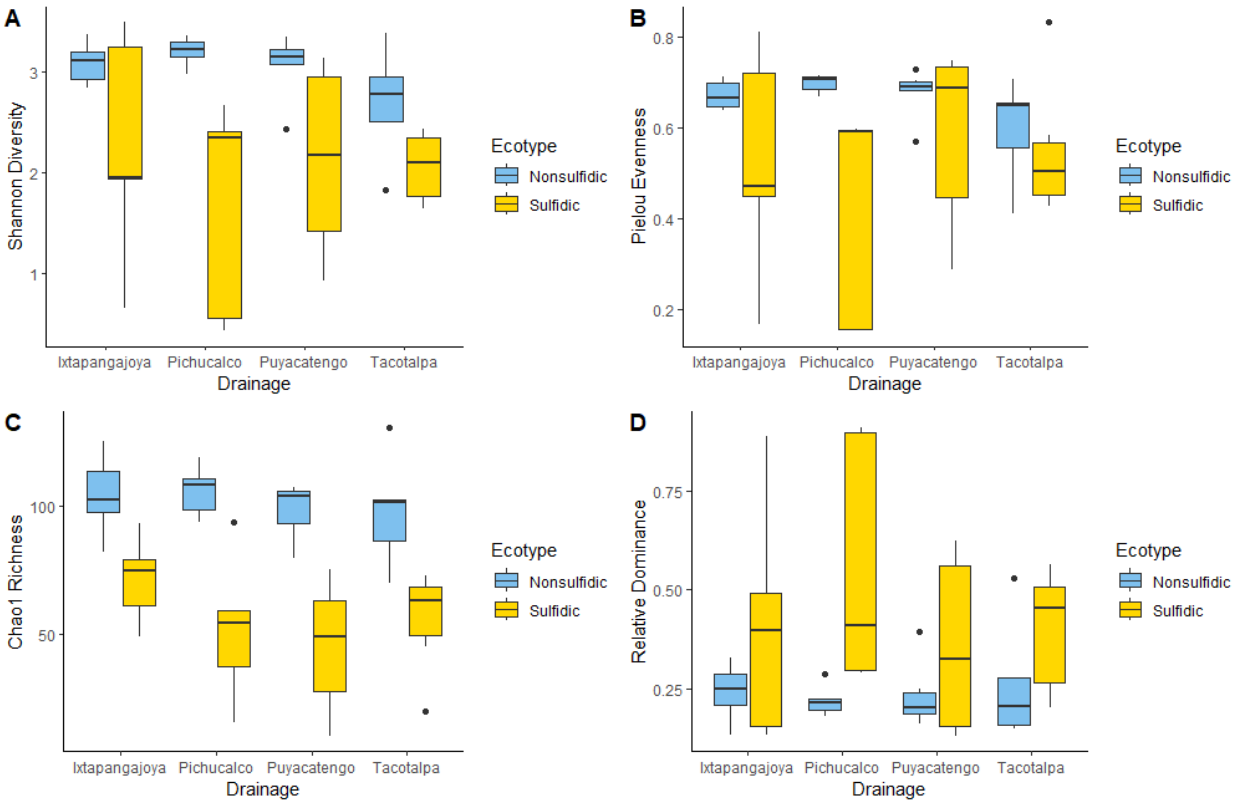
Appendix Figure C-2. Rarefaction curves generated in QIIME2 of the (A) gut microbiome samples from the field, (B) environmental samples from the field, and (C) gut samples from the laboratory showing the number of observed features (sequence variants) relative to sequencing depth.



