Causes and consequences of adaptation to extreme environments

by

John Lodge Coffin

B.S., University of Georgia, 2017

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

2022

Abstract

Adaptation is ubiquitous in nature, yet our mechanistic understanding of how adaptation occurs and its consequences for populations remains lacking in many natural systems. Extreme environments, characterized by strong physiochemical gradients lethal to most organisms, are ideal systems to investigate the causes of adaptation because selective regimes are well-defined, enabling powerful tests of *a priori* predictions about how resident populations cope with the harsh conditions. Comparisons of extremophile populations with closely related populations in benign habitats facilitate studies of adaptive divergence between ecotypes, but how adaptive divergence coincides with the evolution of reproductive isolation and a reduction in gene flow (i.e., ecological speciation occurs) remains poorly studied. I begin by reviewing these concepts and synthesizing the overall findings of my dissertation in Chapter 1. To fill knowledge gaps related to the causes and consequences of adaptation, I studied extremophile populations of two species of livebearing fishes, Western mosquitofish (*Gambusia affinis*) living in a heavy metal polluted stream in Oklahoma, USA and Atlantic mollies (*Poecilia mexicana*) inhabiting freshwater springs rich in toxic hydrogen sulfide (H_2S) in Tabasco and Chiapas, Mexico. Little is known about how *G. affinis* survive in heavy metal contaminated environments, so in Chapter 2, I quantified metal accumulation and gene expression differences between populations of *G. affinis* inhabiting polluted and unpolluted habitats to identify potential mechanisms by which extremophile populations may be able to tolerate and adapt to the toxic environment. I found that *G. affinis* accumulated heavy metals but may be able to achieve tolerance by detoxifying heavy metals and their toxic intermediates by increasing gene expression of antioxidant genes. However, there was little evidence for adaptive divergence between populations of *G. affinis* living in contrasting habitats, indicated by a general lack of population differentiation and local

adaptation. In contrast, the evidence for adaptation of *P. mexicana* to H_2S is much clearer, but how divergence between populations interacts with phenotypic plasticity (e.g., through maternal effects) to shape trait variation remains unclear. Therefore, in Chapter 3, I measured phenotypic divergence between H_2S-adapted and non-adapted populations of *P. mexicana* and determined the relative impacts of genetics and maternal effects on divergence to understand the origins of adaptive trait variation. I found significant functional trait divergence between ecotypes of *P. mexicana* that was primarily caused by population differences, while plasticity from maternal effects played a relatively weaker role, suggesting divergence is primarily a consequence of local adaptation. Finally, there is also evidence that adaptation to H_2S has coincided with strong—though incomplete—reproductive isolation between sulfidic and nonsulfidic populations of *P. mexicana* occurring before copulation, but it remains to be tested whether there are any reproductive barriers acting after mating. To begin to understand the nature of reproductive barriers that may arise after copulation, in Chapter 4, I analyzed the strength of isolation occurring between copulation and fertilization (i.e., postcopulatory prezygotic isolation) by characterizing population differences in sperm competitive traits along a gradient of H_2S and in homotypic and heterotypic ovarian fluid. This allowed me to determine whether sperm competition and/or cryptic female choice may contribute to reproductive isolation and speciation. There was little evidence for reproductive isolation occurring due to sperm competition or cryptic female choice, which suggests that processes after fertilization are likely major barriers to gene flow that contribute to speciation. My dissertation provides empirical data to connect the causes (adaptation to habitats with divergent ecological selective regimes) and consequences (the accumulation of differences that leads to reproductive isolation and reduced gene flow) of

adaptation in extreme environments, which has important implications for understanding the origins of biodiversity through ecological speciation.

Causes and consequences of adaptation to extreme environments

by

John Lodge Coffin

B.S., University of Georgia, 2017

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

2022

Approved by:

Major Professor
Michael Tobler

Copyright

Abstract

Adaptation is ubiquitous in nature, yet our mechanistic understanding of how adaptation occurs and its consequences for populations remains lacking in many natural systems. Extreme environments, characterized by strong physiochemical gradients lethal to most organisms, are ideal systems to investigate the causes of adaptation because selective regimes are well-defined, enabling powerful tests of *a priori* predictions about how resident populations cope with the harsh conditions. Comparisons of extremophile populations with closely related populations in benign habitats facilitate studies of adaptive divergence between ecotypes, but how adaptive divergence coincides with the evolution of reproductive isolation and a reduction in gene flow (i.e., ecological speciation occurs) remains poorly studied. I begin by reviewing these concepts and synthesizing the overall findings of my dissertation in Chapter 1. To fill knowledge gaps related to the causes and consequences of adaptation, I studied extremophile populations of two species of livebearing fishes, Western mosquitofish (*Gambusia affinis*) living in a heavy metal polluted stream in Oklahoma, USA and Atlantic mollies (*Poecilia mexicana*) inhabiting freshwater springs rich in toxic hydrogen sulfide (H_2S) in Tabasco and Chiapas, Mexico. Little is known about how *G. affinis* survive in heavy metal contaminated environments, so in Chapter 2, I quantified metal accumulation and gene expression differences between populations of *G. affinis* inhabiting polluted and unpolluted habitats to identify potential mechanisms by which extremophile populations may be able to tolerate and adapt to the toxic environment. I found that *G. affinis* accumulated heavy metals but may be able to achieve tolerance by detoxifying heavy metals and their toxic intermediates by increasing gene expression of antioxidant genes. However, there was little evidence for adaptive divergence between populations of *G. affinis* living in contrasting habitats, indicated by a general lack of population differentiation and local

adaptation. In contrast, the evidence for adaptation of *P. mexicana* to H_2S is much clearer, but how divergence between populations interacts with phenotypic plasticity (e.g., through maternal effects) to shape trait variation remains unclear. Therefore, in Chapter 3, I measured phenotypic divergence between H_2S-adapted and non-adapted populations of *P. mexicana* and determined the relative impacts of genetics and maternal effects on divergence to understand the origins of adaptive trait variation. I found significant functional trait divergence between ecotypes of *P. mexicana* that was primarily caused by population differences, while plasticity from maternal effects played a relatively weaker role, suggesting divergence is primarily a consequence of local adaptation. Finally, there is also evidence that adaptation to H_2S has coincided with strong—though incomplete—reproductive isolation between sulfidic and nonsulfidic populations of *P. mexicana* occurring before copulation, but it remains to be tested whether there are any reproductive barriers acting after mating. To begin to understand the nature of reproductive barriers that may arise after copulation, in Chapter 4, I analyzed the strength of isolation occurring between copulation and fertilization (i.e., postcopulatory prezygotic isolation) by characterizing population differences in sperm competitive traits along a gradient of H_2S and in homotypic and heterotypic ovarian fluid. This allowed me to determine whether sperm competition and/or cryptic female choice may contribute to reproductive isolation and speciation. There was little evidence for reproductive isolation occurring due to sperm competition or cryptic female choice, which suggests that processes after fertilization are likely major barriers to gene flow that contribute to speciation. My dissertation provides empirical data to connect the causes (adaptation to habitats with divergent ecological selective regimes) and consequences (the accumulation of differences that leads to reproductive isolation and reduced gene flow) of

adaptation in extreme environments, which has important implications for understanding the origins of biodiversity through ecological speciation.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to begin by thanking my broad support system, without which I would not have been able to accomplish what I have in life and in my doctoral studies. This support system begins and ends with my wonderful fiancé Bliss Betzen, who has been with me through thick and thin and always knows what to do and say to inspire me to be my best in all aspects of life. She is truly a breath of fresh air and the spark that helps me find and pursue my passions. I would also like to thank my family for always listening to me excitedly talk about science and encouraging me to keep going. I would specifically like to thank my parents, Tom and Julie Coffin, for their support of education and instilling in me a strong work ethic and thirst for knowledge and inquiry. Graduate school has been one of the most difficult, yet rewarding parts of my life, and while my family is over 1,000 miles away in Atlanta, Georgia, I've always felt welcomed, included, and supported by my soon-to-be in-laws in the Midwest, whom I would also like to thank for being like a family away from home to me. I will also always treasure the friendships I have formed during my time in graduate school.

The Tobler lab at KSU was an awesome group to work with over the last five years. I'd first like to thank my advisor, Michi Tobler, for his patience, his commitment to teaching/communicating, his help and feedback, and for helping me combat my imposter syndrome. The confidence I've gained from working, teaching, and presenting with him has helped me tremendously, and is something I will continue benefitting from as I continue my career. Michi also had a knack for helping me better manage my time to make and keep realistic deadlines. There is also a swarm of Tobler lab team members, including graduate students, awesome undergrads, and post-bac students who just wanted to help out. Without this team to collaborate and commiserate with, there is no way I would have been able to get through the last

five years of graduate school. I'd therefore like to resoundingly thank Bryan Frenette, Ryan Greenway, Henry Camarillo, Nick Barts, Libby Wilson, Cassie Delich, Madison Nobrega, Hannah Hoffman-Colburn, Justine Onnen, Quinlyn La Fon, Michael Laughlin, Amber Arias, Dylan Stockman, Jennie Grill, Katie Rodriguez, and Nathan Herrman.

I would also like to thank my committee, including Mark Ungerer, Keith Gido, and Joanna Kelley. In addition to their support and guidance, I have had stimulating conversations with all of them that have helped me grow as a scientist. I'd specifically like to acknowledge Mark Ungerer, with whom I rotated during the first semester of my Ph. D., and Joanna Kelley, who was instrumental in training me to conduct lab work and bioinformatics.

One of the most rewarding experiences of my time in graduate school has been teaching in the organismic biology, evolution, and ichthyology courses. I'd like to thank the head professors of each of these classes, who helped me learn the basics of pedagogy and spent the time to teach me content with which I was unfamiliar. I'd also like to thank the graduate students I taught with, and especially Sophie Higgs, who taught with me during my first semester of graduate school. I was quite frankly terrified to be thrown into a classroom of students with little background knowledge or experience in a classroom, and Sophie held my hand through it all and helped make teaching fun. While I never taught with them personally, Nick Barts also had a large impact on my teaching, as they taught me effective lesson planning and activity design. In addition to the instructional staff I interacted with, I'd also like to thank the students I've taught and learned from over the last five years. They are what makes teaching fun, exciting, and rewarding.

The research I conducted for this dissertation would not have been possible without the collaborations I formed with groups outside of KSU. I'd first like to thank the Grand River Dam

Authority for their help securing lodging and research space while conducting research for the first chapter of my dissertation. Furthermore, while I was not able to conduct field work in Mexico in person during my Ph. D., collaborations between my lab group and the Centro de Investigación e Innovación para la Enseñanza y Aprendizaje in Teapa, Mexico enabled the research for my second and third chapter to be accomplished. I would also like to acknowledge the organizations that funded my research, including the National Science Foundation, the KSU Biology Graduate Student Association, the KSU Division of Biology, the KSU Graduate School, the KSU College of Arts and Sciences, the US Army Research Office, and the US Department of Education.

Last, but certainly not least, I would like to acknowledge that all of my research was conducted on lands that historically belonged to a variety of Native American groups. KSU sits on the traditional lands of the Kaw people, who were forced onto reservations in Oklahoma by white American settlers. My research in Northeast Oklahoma was also conducted on lands where numerous Native American tribes were forcibly displaced and relocated by white American settlers, including bands from the Osage, Iroquois, Shawnee, Quapaw, Peoria, Miami, Modoc, and Wyandotte tribes, among others. Likewise, the sulfide springs in Southern Mexico are on lands that were ancestrally held by various Olmec and Mayan civilizations until conquest by the Spanish. As the world has come to know in the wake of the Russian invasion of Ukraine, the Israeli occupation of Palestine, and the current colonial holdings of the United States and nations worldwide, colonialism is alive and well even today. It is my hope that by acknowledging that my work was conducted on land that was stolen by colonizers, I can take note of the privileges I have been afforded at the expense of others, work towards supporting and celebrating Native

Americans and other marginalized communities in the United States and abroad, and hopefully inspire others to follow suit.

Dedication

This work is dedicated to my grandma, Janet Coffin, and my Papa, Charlie Smith, both of whom passed away before I completed my doctoral studies but lifted me up to new heights. They were role models in life and love whom I wish to emulate.

Chapter 1 - Introduction and synthesis

Introduction

Understanding how a single, panmictic population of individuals that closely resemble each other can evolve into two or more phenotypically distinct, reproductively isolated populations (i.e., new species) has long baffled biologists. The complex web of eco-evolutionary processes that may contribute to this process, summed up in Darwin's "tangled bank" metaphor (1859), may superficially resemble chaos. The number of interacting species and environmental conditions can seem dizzyingly complex, but as Darwin suggested, this chaos is subject to the natural laws of evolutionary biology. We now have several rigorous frameworks for understanding the processes that generate biodiversity, one of which has been formalized as the process of ecological speciation, whereby adaptive divergence between populations coincides with the evolution of reproductive isolation and the formation of new species (Nosil 2012). This framework has been applied to study the origins of biodiversity across the tree of life (Schluter 1996; Nosil et al. 2008; Jacquemyn et al. 2018; Tobler et al. 2018) because it connects the causes (i.e., divergent selection) and consequences (i.e., speciation through the evolution of reproductive isolation) of adaptation (Rundle and Nosil 2005).

To effectively study the causes and consequences of adaptation, we must first establish the mechanisms underlying adaptation. Identifying the traits that enable survival, and the genes underlying them, is crucial for understanding why and how populations respond to divergent selection. However, not all patterns of trait divergence are due to adaptation. Trait divergence among populations can also be caused by phenotypic plasticity, which occurs when a single genotype gives rise to multiple phenotypes in different environments (Pigliucci 2001). Therefore, understanding whether population differences are due to adaptation requires laboratory

experiments that explicitly test for trait divergence from genetic differentiation (which is indicative of adaptation) and phenotypic plasticity. Once the proximate causes of adaptation and the phenotypic patterns that result are known, we can then address whether and how adaptation has led to the evolution of reproductive isolation and the formation of new species through ecological speciation.

Extreme environments are ideal models for studying these processes because they contain gradients of physiochemical stressors that are lethal to most organisms (Tobler et al. 2018). However, these harsh environments often host extremophile populations that are able to tolerate the conditions and survive (Greenway et al. 2014). Non-adapted populations from benign habitats often die upon exposure (e.g., Madsen et al. 2015), and comparisons between populations in extreme and benign habitats reveal strong, divergent selective regimes that can and often do lead to local adaptation (Graham et al. 2018). This strong divergent ecological selection allows for powerful tests of *a priori* predictions about how adaptation to such extreme conditions occurs (Tobler et al. 2018). I leveraged two extreme environments, one produced by the natural release of toxic hydrogen sulfide (H_2S), and another caused by pollution of heavy metals from intensive mining, to understand generalities of adaptation to extreme environments, the origins of adaptive trait variation, and the evolution of reproductive isolation.