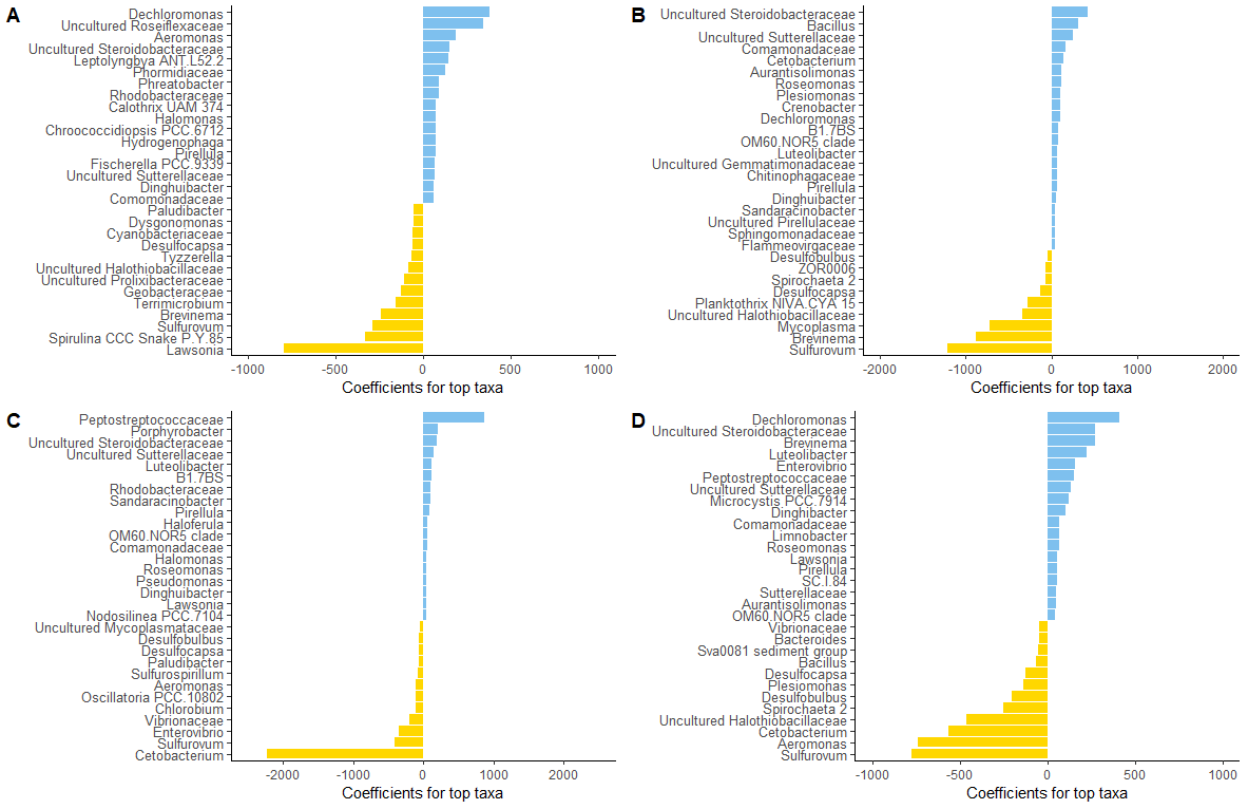
Appendix Figure C-3. Alpha diversity metrics of the gut microbiomes of all *P. mexicana* population pairs, as measured by: A) Shannon diversity index, B) Pielou evenness, C) Chao1 richness, and D) relative dominance.