Sulfide springs in Tabasco and Chiapas, Mexico are extreme environments that occur across separate river drainages, providing a natural laboratory in which to conduct comparative studies on adaptation to extreme environmental conditions. The toxicity of these springs stems from hydrogen sulfide (H_2S), which disrupts the oxidative phosphorylation (OXPHOS) pathway in aerobic organisms (Cooper and Brown 2008). Even at micromolar concentrations, H_2S causes a significant reduction in aerobic ATP production, which rapidly causes death in non-adapted

individuals (Reiffenstein et al. 1992). Across several river drainages of the Grijalva River basin, sulfide springs directly connect to the mainstem of benign, nonsulfidic habitats without any physical barriers separating the habitat types (Plath et al. 2013). Ancestral populations of Atlantic mollies (*Poecilia mexicana*) have repeatedly colonized these sulfidic habitats from adjacent nonsulfidic habitats and adapted to the extreme conditions (Brown et al. 2018). Despite the continued lack of physical barriers, the sulfidic ecotypes of *P. mexicana* have diverged phenotypically and genetically from the nonsulfidic ecotypes, suggesting that ecological speciation is ongoing (Plath et al. 2013). While recent studies on genetic, transcriptomic, and physiological divergence between ecotypes have painted a relatively clear picture of the mechanisms of adaptation to H_2S (Tobler et al. 2008; Kelley et al. 2016; Camarillo et al. 2020; Greenway et al. 2020), it remains unclear how genetic divergence has interacted with phenotypic plasticity to shape the trait divergence observed in nature, and how adaptive divergence leads to reproductive isolation between ecotypes.

Another extreme environment, the Tar Creek Superfund site in Northeastern Oklahoma, USA, is a freshwater tributary of the Neosho River that has been polluted with heavy metals—particularly lead and zinc—from intensive mining and improper waste management practices in the Tri-State Mining District in the 20^{th} century (OWRB 1983). These heavy metals are toxic to most animals because they easily enter the blood stream through the respiratory and gastrointestinal systems and mimic calcium ions to enter cells (Wood et al. 2011a,b). Once inside cells, heavy metal ions can generate free radicals that lead to oxidative damage, which ultimately disrupt homeostasis and lead to disease and death (Hogstrand 2011; Mager 2011). Several fish species remain in Tar Creek, though at very small population sizes (Franssen et al. 2006); the majority of animal life became locally extirpated or migrated from Tar Creek following exposure

to heavy metal waste (Thomas 1984). However, one livebearing fish species, the Western mosquitofish (*Gambusia affinis*), has become abundant in polluted sections of Tar Creek (Franssen et al. 2006). In contrast to the sulfide springs of Southern Mexico, we know relatively little about the mechanisms potentially underlying adaptation to the conditions in Tar Creek, nor whether populations of *G. affinis* from polluted areas are reproductively isolated from populations in benign areas.

Across two genera of livebearing fishes (*Poecilia* and *Gambusia*) inhabiting natural and man-made extreme environments, I compared individuals from extreme habitats to those living in proximate benign habitats to address three questions related to the causes and consequences of adaptation in natural populations: **1) Are all extreme environments—and the responses of populations to different selective pressures—the same?** We have a thorough understanding of how populations of mollies have colonized sulfide springs and adapted to the toxic conditions, but how do populations of other species, like mosquitofish, survive in extreme environments like Tar Creek? Theory predicts that organisms should employ at least one of four potential mechanisms to maintain homeostasis and support the biological processes necessary for life in the face of extreme environmental conditions. Organisms can either minimize uptake or maximize excretion of harmful compounds to maintain non-toxic endogenous concentrations through exclusion (Goffredi et al. 1997; Grieshaber and Völkel 1998). However, if contaminants are able to enter and remain in the body, then organisms can achieve tolerance by converting toxic compounds into less harmful intermediates through detoxification (Hildebrandt and Grieshaber 2008), making molecular changes to the direct toxicity target through resistance (e.g., Pfenninger et al. 2014), and/or making use of different physiological pathways through mitigation (Hildebrandt and Grieshaber 2008). In the second chapter of this dissertation, I used a

systems-level approach to quantify ionomes—the concentrations of all mineral nutrients in an organism (Lahner et al. 2003)—and genome-wide gene expression profiles (i.e., transcriptomes) in mosquitofish living in metal-contaminated and unpolluted habitats to test whether mosquitofish have adapted to life in Tar Creek by employing exclusion and/or detoxification mechanisms to maintain homeostasis.

2) What mechanisms contribute to trait variation observed in nature? Phenotypes are shaped by genetic effects, environmental effects that influence organismal development (e.g., phenotypic plasticity: Pigliucci 2001), and their interactions (Stearns 1992). One special case of phenotypic plasticity that commonly impacts trait expression in nature is maternal effects, which occur when the environment experienced or created by a mother affects her offspring's phenotypes (Mousseau and Fox 1998). Trait variation along environmental gradients, like those found in extreme environments, is common, but to what degree variation is caused by genetic differences between populations or environmentally mediated through plasticity (from maternal effects or direct environmental exposure in offspring) is generally unknown. However, identifying the origins of trait variation is important because the origins of trait divergence are crucial for making inferences about adaptation in nature: if traits are controlled by different factors (i.e., genetic vs. plastic control), then they may respond to selection in different ways (as in Felmy et al. 2022). Genetically mediated traits (i.e., ones with high heritability) can respond strongly to bouts of selection, potentially enabling rapid adaptation (e.g., in response to climate change: Gienapp et al. 2008). Non-heritable traits, like those impacted strongly by plasticity, may not respond to selection, so even strong selective regimes may not cause evolutionary change (Scheiner 2002). However, phenotypic shifts induced by plasticity can nonetheless be adaptive (Passow et al. 2017) and can actually potentiate adaptive evolutionary change if plastic

phenotypes in a population are further from their local optimum (i.e., increasing the strength of directional selection; see Ghalambor et al. 2015). Colonization of sulfide springs by populations of *P. mexicana* has led to divergence between sulfidic and nonsulfidic ecotypes in several traits (see Tobler et al. 2018 for a review), but it remains unclear to what degree these differences are due to population differentiation and/or plasticity. To understand the origins of trait variation in *P. mexicana*, I induced plastic maternal effects by varying resource availability to pregnant *P. mexicana* from sulfidic and nonsulfidic ecotypes and repeatedly quantified complex behavioral and life history traits throughout ontogeny in offspring. Ecotypic comparisons elucidated patterns of population differentiation, while comparisons between maternal food treatments provided insight into the strength of maternal effects (and thus plasticity) on trait divergence.

3) What causes gene flow to be reduced between sulfidic and nonsulfidic ecotypes of *P. mexicana*? In sexually reproducing species, reproductive isolation can evolve prior to mating (i.e., precopulatory barriers: Nosil 2004), after mating but before fertilization (i.e., postcopulatory prezygotic barriers: Devigili et al. 2018), or after fertilization (i.e., postzygotic barriers: Lopez et al. 2000). All of these barriers may act in unison to reduce gene flow, but the timing of each type of barrier relative to the others has important implications for its realized effect on gene flow. For example, if a precopulatory barrier—such as behavioral isolation caused by divergence in mating behaviors between ecotypes—is so strong that different ecotypes never copulate with each other, then postcopulatory prezygotic and postzygotic isolation are not relevant for reducing gene flow in nature, even though they could theoretically evolve. Therefore, many studies first attempt to identify precopulatory barriers and quantify their strength before moving on to postcopulatory prezygotic and then postzygotic barriers.

There is a significant body of evidence documenting strong precopulatory isolation between sulfidic and nonsulfidic ecotypes of *P. mexicana*, but total precopulatory isolation does not correlate with observed patterns of gene flow, suggesting that other barriers after copulation have likely evolved (Plath et al. 2013). Therefore, the fourth chapter of my dissertation examined the presence and strength of postcopulatory prezygotic isolation. Two phenomena that can lead to strong postcopulatory prezygotic isolation are sperm competition and cryptic female choice (Birkhead and Pizzari 2002). Sperm competition occurs whenever ejaculates from two or more males compete with each other to fertilize a female's ova (Parker 1970) and can be mediated by several processes. First, males may have evolved different ejaculate sizes or sperm swimming abilities, which would confer different sperm competitive strengths (Boschetto et al. 2011). Second, females from many poeciliid species—including *P. mexicana*—can store sperm for use in fertilization after copulation (Greven 2011), so males may also have evolved sperm traits that enable sperm survival after ejaculation for a prolonged time (García-González and Simmons 2005; Gasparini and Evans 2013), which is also important for mediating sperm competition success. Cryptic female choice, on the other hand, is a process by which females can preferentially favor the sperm of certain males without actively choosing them (Eberhard 1996). Thus, females can exert control over male-mediated processes like sperm competitive ability to suit their own reproductive needs. If sperm competitive ability is a locally adapted trait, and if cryptic female choice favors sperm from males of the same population over those from males of other populations, then these processes should lead to reproductive isolation. In Chapter 4, I measured several sperm competitive traits (ejaculate size, sperm swimming performance, and sperm longevity) and quantified the effect of cryptic female choice on sperm competition (by comparing sperm competitive ability in related and unrelated ovarian fluid) to evaluate the

presence and strength of postcopulatory prezygotic isolation between sulfidic and nonsulfidic ecotypes of *P. mexicana*.

Synthesizing the causes and consequences of adaptation

Causes—Predictability of adaptive evolution in response to strong divergent selection

Extreme environments are useful in the study of adaptation primarily because of their strong, well-defined selective regimes, which provide insight into the mechanisms that may be required to adapt to and overcome such strong selection. In sulfide springs, the toxicity target of H_2S is well documented; sulfide reversibly binds to complex IV of OXPHOS, which reduces he capacity for aerobic ATP production (Cooper and Brown 2008). H_2S also readily reacts with oxygen, leading to hypoxic conditions in sulfide-rich habitats (Tobler et al. 2011a). This suggests that adaptation should occur through mechanisms that reduce/counteract the toxicity of H_2S and enhance oxygen acquisition ability. Indeed, colonization of sulfide-rich habitats has coincided with molecular modifications in the mitochondria that reduce the binding of H_2S to complex IV of OXPHOS and gene expression modifications that increase sulfide detoxification (Pfenninger et al. 2014; Kelley et al. 2016; Greenway et al. 2020), and has also coincided with behavioral and morphological evolution that increases oxygen availability in hypoxic conditions (exploiting oxygen-rich surface waters through aquatic surface respiration: Plath et al. 2007b; evolution of larger heads and gills: Tobler et al. 2008).

But what happens if the selective regimes and toxicity targets are less clear? As in sulfide springs, heavy metal pollution likely exerts strong viability selection on the aquatic inhabitants of Tar Creek, evidenced by the large die-off of the ichthyofauna in the late twentieth century following mine outflows (Thomas 1984) and the reduced abundance and diversity of fishes in

Tar Creek to this day (Franssen et al. 2006). In the second chapter of my dissertation, I set out to understand how Western mosquitofish (*Gambusia affinis*), by far the most abundant fish in Tar Creek, are able to circumvent this strong viability selection and survive. I found evidence for a putatively adaptive increase in transcription of antioxidant genes in the Tar Creek population compared to the unpolluted reference populations, which could enable tolerance of the toxic conditions. However, my analyses also provided evidence for a lack of local adaptation and population genetic differentiation between habitats, suggesting that mosquitofish may not be adapted to life in Tar Creek. Instead, mosquitofish may be poorly suited for life in heavy metal contaminated habitats, and the population we sampled is merely a sink population that is only present due to constant colonization from other, more favorable habitats. Concomitantly, mosquitofish could be able to tolerate the toxic conditions and survive due to strong patterns of phenotypic plasticity. Integrating my results that found a lack of population genetic differentiation and putatively adaptive increases in antioxidant gene expression suggests that it is likely a combination of these two hypotheses that explains the persistence of populations of mosquitofish in Tar Creek.

Ultimately, the first chapter of my dissertation suggests that strong divergent selection may not always lead to predictable evolutionary changes. This may be because the selective regime and toxicity targets in Tar Creek are not as well-defined as they are in sulfide springs. While the toxicity of H_2S has been clearly documented to affect the mitochondria (Reiffenstein et al. 1992; Cooper and Brown 2008; Olson et al. 2012; Wallace and Wang 2015), the toxicity of various metals does not befall a single cellular component. Instead, oxidative damage from heavy metal exposure can detrimentally impact lipids, proteins, and nucleic acids (Cooke et al. 2003; Davies 2012; Regoli and Giuliani 2014) and disrupt ATP production in the mitochondria

(Belyaeva et al. 2012). Since the toxicity targets of heavy metals are quite broad, we might expect to see similarly broad responses to selection that belong to multiple pathways. However, genomic architecture could constrain the possible ways in which these multiple pathways interact, which could in turn constrain the possible adaptive response to heavy metals in mosquitofish.

Causes—Effects of adaptation and its alternatives on trait expression and divergence

Extreme environments—and ecological gradients in general—often exhibit strong patterns of phenotypic variation between divergent habitats. Many studies quantify phenotypic differences between populations as evidence for adaptation (see Gienapp et al. 2008 and references therein), yet there are a number of alternative mechanisms that can produce phenotypic patterns similar to adaptation, including phenotypic plasticity. Atlantic mollies (*Poecilia mexicana*) in sulfide springs are a prudent example; studies have documented divergence between sulfidic and nonsulfidic ecotypes in nature in a variety of traits, including behavior (Plath et al. 2007b; Lukas et al. 2021), morphology (Tobler et al. 2008), swimming kinematics (Camarillo et al. 2020), metabolic needs (Passow et al. 2015), and life history (Riesch et al. 2014). While field studies are crucial for documenting phenotypic divergence in relevant ecological conditions, they are not sufficient to tease apart whether genetic differentiation (i.e., adaptation) and/or phenotypic plasticity causes divergence. Hence, in the second chapter of my dissertation, I conducted a lab experiment that specifically tested how plasticity mediated by maternal effects and genetic differentiation shape trait expression throughout ontogeny. I found stark divergence in multiple functional traits between sulfidic and nonsulfidic populations of *P. mexicana* that had been in captivity for several generations, and weaker evidence in fewer traits for maternal effects.

Overall, the results from my second chapter suggest that plasticity (in this case, from maternal effects) plays a role in adaptive divergence between ecotypes of *P. mexicana*, but that the majority of phenotypic divergence between ecotypes in nature is driven by local adaptation. Interestingly, maternal effects and genetic differentiation produced trait shifts in the same direction. Particularly because sulfidic and nonsulfidic environments are so different in nature, my data suggests that maternal effects—and perhaps plasticity in general—may enhance trait divergence in nature, which may reinforce genetic adaptation to sulfidic environments.

Consequences—Relating adaptation to sulfide to the emergence of reproductive isolation

Divergent natural selection can lead two populations to adapt to different environments, resulting in the coincidental emergence of reproductive barriers that can give rise to new species (Nosil 2012). Comparing populations from extreme environments to those in proximate benign habitats naturally lends itself to studying the processes of adaptive divergence and speciation. Results from the second chapter of my dissertation corroborated that the consistent phenotypic divergence between ecotypes of *P. mexicana* in the lab and in nature is likely due to adaptation. This is important in the context of ecological speciation because reproductive isolation must evolve as a byproduct of adaptation. Indeed, previous work in sulfide springs has shown that adaptation to sulfide has coincided with significantly reduced gene flow between sulfidic and nonsulfidic ecotypes of *P. mexicana* across multiple independent river drainages (Plath et al. 2013). However, the reproductive barriers that have evolved to cause gene flow to be reduced between these ecotypes remain unclear.

Precopulatory isolation is not strong enough to reduce gene flow to the levels observed in nature (Plath et al. 2013), so mechanisms occurring after copulation are likely occurring to limit

gene flow. In the fourth chapter of my dissertation, I measured several sperm competitive traits and quantified the effect of cryptic female choice on sperm competitive ability and resoundingly found no evidence for postcopulatory prezygotic isolation between ecotypes of *P. mexicana*. Ejaculate size and sperm swimming performance were not different between ecotypes, nor was sperm longevity during storage. Similarly, I found no evidence for cryptic female choice biasing sperm swimming performance in either ecotype. It remains unclear whether other, as yet unaddressed, precopulatory barriers to gene flow exist, or whether isolation after fertilization limits gene flow to the observed levels. Interestingly, nonsulfidic females show strong patterns of assortative mating toward sympatric nonsulfidic males, while they exhibit no preference for or against allopatric nonsulfidic males (Greenway et al. 2016). This is consistent with the reinforcement hypothesis, where secondary contact between diverged allopatric populations reveals costs to hybridization (reviewed in Nosil 2012). If hybridization is costly, then direct selection for female preferences to avoid hybridization should drive the evolution of assortative mating in sympatry, which acts to "reinforce" precopulatory isolation (Servedio and Noor 2003).