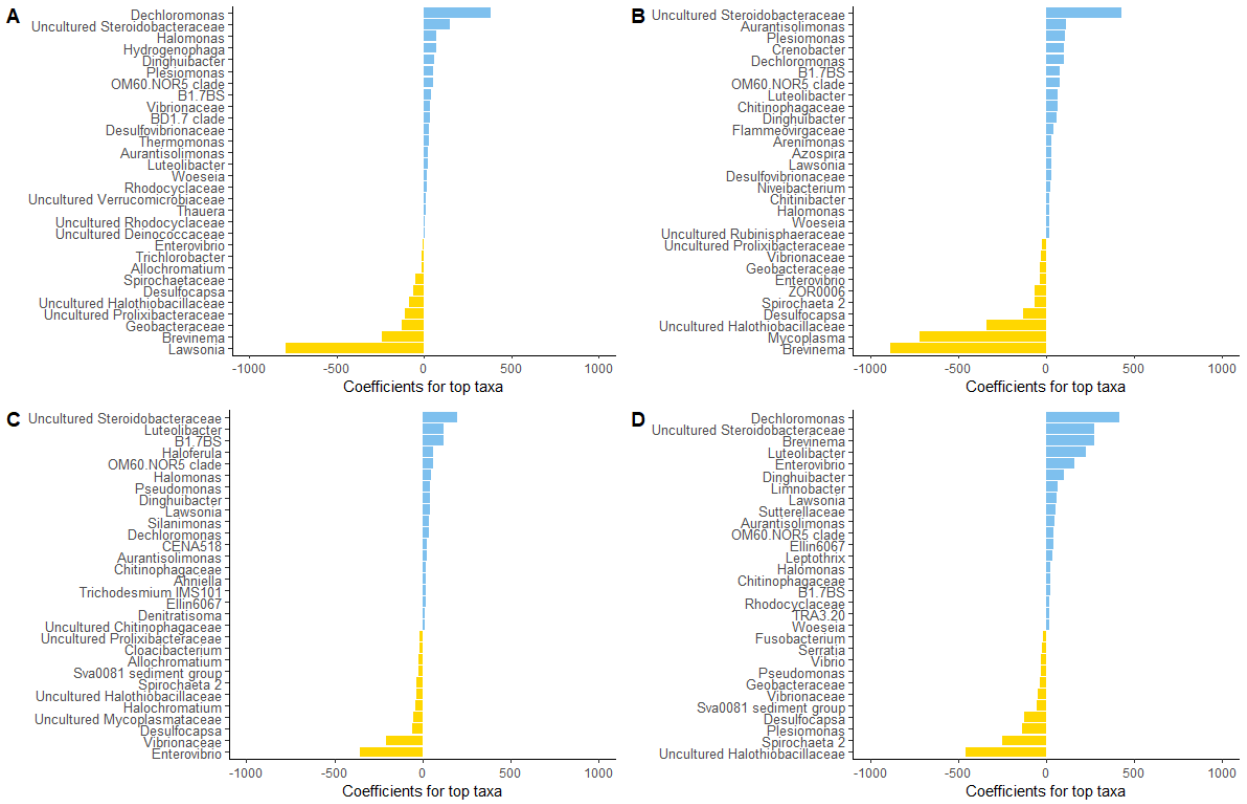
Appendix Figure C-4. NMDS plot based on Bray-Curtis dissimilarity of all gut and environmental samples



Appendix Figure C-5. Alpha diversity metrics of the gut microbiomes of all *P. mexicana* population pairs after environmental correction, as measured by: A) Shannon diversity index, B) Pielou evenness, C) Chao1 richness, and D) relative dominance.



Appendix Figure C-6. The top microbial taxa driving differences between sulfidic and nonsulfidic *P. mexicana* gut microbial community composition in A) the Ixtapangajoya drainage, B) the Pichucalco drainage, C) the Puyacatengo drainage, and D) the Tacotalpa drainage. Taxa in nonsulfidic lineages are indicated by blue bars, and taxa in sulfidic lineages are indicated by yellow bars.



Appendix Figure C-7. The top microbial taxa driving differences between sulfidic and nonsulfidic *P. mexicana* gut microbial community composition after environmental correction in A) the Ixtapangajoya drainage, B) the Pichucalco drainage, C) the Puyacatengo drainage, and D) the Tacotalpa drainage. Taxa in nonsulfidic lineages are indicated by blue bars, and taxa in sulfidic lineages are indicated by yellow bars.

Appendix D - Comparative phylogenetic analysis of Poeciliid microbiomes: effects of evolutionary history and the environment

Appendix D Tables

Appendix Table D-1. Merging and filtering statistics for all samples. Due to large size, this table is included in a Microsoft Excel file called "ElizabethWilson2024_Appendix.xlsx."

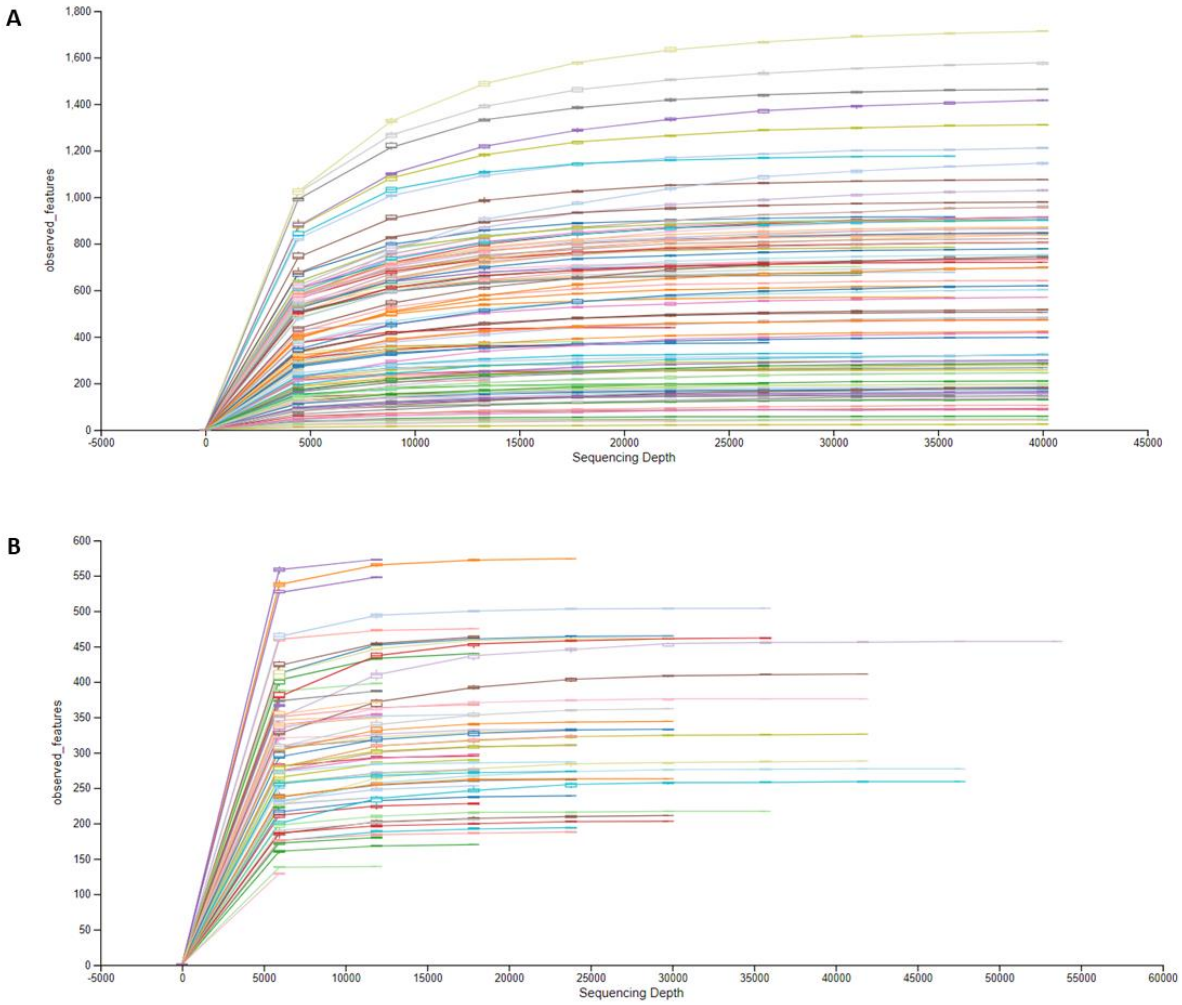
Appendix Table D-2. Read counts after trimming and merging for all of the intestine and environmental samples.