Future work and open questions

Despite the work presented in my dissertation, several questions regarding the causes and consequences of adaptation remain unanswered. While I documented putative mechanisms that may underlie tolerance of heavy metal pollution in mosquitofish in Chapter 2, I also found that Tar Creek individuals did not appear to be locally adapted, and that there was little genetic differentiation between mosquitofish living in Tar Creek and those living in unpolluted habitats. Regardless, mosquitofish are plentiful in Tar Creek, but it remains unclear how and why they remain abundant. Mosquitofish could be poorly adapted to life in Tar Creek and remain in high

abundance due to migration from other habitats, indicative of source-sink population dynamics. Likewise, mosquitofish may be permanent residents of Tar Creek that cope with chronic heavy metal exposure, but it remains to be seen how they may do so. Population genomics studies would help disentangle whether mosquitofish maintain stable populations in Tar Creek, and would benefit from expanding the number of populations of *G. affinis* sampled to get a broader and higher resolution understanding of gene flow genome-wide. Furthermore, particularly if there are permanent mosquitofish populations in Tar Creek, laboratory exposure experiments could help identify genetic and plastic responses to contamination that may shed light on the mechanisms underlying pollution tolerance.

In the third chapter, I found that trait divergence between ecotypes of *P. mexicana* is driven primarily by local adaptation, but there was nonetheless evidence for plasticity from maternal effects furthering trait divergence. This begs the question of whether plasticity caused by maternal effects or current environmental exposure has a larger relative impact on trait expression. To address this question, future work could employ experiments that expose mothers to environmental variation to induce maternal effects and measure trait variation in offspring that are also exposed to environmental variation throughout their ontogeny.

In the fourth chapter of this dissertation, I quantified the strength of postcopulatory prezygotic isolation caused by sperm competition and cryptic female choice. The patterns in these phenotypes that I measured suggested that postcopulatory prezygotic isolation is weak between sulfidic and nonsulfidic ecotypes of *P. mexicana*. However, as shown by Plath et al. (2013), natural and sexual selection against immigrants was significant but not strong enough to limit gene flow to the levels observed in nature, so it is unclear what other barriers have evolved to reduce gene flow between ecotypes. Isolation could arise due to precopulatory mechanisms

that have not yet been addressed, such as ecotypic divergence in competitive ability (Schluter 2001), local adaptation in response to divergent parasite communities (Tobler et al. 2007; Bomblies 2010; Karvonen and Seehausen 2012), and/or differential behaviors and mating preferences driven by the microbiome (Sharon et al. 2010). While I addressed several major drivers of postcopulatory prezygotic isolation in this dissertation, there could also be unaddressed barriers that have evolved between copulation and fertilization. These could include ecotypic differences in reproductive microbiomes that may lead to differential immune responses in males or females to homotypic vs. heterotypic inseminations (Rowe et al. 2020). There could also be strong selection against hybrids indicative of postzygotic isolation, which could arise through multiple mechanisms. First, hybrids could suffer from extrinsic incompatibilities if there is sexual selection against hybrids (Naisbit et al. 2001) or if hybrids have poorer performance than parental ecotypes (Via et al. 2000; Gow et al. 2007; Rajkov et al. 2018). Second, hybrids could also experience intrinsic mito-nuclear incompatibilities that could arise due to the breakup of coevolved mitochondrial and nuclear gene complexes during recombination after hybridization (Tobler et al. 2019). Even slight reductions in gene flow by any of the mechanisms presented here could together sufficiently limit gene flow to cause speciation, so future studies should methodically test for precopulatory, postcopulatory prezygotic, and postzygotic reproductive isolation simultaneously to understand how they jointly impact gene flow and cause speciation. Future work in this area will enable a clearer understanding of the true consequences of adaptation.

Chapter 2 - Impacts of heavy metal pollution on the ionomes and transcriptomes of Western mosquitofish (*Gambusia affinis*)

John L. Coffin, Joanna L. Kelley, Punidan D. Jeyasingh, & Michael Tobler

* Published as:
Coffin, J. L., Kelley, J. L., Jeyasingh, P. D., & Tobler, M. Impacts of heavy metal pollution on the ionomes and transcriptomes of Western mosquitofish (*Gambusia affinis*). *Molecular Ecology* 31(5), 1527-1542. doi: 10.1111/mec.16342.

Abstract

Our understanding of the mechanisms mediating the resilience of organisms to environmental change remains lacking. Heavy metals negatively affect processes at all biological scales, yet organisms inhabiting contaminated environments must maintain homeostasis to survive. Tar Creek in Oklahoma, USA, contains high concentrations of heavy metals and an abundance of Western mosquitofish (*Gambusia affinis*), though several fish species persist at lower frequency. To test hypotheses about the mechanisms mediating the persistence and abundance of mosquitofish in Tar Creek, we integrated ionomic data from seven resident fish species and transcriptomic data from mosquitofish. We predicted that mosquitofish minimize uptake of heavy metals more than other Tar Creek fish inhabitants and induce transcriptional responses to detoxify metals that enter the body, allowing them to persist in Tar Creek at higher density than species that may lack these responses. Tar Creek populations of all seven fish species accumulated heavy metals, suggesting mosquitofish cannot block uptake more efficiently than other species. We found population-level gene expression changes between mosquitofish in Tar Creek and nearby unpolluted sites. Gene expression differences primarily occurred in the gill, where we found upregulation of genes involved with lowering transfer of metal ions from the blood into cells and mitigating free radicals. However, many differentially expressed genes were

not in known metal response pathways, suggesting multifarious selective regimes and/or previously undocumented pathways could impact tolerance in mosquitofish. Our systems-level study identified well characterized and putatively new mechanisms that enable mosquitofish to inhabit heavy metal-contaminated environments.

Keywords: ecological transcriptomics, environmental physiology, ionomics, local adaptation, Tar Creek, tolerance to heavy metals

Introduction

Anthropogenic alterations of the environment have caused the degradation and destruction of ecosystems worldwide (Chapin et al. 2000; Falkowski et al. 2000; Urabe et al. 2010). Pollution is a particularly widespread form of human impact and affects biological systems across all levels of organization. Populations living in disrupted ecosystems typically face three contrasting endpoints: migration toward more suitable habitats (Lenihan et al. 2001), adaptation to altered environmental conditions (Whitehead et al. 2017), or extirpation (Wineland et al. 2019). There are numerous examples of population declines and extirpations caused by exposure to contaminants (e.g., Sánchez-Bayo, Goka, & Hayasaka, 2016; Silva et al., 2017), but some populations are able to persist even in heavily degraded ecosystems (see Oziolor and Matson 2015). Organisms that persist in the presence of environmental contaminants must maintain physiological homeostasis to support critical biological processes, and there are four hypothetical mechanisms by which organisms can do so: exclusion, detoxification, resistance, and mitigation. In animals, exclusion is typically associated with alterations of the integument that minimize or eliminate uptake or maximize excretion of contaminants, thus maintaining low endogenous concentrations (see Goffredi et al. 1997; Grieshaber and Völkel 1998). In contrast, tolerance involves several coping strategies that minimize the pathological effects of contaminants when organisms are unable to exclude them from their bodies. Tolerance could be achieved through: enzymatic detoxification of contaminants into less harmful forms (Grieshaber and Völkel 1998; Hildebrandt and Grieshaber 2008), resistance to toxic effects by modification of direct toxicity targets (as in Pfenninger et al. 2014), or utilization of alternative physiological pathways that are not directly affected by the contaminants (i.e., mitigation; see Hildebrandt and Grieshaber 2008). In most cases, we know little about the mechanisms allowing persistence in polluted

environments. While there are clear predictions regarding the effects of individual contaminants on specific biological pathways (see below and Wood et al. 2011a,b) comparatively little is known about systems-level adjustments in response to multiple contaminants, particularly in a natural setting. As such, we used an agnostic approach to test how two organismal systems (i.e., the transcriptome, and the ionome) respond to differences in the concentration of heavy metals.

Heavy metals (e.g., lead [Pb], zinc [Zn], and cadmium [Cd]) sourced from mining activities are common contaminants in the biosphere (Pacyna and Pacyna 2001). These elements have detrimental effects at all levels of biological organization, notably causing behavioral changes, reproductive damage, and increased mortality at an individual level, with cascading consequences for populations and ecosystems (Harmon 2009; Neuberger et al. 2009). In animals, the most toxic heavy metal ions are divalent cations that enter the body through branchial, cutaneous, and intestinal uptake (Gall, Boyd, & Rajakaruna, 2015; Wood et al., 2011a, 2011b). These ions then enter cells through calcium (Ca) channels by mimicking Ca^{2+} (see Wood, Farrell, & Brauner, 2011a, 2011b for reviews in fishes). Once inside the cell, the toxicity of metal ions is primarily caused by disruption of the redox balance (Birben et al. 2012). Heavy metal ions are oxidized in Fenton-like reactions, which produce hydroxyl radicals from the reaction of a metal (traditionally iron [Fe^{3+}]) and hydrogen peroxide (Fenton and Jackson 1899; Stohs and Bagchi 1995). Oxygen-containing free radicals like hydroxyl are known as reactive oxygen species (ROS) and are highly reactive due to the presence of unpaired electrons (e.g., $^{\bullet}OH$) and/or catalytic activity (e.g., H_2O_2) (Craig, Hogstrand, Wood, & McClelland, 2009; Mager, 2011; McGeer, Niyogi, & Smith, 2011). ROS have naturally beneficial roles in animals by participating in signaling cascades (Fernandez-Marcos and Nóbrega-Pereira 2016) and have even been shown to increase life span in *Drosophila melanogaster* in certain reaction pathways

(see Scialò et al., 2016). However, when ROS concentrations exceed the antioxidant capacity to reduce or bind them, ROS oxidize cellular components, including nucleic acids (Cooke et al. 2003), proteins (Davies 2012), and lipids (Regoli and Giuliani 2014), leading to oxidative damage. Heavy metal ions have also been shown to decrease ATP production by disrupting the electrochemical gradient across the inner mitochondrial membrane (Belyaeva et al. 2012; Kenderešová et al. 2012) and abstracting electrons bound for the electron transport chain (Meyer et al. 2013). Finally, heavy metal toxicity leads to a general suppression of the immune system via direct reduction of macrophage and lymphocyte activity (Sanchez-Dardon et al. 1999) and a decrease in antibody production (O'Neill 1981; Viale and Calamari 1984).

Metazoans have evolved strategies to counteract the toxic effects of heavy metals. Chief among these are antioxidant proteins (e.g., thiol-containing metallothionein and glutathione and radical-scavenging enzymes like superoxide dismutase and catalase) and nonenzymatic antioxidants like vitamins A, C, and E (Birben et al. 2012). There is also evidence of antioxidant regulation by hormones, such as melatonin, which can modulate the expression of antioxidant genes (Steinhilber et al. 1995) and directly scavenge free radicals with its reducing indole group (Poeggeler et al. 1996). These antioxidants are crucial for metazoan life, as many of these proteins are highly conserved, with functional orthologs present in almost all aerobic organisms (Corona and Robinson 2006). Despite these critical antioxidant defenses, heavy metal pollution often raises ROS levels beyond the binding and reductive capacity of natural defenses (Nemmiche 2016). An important question is whether and how individuals use these existing antioxidant mechanisms to persist in polluted habitats.

The Tar Creek Superfund Site in Ottawa County, Oklahoma, USA is characterized by high levels of Cd and Zn stemming from flooded mines and runoff from tailings piles, relics of

intensive mining operations in the area during the first half of the twentieth century (OWRB 1983; United States Environmental Protection Agency 2016). Runoff has caused increases in metal concentrations in the soil and water that pose significant risk to humans and wildlife inhabiting the area, and most tailings piles drain directly into Tar Creek (OWRB 1983; Datin and Cates 2002). Franssen (2009) demonstrated that the water in Tar Creek is extremely hard (i.e., high mineral content), which can buffer the toxic effects of metal cations (Hogstrand 2011). However, even after correcting for the extreme hardness of the water, concentrations of both Cd and Zn in polluted reaches of Tar Creek exceed their respective national recommended water quality criteria for aquatic life from the United States Environmental Protection Agency (Franssen 2009; United States Environmental Protection Agency 2019). Despite ongoing remediation efforts (Cremeans et al. 2019), surface waters of Tar Creek are still polluted with potentially lethal concentrations of heavy metals (Neuberger et al. 2009; United States Environmental Protection Agency 2016). In the fifty years since mining operations ceased, the Tar Creek watershed has exhibited altered fish community composition, decreased species richness, and impaired somatic and reproductive condition of the fishes that do persist (Franssen et al. 2006; Franssen 2009). While there are several fish species still inhabiting Tar Creek, Western mosquitofish (*Gambusia affinis*) are particularly abundant. Mosquitofish have been common in polluted sections of Tar Creek since at least the late 1960s (Branson 1967), and they remain common during sampling today, representing over 80 percent of individuals in polluted reaches (Franssen et al., 2006; personal observations JLC, 2017-2019). There are numerous uncontaminated watersheds in close proximity to Tar Creek (Franssen et al. 2006; Franssen 2009), allowing for comparative analyses that test hypotheses about how mosquitofish can persist in extreme environments like those found in Tar Creek.

Due to the extensive work on the physiological effects of heavy metals, we have a framework to form *a priori* hypotheses about the potential mechanisms that might mediate persistence of mosquitofish in Tar Creek. Here, we specifically asked whether mosquitofish in Tar Creek can actively maintain metal homeostasis through exclusion, and whether there is evidence for coping mechanisms associated with minimizing heavy metal toxicity (i.e., tolerance through detoxification, resistance, or mitigation). We hypothesized that mosquitofish have an increased ability to actively regulate their internal ion composition by blocking uptake of metals from the environment or eliminating metals from the body. Because no element operates in isolation (Baxter 2015), we tested this hypothesis by comparing ionomes—concentrations of all mineral nutrients in an organism (Lahner et al. 2003)—across several fish species that inhabit both polluted and unpolluted watersheds and predicted that metal accumulation would be lower in mosquitofish than in other species that are less abundant in Tar Creek. We interpreted ionomic data using the framework of ecological stoichiometry that makes predictions about the mobility of elements (both essential and toxic) at multiple biological scales, from tissues to communities (Peace et al. 2021). We also hypothesized that mosquitofish could detoxify metals that enter the body by modulating the expression of genes in relevant pathways. We tested this hypothesis by comparing genome-wide gene expression patterns in mosquitofish from polluted and unpolluted habitats. We predicted transcriptional modifications to primarily occur in genes associated with known heavy metal response pathways.

Methods

Overview of study system

The Tar Creek Superfund site consists of approximately 100 square kilometers in Northeastern Oklahoma, centered around the impacted Tar Creek drainage. We selected sites in the Neosho River drainage with minimal environmental differences other than the presence or absence of heavy metals that were previously documented to contain mosquitofish populations that have been used for comparative analyses (Franssen et al. 2006; Franssen 2009). These sites included a polluted section of Tar Creek (36.944° N, 94.854° W) as well as two sites in adjacent uncontaminated drainages, Coal Creek (36.850° N, 94.927° W) and Little Elm Creek (36.910° N, 94.806° W; see Figure 2-1). Coal Creek was used as a reference site for both ionomic and transcriptomic experiments, while Little Elm Creek was only used as a reference site for the transcriptomic experiment. Based on previous water chemistry monitoring in the area (Franssen 2009; United States Environmental Protection Agency 2016), we characterized levels of heavy metal pollution in each site as a binary variable (heavy metals present in Tar Creek or heavy metals absent in reference watersheds). All experimental protocols were approved by Kansas State University's Institutional Animal Care and Use Committee (protocol # 4379).

Do mosquitofish exhibit ionomic patterns indicative of active exclusion or increased elimination?

We tested whether mosquitofish had an increased ability to maintain low internal heavy metal concentrations compared to the other species inhabiting Tar Creek through exclusion or elimination. We predicted clear differences in elemental composition between fish from polluted and unpolluted sites across all species, and that mosquitofish would accumulate heavy metals to

a lesser extent. Lower levels of heavy metal accumulation would suggest that mosquitofish can actively exclude or eliminate them from the body at a higher rate than other species.

Sampling and processing

To test our ionomic predictions, we sampled Tar and Coal Creeks (Table 2-1) and collected individuals of each inhabitant fish species using a 3-mm seine net. At each site, up to seven individuals per species were placed in aerated buckets and transported to a facility operated by the Grand River Dam Authority in Miami, OK for processing. All fish were individually sacrificed by cervical dislocation, and standard length (mm) was recorded for each individual. Visceral organs were removed and discarded using new stainless-steel dissection scissors and forceps. Viscera were removed because their ionomic signatures could be labile, depending on daily fluctuations of behavior or diet, whereas the ionomic signature arising from the remainder of the body is likely more robust to such changes. The remaining parts of the eviscerated carcass were dried in a food dehydrator (National Presto Industries, Eau Claire, Wisconsin, USA) at 55 °C for 48 hours to prevent sample decay. Samples were brought to Kansas State University for further desiccation in a drying oven at 60 °C for at least two weeks, weighed, and digested in a 2:1 mixture of trace metal grade HNO_3 and H_2O_2 for analysis using an inductively coupled plasma optical emission spectrometer (ICP-OES; Thermo Scientific iCAP 7400), following the methods of Rudman et al. (2019) to generate ionomes for each sample. After removing elements from our analysis that were below the limit of detection of the ICP-OES, our ionomic data consisted of mass-standardized concentrations of 24 biologically relevant elements (see Table A-1).

Statistical analyses

If species experienced elemental shifts (in metals or non-metals) in Tar Creek, we predicted that we would be able to identify differences between sites in the concentrations of individual elements and suites of covarying elements. To test these predictions, we first analyzed variation in our overall ionomic dataset using principal component analysis. We used the prcomp function with a correlation matrix in R (version 4.0.0; R Core Team 2020) and retained six principal components with eigenvalues greater than 1, as in Kaiser (1960) (see Table A-1 for loadings for each element along each retained axis). The principal component scores for these six axes were used as dependent variables in subsequent analyses. We compared principal component scores between the contaminated and uncontaminated populations across all species to test for shared ionomic divergence across species between sites. Principal component scores were used as dependent variables in multivariate analysis of variance (MANOVA), and we included 'site' and 'species' as fixed factors. We found high covariance between species and standard length, which violated model assumptions. Hence, we calculated measures of relative standard length, and included this metric as a covariate to account for intraspecific allometry. All possible two-way interaction terms were included, and non-significant effects were removed from the final model. We used a type-III test to partition the sums of squares between the elements of our model using the Manova function from the car package in R (version 3.0-8; Fox & Weisberg, 2019) and Wilks' lambda to approximate *F*-statistics. Relative standard-length was not significant ($F_{6,57}$ = 0.788, P = 0.583), so the main effect and all interactions including the relative standard-length term were removed from the final model. Thus, our final model contained the main effects of site and species and an interaction between site and species. Effect sizes were estimated with a

multivariate analog of partial eta-squared using Wilks' lambda from the etasq function in the heplots package (version 1.3-5; Fox, Friendly, & Monette, 2018).

In the MANOVA, there was a significant interaction between site and species ($F_{36,288.2}$ = 5.338, $P < 0.001$), indicating that the effect of site (polluted vs. unpolluted) varied among species (see results). The significant interaction term might indicate that some species do not exhibit differences across sites while others do, or that site differences are omnipresent but vary in nature and magnitude. To distinguish between these alternative explanations and assess the nature of convergent elemental differentiation between sites across species, we conducted a *post-hoc* canonical analysis of the site term in the MANOVA (as in Langerhans, 2009). We ran a PCA (covariance matrix) on the sums of squares and cross products matrix for the site term, which yielded an eigenvector (called a divergence vector) for PC1 that summarized the linear combination of variables in a single canonical axis that maximized ionomic differences between Tar and Coal Creek individuals across all species, while controlling for all other factors in our model. We then multiplied our matrix of dependent variables by the divergence vector, which yielded a vector of divergence vector scores, with one score for each individual. Total divergence in elemental composition is comprised of both convergent and non-convergent (i.e., unique to a single species) components of elemental differentiation, and divergence vector scores extract only the convergent component. We utilized Student's *t*-tests to compare divergence vector scores between individuals from polluted and unpolluted sites for each species. Our ionomic dataset is archived on Dryad (https://doi.org/10.5061dryad.q2bvq83kw), and all analytical scripts can be found on GitHub (https://github.com/michitobler/TarCreekGambusia).

Are there differences in metal-responsive gene expression patterns?

We tested for gene expression differences between mosquitofish from polluted and unpolluted habitats, focusing on multiple organs in the body. In addition to differences between sites in heavy metal content, there could be other unmeasured environmental differences, so we predicted that among the various environmentally related transcriptional changes, genes involved with heavy metal binding and redox reactions would be upregulated in all organs of individuals from Tar Creek, compared to uncontaminated populations.

Sampling and processing

To quantify gene expression, we captured six adult female *G. affinis* from Tar Creek, Coal Creek, and Little Elm Creek (total $N = 18$ individuals; see Figure 2-1 and Table A-2 for sampling information). Two reference watersheds were utilized to better capture potential spurious, site-specific expression variation that could lead to false inferences when comparing to the contaminated site. Specimens were immediately euthanized by cervical dislocation upon capture, and standard length was measured. Gill, liver, and brain tissues were dissected and placed in separate tubes with 1 ml of RNAlater (Ambion), resulting in a total of 54 tissue samples. Samples were stored at 4 °C for approximately two weeks, at which point they were stored at -80 °C until further processing.

RNA extraction, library preparation, and sequencing

We followed RNA extraction (Macherey-Nagel Nucleospin RNA kit) and library preparation (New England Biolabs NEBNext Ultra Directional RNA Library Prep Kit with NEBNext Poly(A) mRNA Magnetic Isolation Module) protocols previously established and utilized for

other members of the family Poeciliidae (Passow et al. 2017), with the following minor modifications. During RNA extraction, the tissue sample input was smaller (between 3 and 25 mg of RNAlater-preserved tissue) because some samples (especially brains) of *G. affinis* were smaller than those of the species studied by Passow et al. (2017). Furthermore, libraries were enriched with twelve cycles of PCR amplification of adaptor ligated DNA. All library samples were stored at 4 °C to minimize freeze-thaw cycles. After assessing library quality on an Agilent Bioanalyzer High Sensitivity DNA chip, libraries were pooled into two groups of 27 samples that had equal proportions of samples from each site and tissue to minimize lane effects during sequencing. Paired-end, 100-bp sequencing was completed on two lanes of an Illumina HiSeq 2500 (Admera Health Biopharma Services, South Plainfield, NJ, USA) with one pool of 27 samples per lane.

Trimming, mapping, and assembly

Of the 54 samples sequenced, two Tar Creek gill samples and one Tar Creek brain sample were removed from the analysis due to improper indexing or low sequencing output. Thus, we analyzed gene expression in 16 gill samples, 18 liver samples, and 17 brain samples, for a total of 51 raw transcriptomes. Low quality reads (Phred score < 24) and adapter sequences were trimmed from the raw RNA-seq reads using TrimGalore! (version 0.4.1; Krueger, 2014). Only paired reads longer than 50 bp were retained. Across all tissues and sites, each sample had an average of 25,106,758 ± 9,999,182 reads. To reduce confounding effects of library size, we randomly subsampled 26 million reads (13 million paired-end reads) from 24 libraries that had significantly higher read counts than the other libraries using seqtk (Li 2013b). After subsampling, the samples had an average of 21,502,509 ± 5,314,437 reads. The remaining high

quality reads were mapped to the southern platyfish (*Xiphophorus maculatus*) reference genome (Schartl et al. 2013) using BWA-MEM (version 0.7.17; Li, 2013a) with default parameters. Reads mapped to a genome assembly of *G. affinis* (Hoffberg et al. 2018) yielded similar results, but the better assembly and annotation of the *X. maculatus* genome led to improved functional interpretation of gene expression results.

Using fixmate from samtools (version 1.2; Li et al., 2009), we added read groups to each alignment, removed soft-clipped bases, sorted aligned reads in coordinate order, and ensured all paired read alignments contained correct mate information. We used Stringtie (version 1.3.2d; Pertea et al., 2015) to extract expressed transcripts for each sample, and then used a python script (prepDE.py, provided with Stringtie) to generate a count matrix with all 27,266 genes for each sample. Table A-2 contains read counts and mapping efficiency throughout our analytical pipeline.

Quantification of differential expression

We removed genes with low expression (fewer than 2 counts per million [cpm], and present in fewer than 15 of the 51 transcriptomes), which removed 8,573 genes and brought the total number of genes in the dataset to 18,693. We then compared expression between Tar Creek and both uncontaminated creeks in each tissue separately. Using the edgeR package from Bioconductor (version 3.30.0; Robinson, McCarthy, & Smyth, 2010), we created a DGEList and calculated normalized factors based on library size for each sample. We then separated the data by tissue and created a design matrix (using the model.matrix function) for each tissue type separately, which consisted of an intercept of zero and site as the independent variable. For each tissue, we used the design matrix to calculate common and tag-wise dispersion across all three

sites using estimateDisp. Then we fit a negative binomial generalized linear model to the normalized counts for each tag, using the glmFit function. We generated contrast matrices using the makeContrasts function in the limma R package (version 3.44.1; Ritchie et al. 2015) to compare expression in each tissue between all sites in a pairwise manner. Using each contrast matrix of comparisons and the fitted negative binomial models for each tag, we tested for statistically significant differences in gene expression using likelihood ratio tests (with the glmLRT function), while accounting for error due to multiple testing with a false discovery rate <0.05 using the Benjamini-Hochberg correction (Benjamini and Hochberg 1995). To eliminate spurious gene expression differences due to inherent site-specific differences other than the factor of interest in our study (i.e., the presence of heavy metals), we used only genes that were significantly differentially expressed in the same direction in both unpolluted sites for our analyses (Tables A-3 and A-4). To visualize expression variation among sites in shared differentially expressed genes, we performed a discriminant analysis of principal components (DAPC) using the adegenet package in R (version 2.1.4; Jombart, 2008). We separated the differentially expressed genes by tissue and used the dapc function to linearly transform the count data using PCA and perform a linear discriminant analysis on a chosen subset of the principal components. For all tissues, we retained eight principal components and two discriminant functions.

Functional annotation of differentially expressed genes

To understand the functions of differentially expressed genes between polluted and unpolluted sites, we annotated the genes from the *X. maculatus* reference set with functional information from human orthologs in the SwissProt database (http://ca.expasy.org/sprot/), following Passow

et al. (2017). For our analyses, we used SwissProt annotations from two lists of genes in our dataset—consistently differentially expressed candidate genes in mosquitofish (the target set) and genes from the *X. maculatus* reference genome that were retained after filtering out lowly expressed genes (the background set). Each list of candidate gene annotations was uploaded as an unranked list to GOrilla (Eden et al. 2009), which identifies the Gene Ontology (GO; The Gene Ontology Consortium, 2004) terms associated with each SwissProt entry provided and tests for enrichment of GO terms in the target set relative to the background set. We restricted our analysis to GO terms related to biological processes. There were 11,186 genes in the background set that were uniquely associated with at least one of these GO terms. Using GOrilla, we tested for enrichment of GO terms with FDR <0.05. Of these enriched GO terms, we retained those with an enrichment score >2, and those that were associated with at least 5 genes in our target gene set, resulting in 621 total enriched GO terms (see Table A-5 for enriched terms in each tissue). Up- and downregulated genes in each tissue were analyzed separately.

Results

Mosquitofish exhibit a different shift in ionomic composition than other Tar Creek inhabitants

We generated ionomes of 4 to 7 individuals in each population of 7 species that inhabit Tar Creek (see Table 2-1). If mosquitofish can exclude metals, we predicted that all species except mosquitofish would experience a significant ionomic shift. To determine whether species exhibited ionomic shifts and which elements drove divergence between habitats, we analyzed ionomes between polluted and unpolluted populations across all species simultaneously. After principal component analysis, we retained 6 axes that cumulatively explained 81.8 % of the

variance in elemental composition (Table A-1). The first principal component explained 38.1 % of the variation in overall elemental composition (Figure 2-2), and principal component scores along this axis were most positively correlated with concentrations of Be, Bi, Ca, Co, Cr, S, Tl, Zn, though all elements had a positive loading along this axis. The second principal component explained 18.4 % of the variance in elemental composition. Positive scores along PC2 were associated with higher concentrations of B, Cd, Fe, and Si and lower concentrations of Ba, Mn, Ni, and Sr. Principal component 3, which explained 9.0 % of the overall variance, was positively correlated with As, Ca, S, and Zn, and negatively correlated with Co, Se, Sr, and V. There was significant segregation of samples from different sites along each of the first three principal components (Student's *t*-tests; PC1: $t_{89} = 4.574$, $P < 0.001$; PC2: $t_{89} = -7.415$, $P < 0.001$; PC3: $t_{89} = -4.583$, $P < 0.001$), and Tar Creek samples tended to have higher scores along PC1 and lower scores along PC2 and PC3. Segregation along these three axes indicated that Tar Creek individuals were enriched in As, B, Ca, Cd, Cr, Fe, S, and Zn, relative to Coal Creek individuals.

Analyses of principal component scores in the MANOVA (Table 2-2) identified a significant effect of site ($F_{6,65} = 6.978$, $P < 0.001$, $\eta_P^2 = 0.392$), indicating that—irrespective of species—Tar Creek and Coal Creek individuals differed in ionomic composition, and that site explained approximately 39 % of the non-error variance in ionomic composition after controlling for other factors in the model. This variance can be thought of as the shared ionomic divergence between sites among all species. We also found that ionomic composition varied by species ($F_{36,288.2} = 6.785$, $P < 0.001$, $\eta_P^2 = 0.362$), and there was a significant interaction effect between site and species ($F_{36,288.2} = 5.338$, $P < 0.001$, $\eta_P^2 = 0.312$), demonstrating that there were differences in how each population pair responded to heavy metal stress. *Post-hoc* comparisons of the divergence vector scores from the canonical axis from the site term of our model showed

significant differentiation in divergence vector scores between sites in every species except mosquitofish (for *Gambusia affinis*: $t_{10} = 1.891$, $P = 0.088$; for other species, see Figure 2-3), suggesting the presence of non-convergent ionomic changes in mosquitofish and that all other species exhibit significant shared ionomic variation.