Sample type	Number of samples	Total reads	Minimum reads per sample	Maximum reads per sample
Intestine	132	6,157,994	2,817	122,251
Environment	67	1,601,043	133	53,553

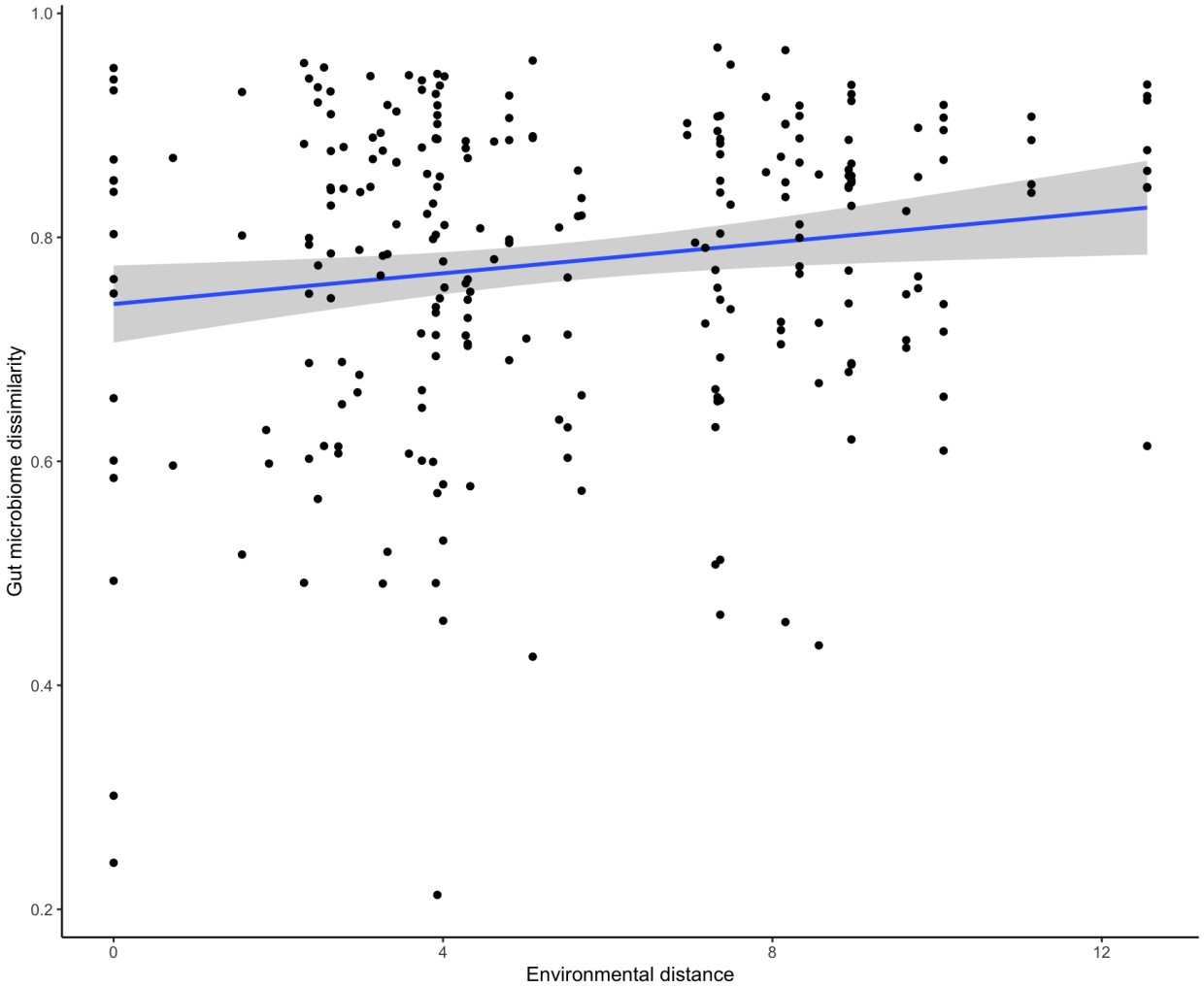
Appendix Table D-3. Convergence metrics of the primarily host-associated microbiomes, indicating the degree of convergence in sulfidic fish gut microbiomes relative to nonsulfidic lineages.

	***C* score**	***P*-value**
C1	0.459689	0.0001
C2	0.444779	0.0000
C3	0.188608	0.0020
C4	0.002732	0.0307

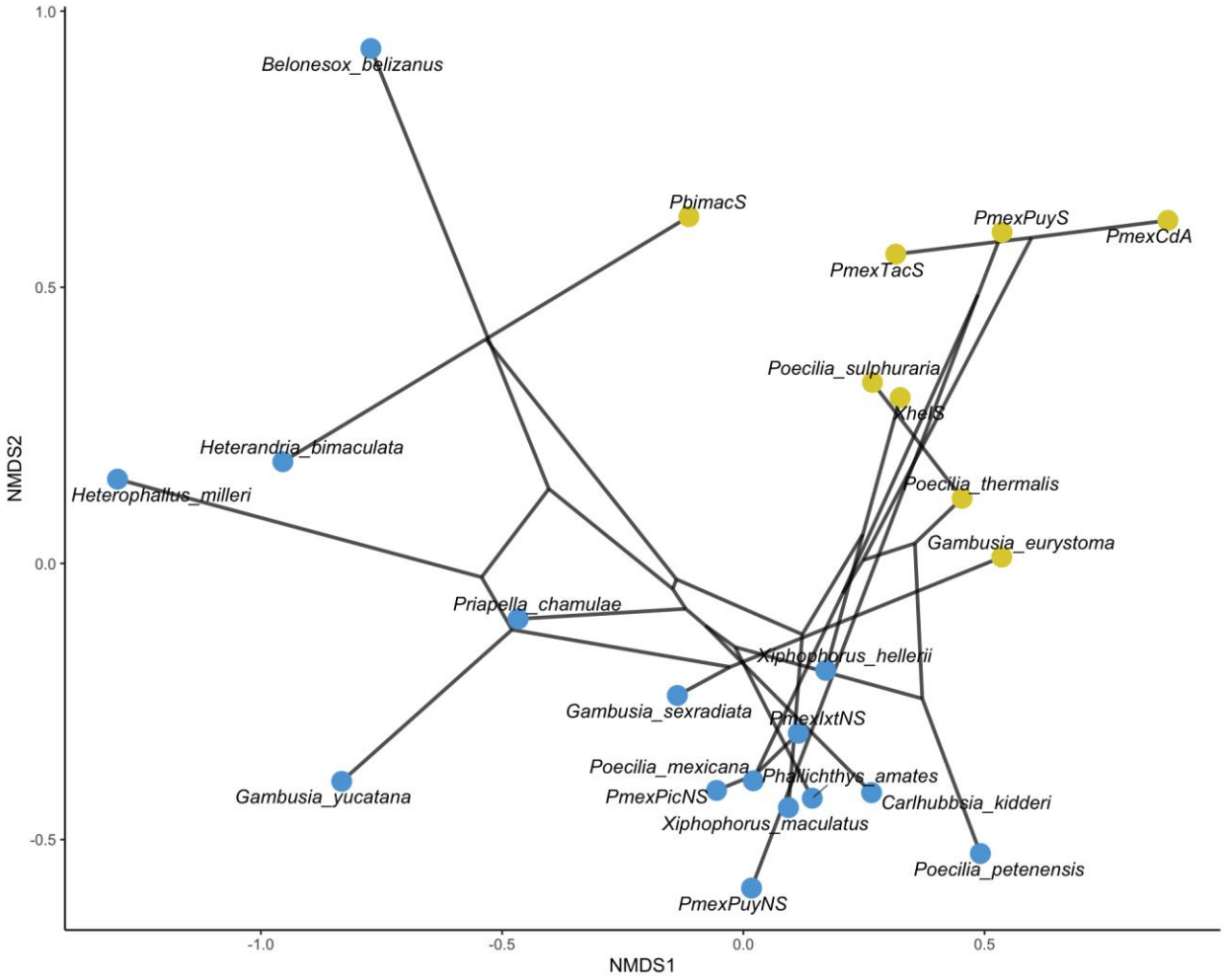
Appendix D Figures



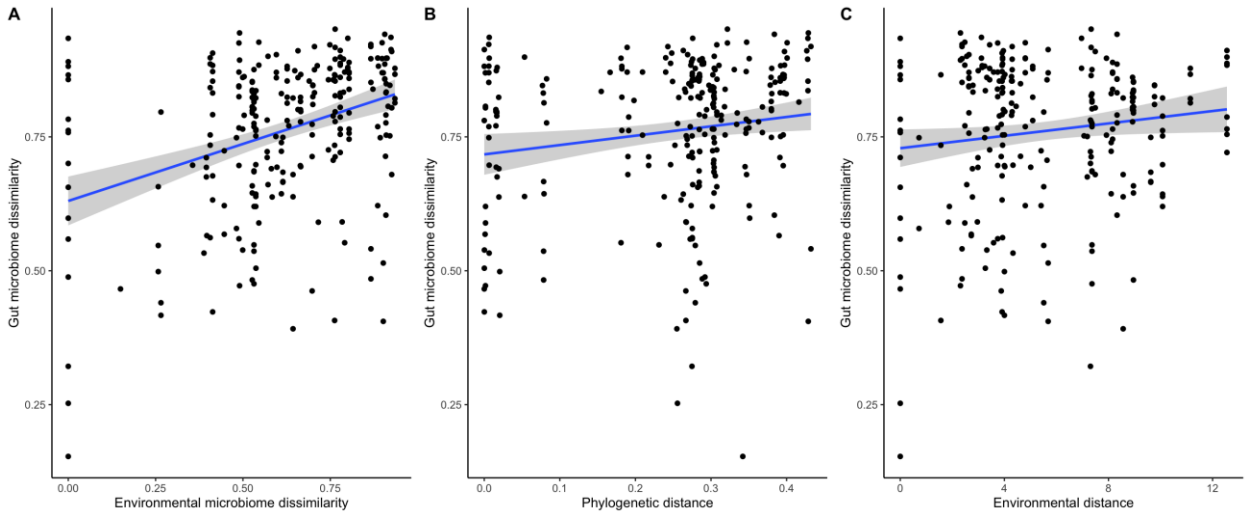
Appendix Figure D-1. QIIME2 rarefaction curves for A) the gut microbiome data and B) the environmental microbiome data showing the number of observed features (sequence variants) relative to sequencing depth.



Appendix Figure D-2. Scatterplot showing gut microbiome dissimilarity vs. environmental distance (based on bioclimatic variables), for which there was no significant correlation (Mantel correlation: 0.033; $P = 0.5959$).



Appendix Figure D-3. NMDS based on Bray-Curtis dissimilarity showing the primarily host-associated gut microbiome variation of all the lineages (after filtering out environmental microbes). The points are NMDS coordinate averages of each lineage's microbiome composition, and the lines connecting the points reflect the host phylogenetic relationships.



Appendix Figure D-4. Correlations between gut microbiome dissimilarity of the primarily host-associated microbiomes and A) environmental microbiome dissimilarity (Mantel correlation: 0.358; $P < 0.0001$) and B) host phylogenetic distance (Mantel correlation: 0.204; $P = 0.0065$). C) There was no significant correlation between gut microbiome dissimilarity and environmental distance (Mantel correlation: –0.028; $P = 0.7201$).