Significant differential gene expression in each tissue

Due to the inherent complexity of environmental differences between our sites, we analyzed genes that were differentially expressed in the same direction in both uncontaminated sites compared to Tar Creek. We found 2,388 shared differentially expressed genes in gill samples, 499 in liver samples, and only 18 in brain samples (FDR < 0.05; Figure 2-4A). A separate ionomic analysis of gill, liver, and brain tissues solely in mosquitofish closely mirrored these results (see Figure A-1), showing divergence in elemental composition in gills and livers, but not in brains. The expression patterns of each mapped gene can be found in Table A-3, and visualization of divergence in gene expression across sites in each tissue can be found in plots of linear discriminant scores in Figure 2-4.

In addition to understanding how expression differed at each gene as a function of metal exposure, we were also interested in understanding whether any genes had shared expression responses to metal exposure. Co-expressed genes may be regulated by the same transcriptional machinery or be related functionally, as they likely belong to the same biological pathway (Langfelder & Horvath, 2007). We generated weighted gene co-expression networks using weighted gene co-expression network analysis (i.e., WGCNA; see online supplement for methods) and obtained modules of genes that showed similar expression profiles and were correlated with the presence of environmental heavy metals. We found that our network-based

analyses of gene expression largely matched our differential expression analyses. We found nine modules of co-expressed genes across all three tissues sampled that were associated with the presence of environmental heavy metals (Figure A-2), but functional analysis of genes within these modules did not identify any genes with known functions related to heavy metal homeostasis (see online supplement for details).

Antioxidant genes are upregulated in the gill in response to heavy metal stress

To understand the functional implications of the observed gene expression patterns, we annotated the differentially expressed genes with human SwissProt accessions and biological process GO terms. We found that several known antioxidant genes were upregulated in gill tissues, including metallothionein, glutaredoxin, nucleoredoxin, glutathione peroxidase, and glutathione S-transferase (see Table A-4 for SwissProt annotations of differentially expressed genes). Products of these genes have been shown to guard against oxidative stress by binding metal ions prior to the production of ROS or rendering free radicals inert by donating spare electrons (Hogstrand 2011; Mager 2011).

Using a broader systems-level approach to understand patterns of differential expression, we expected to see enriched GO terms and modules of co-expressed genes (see online supplement) related to heavy metal exposure. We found enrichment of eight GO terms in the upregulated genes in the gill, 383 terms in downregulated genes in the gill, and 230 terms in downregulated genes in the liver of fish from the polluted site (Table A-5). There was no significant enrichment for upregulated genes in the liver or for up- or down-regulated genes in the brain. To search the enriched GO terms for biological processes related to our *a priori* hypotheses regarding heavy metals, we queried the Gene Ontology database using AmiGO2 and

searched for “metal” (http://amigo.geneontology.org/amigo; accessed 16 September 2021), which yielded a list of 66 biological process GO terms broadly related to metals (see Table A-6). We intersected this list with the 621 enriched GO terms and found enrichment of three biological process GO terms related to metals. There were 39 genes that were downregulated in gill samples that were associated with the “regulation of metal ion transport” GO biological process, and 57 such genes that were associated with “cellular metal ion homeostasis”. The “cellular metal ion homeostasis” GO term was also enriched in downregulated genes in the liver, which was driven by downregulation of 24 genes in the liver. Finally, there were 25 genes that were downregulated in the liver that were associated with the “metal ion homeostasis” GO term.

The majority of the enriched GO terms were related to diverse biological processes not typically invoked in responses to heavy metals. The top five biological processes over-represented in upregulated genes in the gill were angiogenesis, regulation of angiogenesis, regulation of nuclear cell cycle DNA replication, regulation of vasculature development, and neuron projection guidance. In downregulated genes in the gill, the top five biological processes that were enriched in our dataset were immune system process, immune response, cell activation, defense response, and leukocyte activation, which was largely mirrored by the top five GO terms in downregulated genes in the liver—immune system process, immune response, defense response, signal transduction, and positive regulation of immune system process. Annotation and enrichment results for each tissue are summarized in Table A-5.

Discussion

Environmental perturbations represent a significant challenge to the maintenance of homeostasis. While it is clear that organisms residing in polluted environments must maintain homeostasis

despite the suboptimal conditions, we have a limited understanding of the mechanisms by which they are able to do so. We utilized an integrative approach to understand whether mosquitofish are able to actively maintain homeostasis by blocking metal uptake from the environment or tolerate internal metal accumulation through transcriptional changes. We compared whole body elemental composition and transcriptome changes between populations of *G. affinis* that inhabit a heavy metal-polluted watershed and nearby unpolluted habitats to understand the mechanisms behind the persistence of these fish in the polluted watershed. We found evidence for the accumulation of multiple elements in all species in Tar Creek, but mosquitofish exhibited a pattern of ionomic divergence distinct from that of other species. We also identified significant upregulation of several antioxidant genes that are known to minimize cellular intake of metal ions and maximize detoxification of ROS. Overall, this study uncovered putative molecular mechanisms underlying the persistence of mosquitofish in Tar Creek.

Maintenance of homeostasis through exclusion of metals from the body

We found evidence for multivariate ionomic shifts in the fishes of Tar Creek, where Tar Creek individuals across all species were generally enriched in As, B, Ca, Cd, Cr, Fe, S, and Zn, and relatively depleted in Al, Ba, Co, Mn, Ni, Se, Sr, and V (Figure 2-2). This suggests that the water chemistry in Tar Creek is associated with changes in tightly regulated suites of elements, which has implications for how fishes build and maintain their bodies in heavy metal-contaminated water. Each species, including mosquitofish, experienced a shift in ionomic composition between populations, but the shift in mosquitofish was unique. Mosquitofish were the only species that did not show a significant difference in the site-specific divergence vector scores (i.e., the convergent differences in elemental composition across species); all other species showed

evidence of significant convergence in ionomic shifts between sites (Figure 2-3). To better understand the unique nature of divergence in mosquitofish compared to the other species, we plotted the average principal component scores for each population along the first three principal component axes (Figure 2-5). Along each PC axis, the non-convergent nature of ionomic shifts in mosquitofish becomes apparent, as both populations of mosquitofish appear substantially different from the other species in their respective habitats. Furthermore, despite showing significant differences in divergence vector scores between sites (see Figure 2-3), *Pimephales notatus* also appeared to show large non-convergent differences in multivariate space, which warrants further study. These findings suggest that there are unique ionomic shifts in mosquitofish, but that metals are nonetheless accumulated, demonstrating that mosquitofish—as all other species in this study—lack the ability to actively maintain metal homeostasis by blocking uptake from the environment.

Other animals also show negligible evidence for exclusion of heavy metals, though several studies have found evidence of such mechanisms in the metal-chelating properties of algal cell walls (e.g., Pinto et al., 2003). Originally, we considered the possibility that mosquitofish may possess novel adaptations to limit uptake of environmental contaminants, and that mosquitofish would be the only species to not exhibit elemental differences, thus explaining the higher frequency of mosquitofish in Tar Creek compared to other resident species. However, mosquitofish still differed in ionomic composition between sites (see Figure 2-2 and Table A-7). The bioaccumulation of heavy metals in mosquitofish suggests that exclusion is not used as a mechanism to maintain physiological homeostasis.

Maintenance of homeostasis through detoxification

Persistence of mosquitofish in Tar Creek could also be mediated by modulating gene expression—and ultimately physiological function—to mitigate the adverse effects of heavy metal exposure and increased internal concentrations, allowing mosquitofish to maintain homeostasis through detoxification. In each tissue, we identified genes that were significantly differentially expressed in the same direction in both unpolluted sites, with the majority of these shared differentially expressed genes found in gill tissues.

In gills, where the highest accumulation of metals occurred, we found a coordinated downregulation of the *ZIP3* gene, which codes for a zinc transport protein that allows zinc to enter the cytoplasm, and upregulation of several antioxidant proteins (Tables A-3 and A-4). We also found upregulation of enzymes that are able to reduce ROS into less harmful intermediates (e.g., glutathione peroxidase and several isoforms of glutathione S-transferase), as well as metal-binding proteins such as metallothionein (see Tables A-3 and A-4), which matches with established patterns of gene expression and protein activity following heavy metal exposure in fishes (Basha and Rani 2003; Woo et al. 2009; Uren Webster et al. 2013; Ransberry et al. 2015). Counterintuitively, peroxiredoxin-6, which is important in the reduction of hydrogen peroxide with glutathione (Chen, Dodia, Feinstein, Jain, & Fisher, 2000), was downregulated in the gill, suggesting that some canonical antioxidant proteins may be triggered or repressed upon metal exposure, depending on the context of exposure. Consistent with our data, previous studies have found evidence for upregulation of some, but not all, antioxidant genes following heavy metal exposure, and the antioxidants identified varied by study. There are few studies that found evidence for differential expression of all antioxidants known to play a role in heavy metal homeostasis, suggesting that there are likely distinct responses to different mixtures of metals

and that there could be functional redundancy in antioxidant responses. Alternatively, species-specific evolutionary histories could produce different genomic architectures, leading to historical contingencies in terms of which genes are responsive to heavy metal exposure in different species.

While the whole-body ionomic data showed that mosquitofish experience an influx of metal from the environment into the body, the gene expression results identified downregulation of genes coding for proteins related to transport of ions from the blood into cells and upregulation of several genes that code for antioxidant proteins. If the gene expression patterns noted here lead to proteomic differences, then the amount of heavy metals that enter cells from the bloodstream may be limited, thus lessening the adverse effects of the heavy metals once they are in the cell.

Liver and brain tissues did not exhibit the transcriptional responses to heavy metal exposure that we hypothesized or observed in the gills. While there was significant accumulation of metals and many genes were differentially expressed in the liver, the vast majority of the gene-by-gene and systems-level (GO enrichment and WGCNA) expression patterns we identified revealed downregulation of genes related to Ca^{2+} homeostasis (see Tables A-3, A-5, and A-8) and a general suppression of the immune system, which is a common response when organisms are stressed, by heavy metals or otherwise (Sanchez-Dardon et al. 1999). In contrast to our findings in the gill, however, we did not find significant differential expression of any genes known to play a role in heavy metal responses. This discrepancy in responses between tissues could be caused by gills being in direct contact with heavy metals in the environment.

The majority of transcriptional differences in the liver did not appear to be related to heavy metal homeostasis, but we found evidence of differential expression of at least one gene in

the liver related to metabolism with known ties to heavy metal stress. There is ample evidence of disruption of mitochondrial function upon exposure to heavy metals (e.g., Belyaeva et al. 2012; Kenderešová et al. 2012; Meyer et al. 2013), and we found significant upregulation of cytochrome c, which is a protein associated with the inner mitochondrial membrane that shuttles electrons from complex III to complex IV of the electron transport chain (Hüttemann et al. 2011). In addition to furthering aerobic respiration through the oxidative phosphorylation pathway, the upregulation of cytochrome c could also represent an antioxidant response, as cytochrome c has been shown to reduce superoxide radical, thus diminishing total cellular ROS (Skulachev 1998; Andreyev et al. 2005). We hypothesize that the upregulation of cytochrome c could be an adaptive mechanism by which mosquitofish are able to maintain aerobic respiration despite the disruptions to mitochondrial function caused by heavy metals. However, further studies specifically focusing on mitochondrial physiology are required to further explore and validate this hypothesis.

Addressing alternative hypotheses

In addition to the documented upregulation of key antioxidant genes in the gill, we also saw modulation of hundreds of genes across tissues that did not appear to be related to heavy metal stress and homeostasis. There are two non-mutually exclusive hypotheses that could explain these patterns. First, mosquitofish in Tar Creek are likely under stress, which has cascading effects genome-wide (Aluru and Vijayan 2009). Likewise, the presence of heavy metals may not be the only strong source of selection between polluted and unpolluted habitats. The Tar Creek watershed is not a static system with constant input of a few major pollutants; it is a dynamic patchwork of fluctuating heavy metal concentrations, pH, oxygen and food availability,

ecological interactions, and human impact. While the unpolluted sites were selected carefully to minimize environmental differences other than heavy metal pollution, there may be unmeasured ecological differences between sites that could cause physiological responses in inhabitants. There could be numerous stressors other than heavy metals that differentially impact the organisms inhabiting Tar Creek compared to the unpolluted sites, which could lead to diverse gene expression responses that do not belong to a single pathway. This pattern has been seen in other poeciliid species inhabiting extreme environments (Passow et al. 2017; Tobler et al. 2018). Some abiotic stressors are environmentally correlated (e.g., hypoxia and toxic hydrogen sulfide are positively correlated in freshwater: Chen & Morris, 1972; Cline & Richards, 1969); individuals coping with toxicity from one physiochemical stressor must sometimes also survive in suboptimal conditions, leading to broad gene expression responses. In addition to selection from physiochemical stressors, we likewise see reduced species richness in many extreme environments (Greenway et al. 2014), which has been shown to impact behaviors (Plath et al. 2007b; Tobler et al. 2009a) and ecological interactions (Tobler and Plath 2011; Tobler et al. 2015; Laske et al. 2018). We posit that similar mechanisms could be at play in Tar Creek, leading to broad stress-related gene expression patterns. Future work should focus on the interplay of these multifarious interactions and include controlled common garden laboratory experiments.

In addition to the transcriptional modifications that could enable survival of *G. affinis* in Tar Creek, mosquitofish may also be able to persist in heavy metal-contaminated habitats due to species-specific ecological attributes that went unmeasured in our experiments. Mosquitofish have a high reproductive potential with long reproductive seasons (Pyke 2005), representing an opportunistic life history strategy that is particularly well suited for repopulation of disturbed

habitats (Winemiller and Rose 1992). Furthermore, as the only livebearing fish species in Tar Creek (Franssen et al. 2006), internal gestation could enable persistence of mosquitofish by providing a buffer between the developing embryos and the toxic environment (Greenway et al. 2014), though this remains to be tested experimentally. Finally, because of the life history traits noted above and the high connectivity between Tar Creek and neighboring uncontaminated sites, mosquitofish may not be adapted (generally or locally) for life in Tar Creek at all; instead, they may be a sink population that merely persists due to continuous migration from populations in unpolluted habitats. Indeed, an examination of population genetic variation in our transcriptome data (see online supplement for details) indicated that there was only one genetic cluster represented in our data, suggesting there is no genetic differentiation or population structure in the genomic regions captured by the transcriptome. Future work to address the mechanistic basis of pollution tolerance should utilize common garden laboratory exposure experiments to identify local adaptation through plastic and genetic responses to pollution and population genomic data to understand patterns of gene flow genome-wide. An open question that remains is whether gene expression variation in this system is purely driven by environmental differences among populations or whether genetic and/or epigenetic differences in gene regulation across these subpopulations also play a role.

Conclusions and other considerations

Our analyses of elemental differentiation across multiple species inhabiting Tar and Coal Creeks demonstrated that all species experienced an ionomic shift that included accumulation of heavy metals; however, the axis of ionomic divergence was different in mosquitofish than other species, suggesting mosquitofish cannot limit uptake of metals from the environment. We also

found evidence of differential gene expression in gill, liver, and brain tissues between individuals from polluted and unpolluted sites, providing a first look into transcriptional mechanisms that could mediate tolerance of mosquitofish to life in heavy metal-contaminated environments. The gills exhibited the largest transcriptomic response and appeared to be involved in mediating heavy metal tolerance by upregulating genes involved with limiting uptake of heavy metal ions from the blood and binding and reducing free radicals. However, several known antioxidant proteins (e.g., superoxide dismutase, catalase, and glutathione) were not differentially expressed in our study, and we also found differential expression of hundreds of genes with diverse functions other than metal homeostasis. This indicates the inherent complexity of organismal responses to diverse mixtures of pollutants like those found in Tar Creek and highlights the need for further research to understand the intricacies of responses to contaminant exposure. Despite the lack of data on local adaptation in this system, the ionomic and transcriptomic shifts noted in the present study likely play a major role in pollution tolerance in mosquitofish. Building on our existing systems framework by integrating genomic data would vastly improve our understanding of how mosquitofish maintain homeostasis and survive in Tar Creek.

Pollution tolerance is a complex phenotype involving the interplay of atoms, genes, RNA and protein molecules, tissues, organisms, and ultimately larger ecological entities. Systems-level approaches have been instrumental to our understanding of these responses by integrating multiple lower-level components to elucidate phenomena at higher levels of biological organization (Pinu et al. 2019). By analyzing multiple elements simultaneously, our study improves the resolution over modular studies of individual contaminants, which may lack biological significance, since contaminants rarely exist in isolation (Escher, Stapleton, & Schymanski, 2020; Wood, 2011). The results of this study—and particularly the complex

changes in gene expression—demonstrate the need for future studies to test for effects of multiple concomitant stressors simultaneously.

Acknowledgements

We are deeply indebted to the Native American tribal communities of northeastern Oklahoma, on whose ancestral land the majority of this work was accomplished. We also thank the local landowners of Commerce and Miami, OK for access to our study sites and the Grand River Dam Authority for logistic support and accommodations. Without the assistance provided by E. J. Wilson, H. Camarillo, J. Grill, M. Laughlin, A. Arias, and D. Stockman in the field and laboratory, this work would not have been possible. We would also like to thank R. Sherman, who generated the ionomes. Permits were granted by the Oklahoma Department of Wildlife Conservation. This research was supported by the National Science Foundation (IOS-1557860 and IOS-1931657 to MT). Additional support was provided to JLC through a Graduate Assistance in Areas of National Need (GAANN) fellowship and travel awards through the Biology Graduate Student Association, College of Arts and Sciences, and Graduate School of Kansas State University.

Data Accessibility

All transcriptomes have been deposited in GenBank (BioProject accession: PRJNA707024). The Xiphophorus maculatus assembly used as a reference genome for our analyses was also obtained from Genbank (BioProject accession: PRJNA72525). Raw ionomic and transcriptomic datasets are archived on Dryad (https://doi.org/10.5061/dryad.q2bvq83kw), and scripts used for all analyses are available on GitHub (https://github.com/michitobler/TarCreekGambusia).

Author Contributions

JLC and MT designed the experiments and collected samples in the field. Reagents were provided by JLK, PDJ, and MT. Computing resources were provided by JLK and MT. JLC, JLK, and PDJ conducted laboratory work. JLC and MT conducted the analyses and wrote the manuscript. All authors made substantial contributions to revisions of the manuscript.

Figures

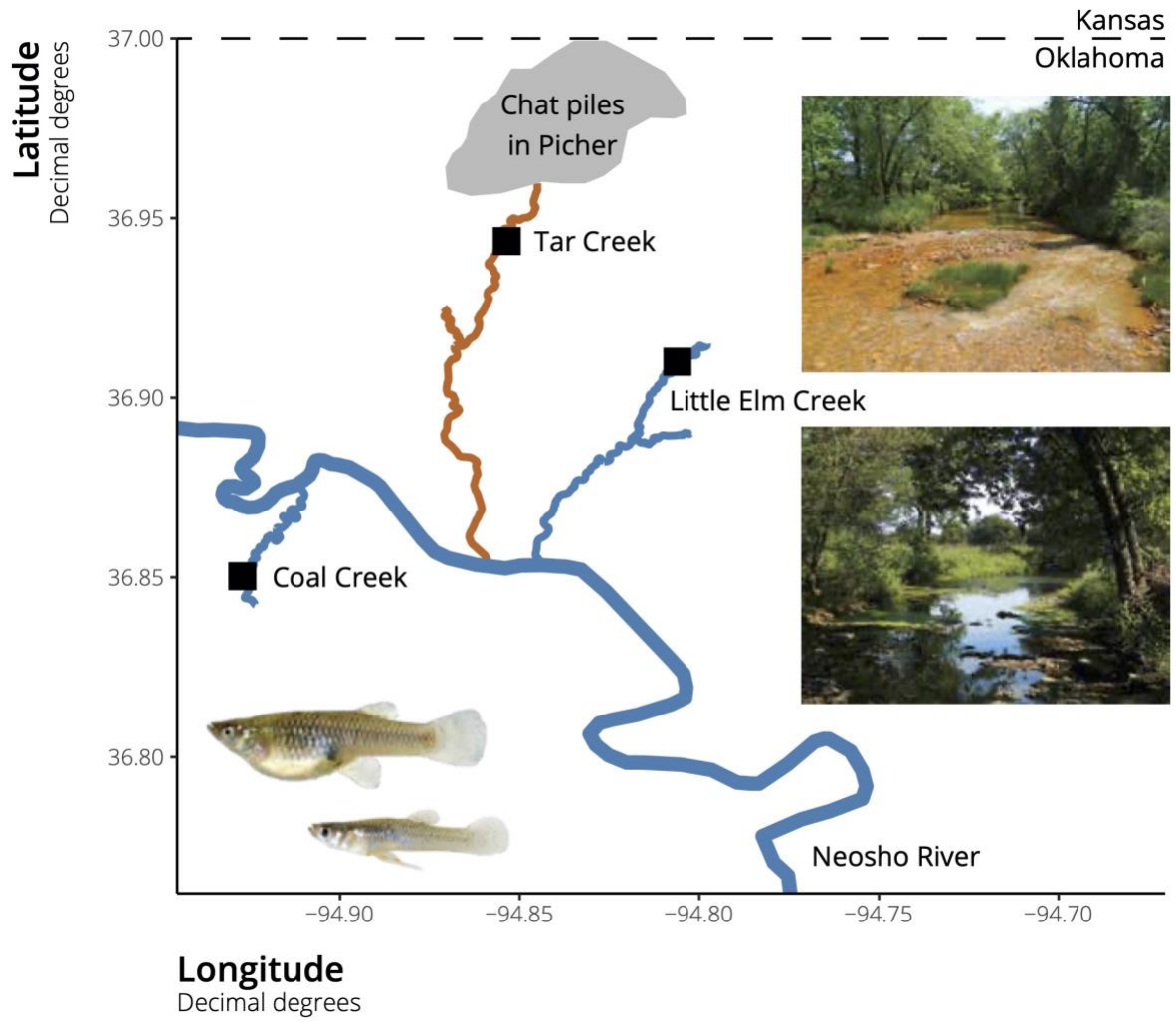

Figure 2-1: Map of the Tar Creek Superfund Site. Acid mine drainage directly enters Tar Creek through two routes: i) runoff from surface piles of mining waste known as chat piles in the area of Picher, OK, and ii) through boreholes to inundated mines throughout the area. The Neosho River and tributaries sampled in this study (Tar, Coal, and Little Elm Creeks) are represented by latitude and longitude digitized from Google Earth. Black squares represent sampling locations on each tributary. A male and female *G. affinis* are pictured in the bottom left, and Tar Creek (top-right) and Coal Creek (bottom-right) are pictured for reference.

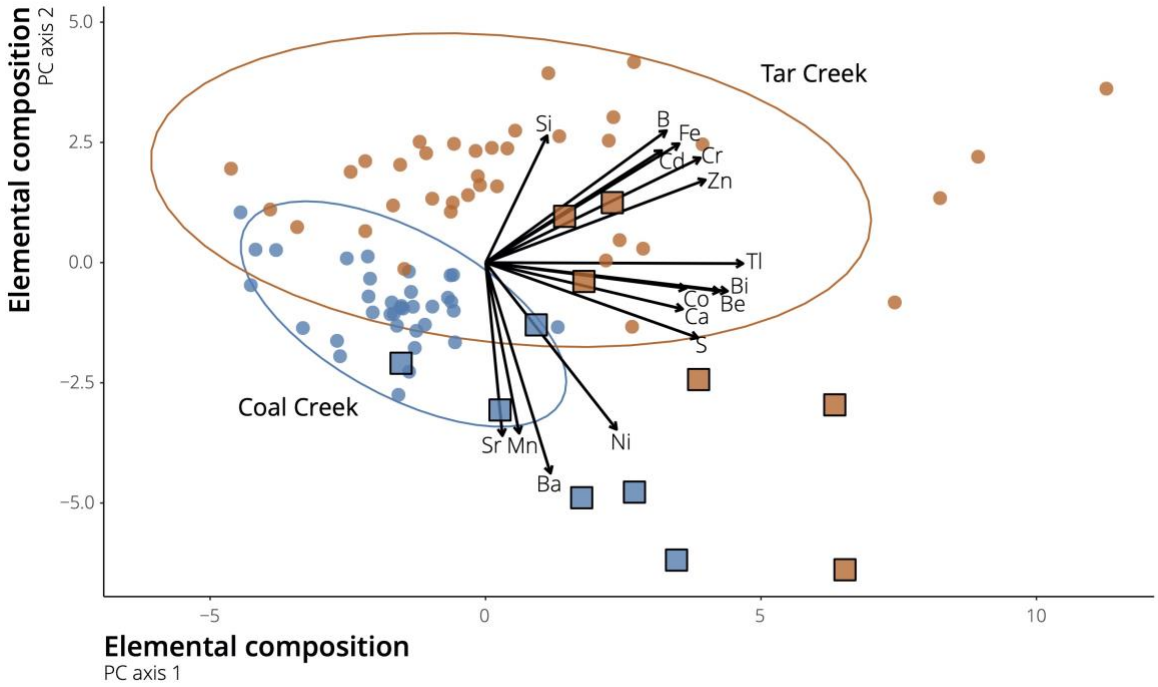

Figure 2-2: Plot of individual elemental principal component scores along the first two principal component axes. Observations in orange are from Tar Creek, and observations in blue are from Coal Creek. Mosquitofish from both sites are represented by squares, while all other species are shown as circles. Ellipses show 95 % confidence intervals in both dimensions for each site. Loadings are shown in black for the eight elements with the largest absolute loadings along each axis. Loadings were calculated as the eigenvector multiplied by the square root of the eigenvalue for that axis. The magnitude of each loadings arrow was magnified by a factor of five for ease of viewing.

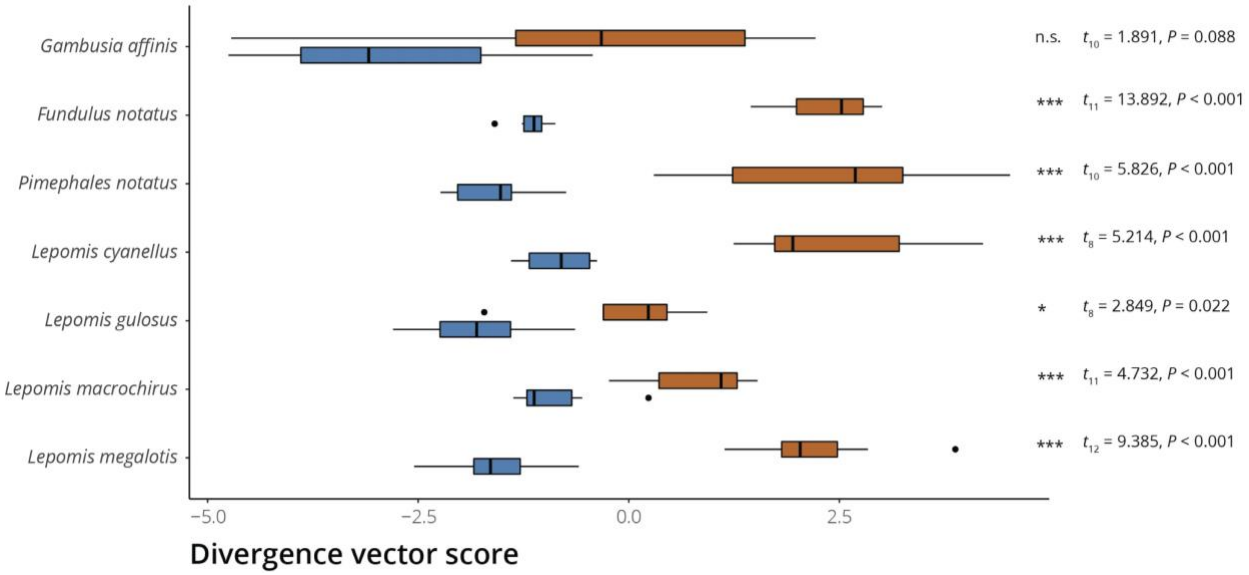

Figure 2-3: Boxplots of divergence vector scores comparing convergent elemental differentiation between sites across all species. To obtain divergence vector scores, we used multivariate analysis of variance to test for the main effects of site and species and the interaction between site and species on the principal component scores from the first six axes. We ran a PCA on the sums of squares and cross products matrix for the 'Site' term from this MANOVA and multiplied the resulting eigenvector for PC1 by the matrix of principal component scores for the first six axes to obtain divergence vector scores for each individual. See the Methods section and Table 2-2 for more details. Scores were plotted for species (shown on the y-axis) and sites (blue corresponding to Coal Creek and orange corresponding to Tar Creek individuals) separately to show differences between sites in each species. Separate *t*-tests were conducted on each species, and statistics and *P*-values are shown at right. *Gambusia affinis* is the only species that did not have significantly different divergence vector scores between sites.

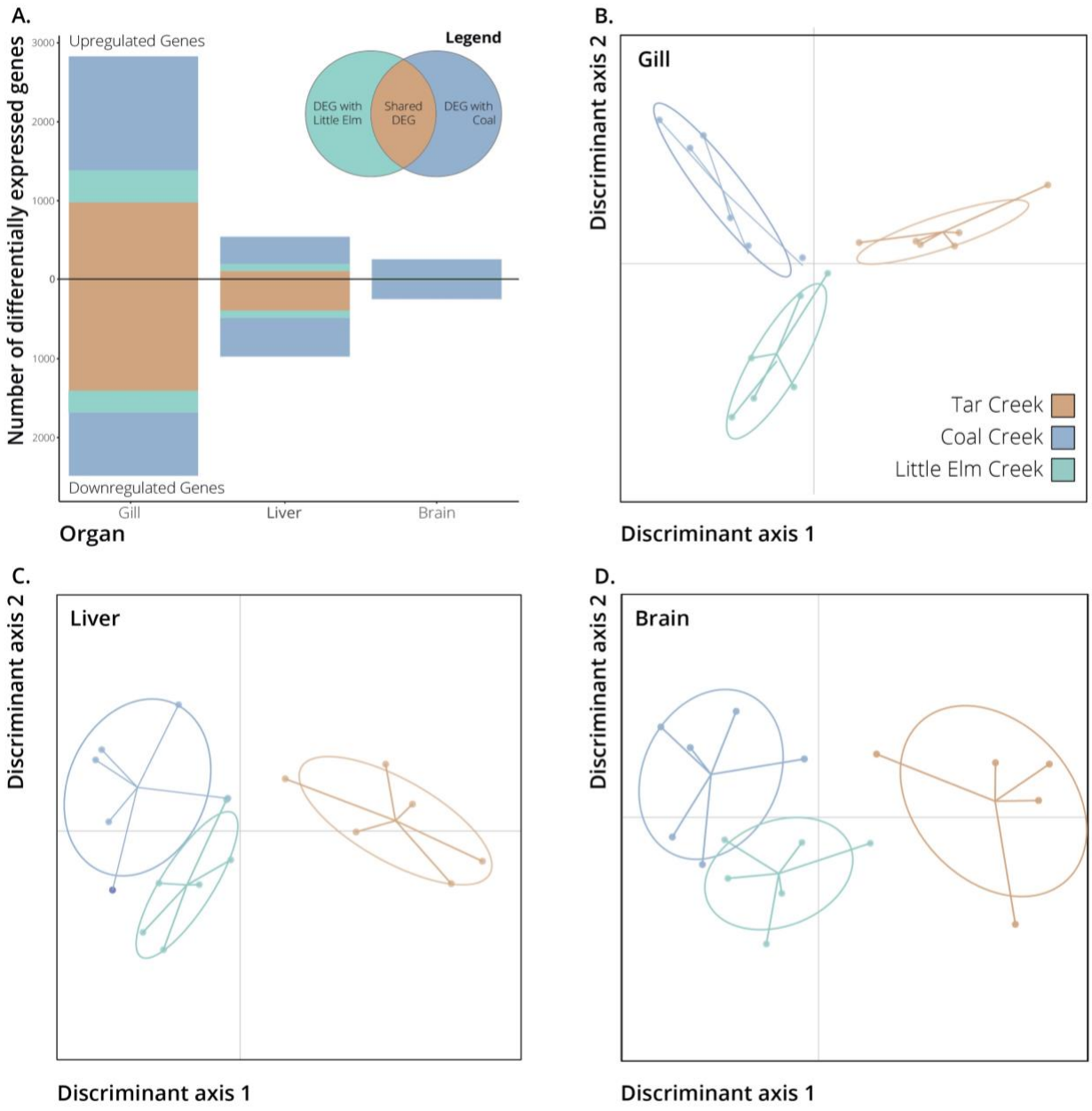

Figure 2-4: (A) Stacked bar plot of the number of differentially expressed genes (DEG) in each tissue. Genes above zero on the y-axis were upregulated in Tar Creek individuals, and genes below zero were downregulated in Tar Creek individuals. The orange portion of each bar shows the number of genes that were differentially expressed in both Coal and Little Elm Creeks in the same direction, compared to Tar Creek. The blue portion of bars signifies the number of genes differentially expressed only in Coal Creek individuals compared to Tar Creek individuals. The turquoise portion of each barplot represents genes that were only differentially expressed in Little Elm Creek compared to Tar Creek. Plots of discriminant scores from a discriminant analysis of principal components (DAPC) to visualize multivariate divergence in gene expression in (B) gill, (C) liver, and (D) brain samples.

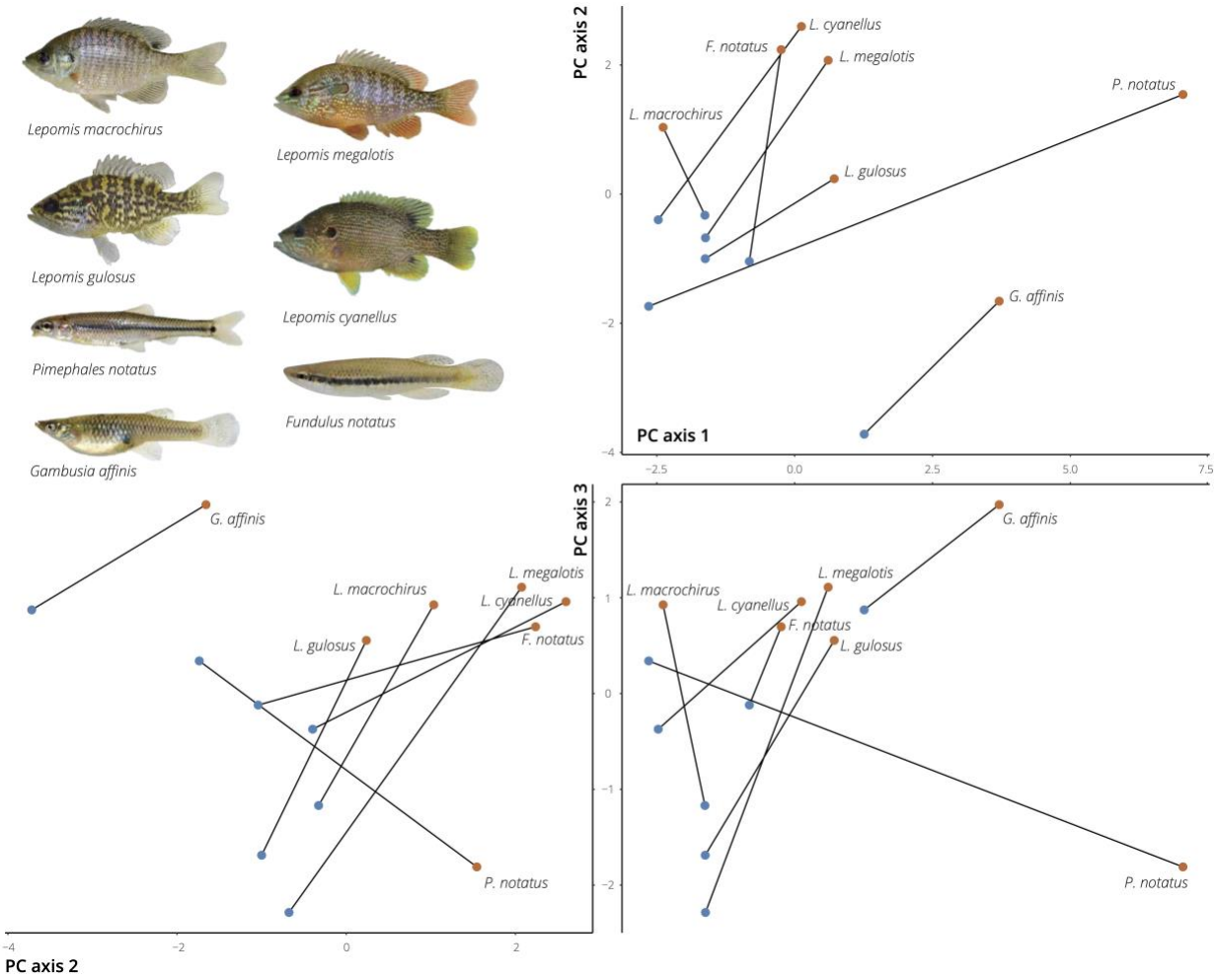

Figure 2-5: Images of each sampled species are included in the top left. Plot of average principal component scores generated from whole organism ionomes for each population in each of the first three principal component axes. Populations of the same species are connected with a line to demonstrate species-specific differences. Populations from Tar Creek are colored orange, while Coal Creek populations are colored blue.

Tables

Table 2-1: Sample sizes for ionomic data and taxonomic information for each species in both Tar and Coal Creeks

Family	Species	N_{Tar}	N_{Coal}
Poeciliidae	*Gambusia affinis*	6	6
Fundulidae	*Fundulus notatus*	7	6
Cyprinidae	*Pimephales notatus*	6	6
Centrarchidae	*Lepomis cyanellus*	6	4
Centrarchidae	*Lepomis gulosus*	4	6
Centrarchidae	*Lepomis macrochirus*	7	6
Centrarchidae	*Lepomis megalotis*	7	7

Table 2-2: Type-III multivariate analysis of variance summary table comparing principal component scores of ionomic differentiation between sites and species. We used the Wilks' lambda test statistic and estimated effect sizes with a multivariate analog of η_p^2, as calculated with the Wilks' lambda test statistic (Λ): $\eta_p^2 = 1 - \Lambda^{1/s}$, where s is the number of levels of each factor minus 1, or the number of dependent variables, whichever is smaller.

	df	Wilks' Λ	approx F	num df	den df	Pr(>F)	η_p^2
Intercept	1	0.64	6.083	6	65	<0.001	0.360
Site	1	0.608	6.978	6	65	<0.001	0.392
Species	6	0.068	6.785	36	288.2	<0.001	0.362
Site × Species	6	0.106	5.338	36	288.2	<0.001	0.312

Chapter 3 - Genetics and resource availability shape life history and behavioral divergence between ecotypes of Atlantic mollies (*Poecilia mexicana*)

John L. Coffin & Michael Tobler

Abstract

Phenotypic variation in nature is common along environmental gradients, but it is unknown whether it results from genetic differentiation between populations (i.e., adaptation) or population-specific effects of environmental exposure (i.e., phenotypic plasticity). Lab studies are crucial for disentangling these alternatives because they enable controlled environmental manipulations to understand the relative roles of genetic differentiation and plasticity on trait variation. We studied populations of Atlantic mollies (*Poecilia mexicana*) in Southern Mexico that have colonized streams rich in hydrogen sulfide (H_2S), a respiratory toxin that disrupts aerobic ATP production. Colonization of H_2S-rich habitats has coincided with stark phenotypic differentiation between sulfidic and nonsulfidic ecotypes of *P. mexicana*, but it is unclear whether trait variation is caused by adaptation or plasticity. We varied food availability to pregnant *P. mexicana* to induce maternal effects, a form of plasticity that occurs when a mother's environment impacts phenotypes in her offspring, and repeatedly measured several traits throughout offspring ontogeny. Genetic differentiation affected most of the traits we measured, as sulfidic offspring tended to be born larger, mature later, have lower burst swimming performance, be more exploratory, and feed less accurately. Likewise, maternal effects impacted several key functional traits; offspring from poorly provisioned mothers tended to be born larger and be more exploratory. Population differences and maternal effects (when both were present)

were additive, and while population differences had a larger effect than maternal effects, there was no interaction between populations and food treatments. This suggests that the majority of phenotypic divergence between ecotypes of *P. mexicana* in nature is caused by local adaptation, and that plasticity (e.g., from maternal effects) heightens—but does not cause—differences between populations.

Keywords: ecological gradients, extreme environments, ecological speciation, trait divergence, maternal effects, population differentiation, resource availability, adaptation, Poeciliidae

Introduction

Phenotypic variation is at the heart of investigations of adaptation because it links the cause (natural selection) to the consequence (genotypic change) of adaptive evolution (Lande and Arnold 1983). We have known that inheritance causes resemblance between parents, offspring, and siblings for over a century (Mendel 1865), reflecting a genetic component to phenotypic variation. However, not all trait variation is caused by genetic variation; it can also arise from phenotypic plasticity, where a single genotype gives rise to alternate phenotypes in response to internal or environmental cues (Pigliucci 2001). Across the tree of life, developing organisms can sense the environment, and plasticity enables them to enter alternative developmental trajectories (Weiss 2019). In addition, the environment experienced by parents can affect phenotypes in their offspring (i.e., parental effects: Uller 2008), representing a special case of plasticity that is shaped across generational boundaries. Phenotypic variation in natural systems can therefore arise from genetic differences among individuals, plasticity induced by individual exposure to different environmental conditions, plasticity induced by parental effects, and their interactions (Stearns 1992). We often know little about the origins of phenotypic variation, even though it critically shapes our inference of adaptation in natural populations.

Plasticity induced by parental effects is particularly strong from mothers due to their higher reproductive investment and, in viviparous species, the physically intimate relationship between mothers and their developing embryos (Lindholm et al. 2006; Wolf and Wade 2009). Such maternal effects are widespread in nature (Mousseau and Fox 1998) and can impact trait expression and evolution (Rossiter 1996; Wilson et al. 2004; Beckerman et al. 2006). Maternal effects can be adaptive if the expression of offspring traits is biased to match the environment experienced by the mother (e.g., Marshall and Uller 2007), or if mothers in good condition are

able to endow phenotypes that provide a competitive advantage to their offspring in any environment (e.g., Grafen 1988; Monaghan 2008; Van Allen et al. 2021). However, maternal effects can be maladaptive and produce mismatches between offspring phenotype and environment, as documented in some organisms responding to anthropogenic climate change (Jacob et al 2015; Marshall & Burgess 2015). Maladaptive maternal effects can also be related to stress, where physiological stress responses in mothers have unintended negative consequences on offspring (MacLeod et al. 2021). Regardless of whether maternal effects are adaptive, they are important biological phenomena that need careful attention and explicit accounting in evolutionary analyses due to the non-genetic effects on phenotypic expression.

Phenotypic variation in nature is common along environmental gradients, but it is often unclear whether it is caused by genetic differentiation among populations or plastic effects that arise from population-specific environmental exposure histories experienced by mothers or their offspring directly. For example, freshwater streams rich in toxic hydrogen sulfide (H_2S) in the Grijalva River basin of Tabasco and Chiapas, Mexico are extreme environments that are connected to adjacent nontoxic streams, and stark phenotypic gradients can be observed in fish occupying both habitats in as little as a few meters. Sulfide springs are complex ecosystems with several correlated sources of selection. H_2S is toxic because it disrupts aerobic ATP production (Reiffenstein et al. 1992; Cooper and Brown 2008), but habitats rich in H_2S also differ from nonsulfidic habitats in other aspects of water chemistry (dissolved oxygen concentrations, pH, and salinity), resource availability, and the presence of competitors and predators (Riesch et al. 2010a; Tobler et al. 2011a, 2015; Greenway et al. 2014). Populations of Atlantic mollies (*Poecilia mexicana*), a species of livebearing fish, have independently colonized and adapted to sulfidic streams across multiple river drainages, and previous studies have documented how

populations from different habitat types differ in morphology and locomotion (Tobler and Hastings 2011; Camarillo et al. 2020), behavior (Plath et al. 2007b; Lukas et al. 2021), physiology (Barts et al., 2018; Greenway et al., 2020; Tobler et al., 2011), and life history traits (Riesch et al. 2011a,b).

Phenotypic divergence between sulfidic and nonsulfidic mollies likely has a significant genetic component, because it coincides with strong genetic differentiation between ecotypes, even though there are no physical barriers separating populations in sulfidic and nonsulfidic habitats (Palacios et al., 2013; Plath et al., 2013; Riesch et al., 2016). However, trait variation between ecotypes likely also has an environmental component; while population differentiation persists in captive populations reared in common garden conditions in the lab (Tobler et al. 2016; Greenway et al. 2020), there is also evidence for plasticity caused by short-term exposure to different environmental conditions (Passow et al. 2017). In contrast, the impact of maternal effects on offspring trait expression remains to be investigated (Tobler et al. 2018). Hence, we tested how genetic differentiation, maternal effects, and their interactions shape phenotypic expression in *P. mexicana* ecotypes from sulfidic and nonsulfidic habitats. To induce maternal effects, we manipulated resource availability to pregnant mothers, because natural populations vary substantially in nutritional state (Tobler 2008b). Specifically, mollies in sulfidic habitats have significantly reduced body condition (both when inferred through length-weight regression and body fat content analysis; see Tobler et al. 2006; Tobler 2008b) because of constraints associated with resource acquisition. H_2S exacerbates hypoxia, and fish from sulfidic habitats have to trade off performing aquatic surface respiration to access better-oxygenated surface waters with benthic foraging (Tobler et al. 2009a). In addition, the resource environment

experienced by mothers has been shown to impact offspring trait expression in other species (Reznick et al. 1996a; Altmann and Alberts 2005; Boots and Roberts 2012).

To quantify population differentiation and maternal effects in populations of *P. mexicana*, we followed offspring families from birth to the onset of maturation and repeatedly quantified a host of complex phenotypic traits throughout ontogeny. Focal traits included brood size, size at birth, and age at maturity and ontogenetic variation in growth rates, burst swimming, exploratory behavior, and feeding accuracy. We chose these traits because they likely have implications for fitness, and we have prior knowledge for many of them from natural populations, providing us with a strong framework to make *a priori* predictions. Based on these data, our experiments sought to address four specific questions: 1) Is there evidence for differences in phenotypic traits between populations from sulfidic and nonsulfidic habitats that persist in fish reared in a common garden environment for multiple generations? Divergence in functional traits between populations regardless of maternal food treatments would indicate that trait differentiation is due to genetic variation between populations. 2) Is there evidence for maternal effects in response to resource availability? Significant differences in offspring traits across maternal food treatments, irrespective of population of origin, would suggest resource-induced maternal effects. 3) How do functional traits vary throughout ontogeny, and how do population differences and maternal effects interact with ontogeny? Age is a major determinant of expression for many traits (Hegyi et al. 2006; Zhang et al. 2015), but how population differences and maternal effects impact trait expression through ontogeny is less clear. In other poeciliids, maternal effects tend to be present at birth and decline with age (Lindholm et al. 2006). In contrast, population differences between sulfidic and nonsulfidic *P. mexicana* are stark in adults, suggesting differences may emerge early in life and even increase throughout ontogeny

(as in Riesch et al. 2011b). Accordingly, we predicted that age would impact most of the traits measured, that population differentiation would increase with age, and that maternal effects would only be observable at birth and diminish with age. 4) How do population differences interact with maternal effects? A difference in how each population responds to variation in maternal resource availability would indicate a genotype-by-environment interaction. The two populations are different in multivariate trait space in nature (Riesch et al. 2016b), but the magnitude of differences is less pronounced—though still significant—in captive-bred stocks (Tobler et al. 2008; Passow et al. 2015). We predicted that trait variation induced by maternal effects would be additive to trait variation associated with population differentiation, causing further divergence between nonsulfidic (typically in a "high-food" equivalent resource environment in nature) and sulfidic (typically in a "low-food" equivalent resource environment in nature) populations that resembles patterns of variation found in the wild. Alternatively, low-resource traits may be fixed in the sulfidic population because sulfidic individuals are constantly food stressed in nature (Wagner et al. 1997). If true, maternal effects may be weaker in the sulfidic population.

Methods

Experimental overview

For our experiments, we used two lab-reared populations of *Poecilia mexicana* originating from wild-caught relatives in the Tacotalpa River drainage. One population originated from a sulfide spring complex called El Azufre I (according to Plath et al., 2013; hereafter referred to as the sulfidic population or ecotype), and the other population was from a nonsulfidic stream connected to the mainstem of the Tacotalpa River, called Arroyo Bonita (according to Plath et

al., 2013; hereafter referred to as the nonsulfidic population or ecotype). Both populations were reared in 150-gallon stock tanks filled with a raised bed of aquatic plants, and filtered tap water. Tanks were fed *ad libitum* twice daily with commercial dry fish food (Purina), and 50 % of the water was exchanged weekly.

From each tank, 30 females were caught with a dipnet and isolated in a 5-gallon tank with an aerating filter and a large bundle of plastic mesh to help fry hide once they were born. Females were fed twice daily, with aquatic gel diet for omnivorous fish (Mazuri) in the morning and freshly hatched *Artemia* nauplii (Brine Shrimp Direct) in the evening. Females were randomly assigned to either a "high-food" diet (0.32 ml) or a "low-food" diet (0.08 ml), with equal numbers of females being assigned to each group from each population. Each tank was checked daily for fry hiding in the plastic mesh. Not all females gave birth during our experiment; 17 sulfidic females gave birth (11 from the high food treatment and 6 from the low food treatment) and 22 nonsulfidic females produced a brood (12 from the high food treatment and 10 from the low food treatment). To minimize density-dependent effects, we randomly selected 15 individuals (if available) from each family to remain in the experiments and returned the remaining individuals to their respective stock tank. The brood size, date of birth for each family, the length of time that the female was in the food treatment (days), and the mother's standard length (distance from the anterior tip of the snout to the caudal peduncle, mm) were recorded. We accidentally failed to record standard length for five females, so we used the respective population average standard length for those individuals. All fry received similar quantities of food; they were fed *ad libitum* twice daily with a mixture of decapsulated brine shrimp eggs (Brine Shrimp Direct) and dry food.

To address our questions, we then quantified offspring phenotypes throughout development for life history traits (size at birth, weekly growth rate, and age at maturity; measured once) and behavioral traits (burst swimming, exploratory behavior, and feeding accuracy; measured approximately weekly) that likely have implications for fitness (see Figure 3-1) and tested whether there were differences in these phenotypes between maternal food treatments (i.e., maternal effects) and/or between populations (i.e., sulfidic vs. nonsulfidic, indicating effects of evolved population differences that persist after several generations in captivity). All analyses were conducted in R 4.0.5 (R Core Team 2020).

Size at birth and growth rate

To measure size at birth, we photographed each family from above on the day of their birth with a Nikon D90 digital camera fitted with an AF-S Micro NIKKOR 105mm f/2.8 lens. A ruler was included in the background of each image. Images were imported to ImageJ (version 1.53; Schneider, Rasband, & Eliceiri, 2012) and calibrated by setting the scale. We measured the standard length (mm) of each offspring in the family. These measurements were averaged across all members of the family to obtain a single mean size at birth for each family. This measurement was completed weekly, and the measurement for each family was subtracted from that family's measurement from the week prior to obtain the average weekly growth rate in mm. Then, to account for potential allometric differences in growth rate, we converted the average weekly growth rate in mm to a proportional growth rate (in terms of body lengths) by dividing the average weekly growth rate by average body size the week prior.

Age at maturity

We measured the age at maturity for each family using morphological characteristics of sexual maturity. While juvenile livebearers of different sexes are difficult to distinguish, at the onset of sexual maturity, the anal fin is modified into an intromittent organ called a gonopodium in males, while it remains unmodified in females (Chambers 1987). We therefore measured age at maturity as the length of time (days) after birth that it took to observe the first visual signs of the development of the gonopodium.

Burst swimming

Most fish avoid predation with a highly conserved reflexive escape response that causes the head to move away from the stimulus, bending the body into a 'C' shape (Eaton et al. 1977). Then, a strong stroke of the caudal fin starts the movement away from the stimulus (Domencini and Blake 1997). This process (bending into a 'C' and propelling away from the stimulus) is known as a 'C-start' response and is frequently used as a metric of escape performance in fish (Ghalambor et al. 2004; Camarillo et al. 2020). To record this burst swimming behavior, we placed one individual from each family in a glass petri dish (9 cm in diameter, containing 2 cm of water) with opaque sides suspended above an angled mirror, providing a ventral view of each fish. After a 5 min acclimation, we quickly struck the surface of the water within a body length of the posterior of the fish with a blunted metal probe. We recorded the movement of the fish from below with a Sony NEX-FS700R camcorder at 60 frames per second with 1080p resolution. We converted the resulting .mts files into .mp4 files (to enhance compatibility with downstream applications) with FFmpeg (Tomar 2006).

We used DLTdv8 (Hedrick 2008) to digitize the 2-D location of the isthmus (i.e., the area on the ventral surface of the head where the gill operculae converge) of the fish in each frame. Digitized points were then used to calculate the maximum velocity (v_{max}, mm/s), maximum acceleration (a_{max}, mm/s^2), and net distance traveled (d_{net}, mm of displacement within 1/12th of a second after the C-start). To calculate v_{max}, we calculated the straight-line distance between each pair of successive digitized points (i.e., the hypotenuse), divided this distance by the inverse of the frame rate (60 fps), and found the maximum value between any two points. a_{max} was calculated in a similar way by subtracting the velocity value at each point from the velocity value at the point immediately preceding it and finding the maximum value. d_{net} was calculated by recording the 2-D position immediately after the fish ends the C-start with a single stroke of the caudal fin, and then recording the 2-D position 1/12th of a second later (5 video frames later) and calculating the straight-line distance between the two points. To reduce the dimensionality of this dataset, we then conducted a principal component analysis (PCA) using the prcomp function and a correlation matrix. There was one principal component with an eigenvalue greater than 1 (explained 85.6 % of the total variance), which was retained as the dependent variable. Positive scores along this principal component axis were associated with higher velocity, acceleration, and distance traveled (Table B-1). All mathematical operations were conducted using packages that came with the base distribution of R.

Exploratory behavior

We used an open field test to quantify the exploratory tendencies of fry. We filled a 9 cm Styrofoam cup with 3 cm of water and covered the arena with a sheet of glass coated in anti-fog solution (rain-x interior glass anti-fog liquid). We randomly selected one individual from each

family and placed it in the arena undisturbed for a 5 min acclimation period, after which time we recorded 5 min of video from above with a GoPro Hero 4 (1080p resolution, linear field of view, 30 fps).

We used idTracker (version 2.1, bundled with 64-bit Matlab Compiler Runtime 8.3; Pérez-Escudero et al. 2014) to track the 2-D location of the fish automatically for the entire 5 min recording. We set the number of individuals to be 1 and manually determined the intensity threshold (0.5-0.8) and minimum size (40-250 pixels) based on each video. We imported a still frame from each video into ImageJ to measure the centroid coordinates and arena radius to understand the distance from the center in each frame.

Using the 2-D coordinates in each frame, we calculated several metrics of motion that we used as proxies for exploratory behavior. We calculated distance traveled between each pair of successive points, the velocity, the acceleration, and the total cumulative distance traveled (d_{total}, mm), as above. We then calculated v_{avg} (average velocity, mm/s), v_{max} (maximum velocity, mm/s), and a_{max} by finding the mean and/or maximum of all velocity and acceleration values. Finally, we calculated the proportional average distance from the center of the arena (d_{center}, dimensionless) by calculating the distance from the fish's location to the centroid and dividing this value by the arena radius. Videos were excluded if the fish was completely still in all frames. To reduce the dimensionality and observe the correlation structure within this dataset, we ran a PCA on v_{avg}, v_{max}, a_{max}, d_{total}, and d_{center}, and used a correlation matrix. We retained scores along the first PC axis (explained 56.9 % of the total variance) as a composite exploratory behavior score. As shown in Table B-1, higher PC1 scores were associated with more exploratory behavior, with positive loadings for each input variable.

Feeding accuracy

To measure feeding accuracy, we starved the individuals of each tank for 24 h prior to the experiment and placed one randomly selected fry in a viewing tank for a 5 min acclimation period. The viewing tank was constructed by connecting two 10 cm × 10 cm sheets of glass ($1/8^{th}$ inch thickness) to either side of 1 cm diameter glass rods with waterproof silicone. The rear wall of the viewing tank was covered with a black sheet of plastic to enhance contrast between the background and the fish in the tank. The feeding solution consisted of 1 g of freshly hatched, live *Artemia* nauplii diluted into 100 ml of filtered tap water. After the 5-min acclimation, we added 0.08 ml of the feeding solution to the viewing tank and recorded 5 min of video with the camcorder. We analyzed the video frame-by-frame using BORIS (version 7.13; Friard & Gamba, 2016) and recorded the number of successful strikes (a feeding strike that ends in consumption of the food item) and unsuccessful strikes (a feeding strike that either misses the food item or results in regurgitation of the food item). We divided the number of successful strikes by the total number of strikes to get an overall proportional estimate of feeding accuracy, which was arcsine-square-root transformed prior to analysis.

Statistical analyses

There were many potential sources of variance in our experimental design; other than the effects of interest for our study (population, maternal food treatment, and their interaction), trait variation could also arise from differences in fry age, maternal standard length, the duration of food treatment, and brood size. We used a model selection approach to find the model that was best supported with our data from each experiment separately. For each phenotype, we created a global model that contained all possible effects. For phenotypes that were only measured once

for each family (size at birth, brood size, and age at maturity), we used a general linear model, and for phenotypes that were measured repeatedly through development (growth rate, burst swimming, exploratory behavior, feeding accuracy), we used a linear mixed model implemented with the lmer function from the lme4 package (version 1.1-26; Bates et al. 2015) that contained 'Family' as a random effect. We then used the dredge function from the MuMIn package (Bartón 2020) to create a model selection table based on the effects contained in the global model, with different models ranked and weighted based on AICc. Full model selection tables are available for each phenotype in Table B-2. To avoid overfitting, we limited the models to a maximum of four terms. We chose the top-supported model for each trait and analyzed it with either type-III analysis of variance (ANOVA) or Wald's chi-square tests (see Table 3-1). Quantification and visualization of effects was accomplished by calculating and plotting estimated marginal means for the effects of 'Population' and/or 'Food Treatment', depending on the best-supported model, using the Effect function in the effects package (Fox and Weisberg 2018, 2019a) (Figures 3-2 and 3-3). To aid in drawing inferences from our models, we generated 95 % confidence intervals (CI) for model coefficients in our top models using the confint function form the base R distribution. Effect sizes were calculated as partial eta-squared (η_p^2), which represents the proportion of variance explained by a particular variable after accounting for that explained by all other variables. We calculated η_p^2 with the etasq function (Fox et al. 2018) for general linear models or the eta_squared function (Ben-Shachar et al. 2020) for linear mixed models.

Multivariate analysis

Since selection ultimately acts on complex, multivariate phenotypes (Lande and Arnold 1983), we sought to understand how the traits measured for our analyses vary and covary to jointly

shape a single multivariate phenotype. To do so, we averaged each phenotype across all ages for each tank. We used this approach rather than including age as a covariate and analyzing raw phenotypic scores because of timing and logistical constraints that made it impossible to conduct each experiment on offspring that were the same age. We analyzed the averaged phenotypes with a PCA (correlation matrix) and used scores along the first two principal components—which had eigenvalues greater than 1—as dependent variables. We analyzed principal component scores along each axis separately because the axes were orthogonal. For each axis, we created a global model containing all possible effects and selected the best-supported model based on AICc, as above. Code to reproduce all analyses can be found on GitHub (https://github.com/michitobler/common-garden).

Results

We measured seven functional traits in offspring from sulfidic and nonsulfidic populations throughout ontogeny. Model selection tables for the analysis of each trait at birth and separately across all ages can be found in Table B-2, and the best-supported models are summarized in Table 3-1. For brevity, we will present results in the context of our hypotheses outlined in the Introduction, focusing on how all traits vary through ontogeny in terms of population differentiation, maternal effects, and their interaction, rather than presenting the results for each trait separately.

Is there evidence for population differentiation?

We found evidence for significant population differences in five of the seven traits measured: size at birth, age at maturity, burst swimming, exploratory behavior, and feeding accuracy (but

not brood size and growth rate). Age at maturity, burst swimming, and feeding accuracy were significantly different between ecotypes, but not between food treatments, whereas there were significant effects of 'Population' and 'Food Treatment' (but no interactions) on size at birth and exploratory behavior (Table 3-1). Sulfidic individuals were born 9.23 % larger [$F_{1,35}$ = 20.032; P < 0.001; estimated marginal mean for sulfidic fish (EMM_S) = 10.03 mm, estimated marginal mean for nonsulfidic fish (EMM_{NS}) = 9.18 mm; η_p^2 = 0.36] (Figure 3-2A), matured an average of 10.53 days later [$F_{1,26}$ = 4.420; P = 0.045; EMM_S = 54.22 days, EMM_{NS} = 43.68 days; , η_p^2 = 0.15], had significantly lower burst swimming performance [$\chi^2(1)$= 4.308; P = 0.038; CI: (-0.94, -0.03); η_p^2 = 0.03], were more exploratory [$\chi^2(1)$ = 7.513; P = 0.006; CI: (0.29, 1.70); η_p^2 = 0.28], and were 7.15 % less accurate during feeding [$\chi^2(1)$ = 4.645; P = 0.031; EMM_S = 0.72, EMM_{NS} = 0.80; η_p^2 = 0.04] (Figure 3-3).

Is there evidence for maternal effects in response to resource availability?

To test whether maternal effects induced by resource availability during pregnancy impact functional traits, we compared each phenotype between offspring born to mothers who experienced a high-food environment and mothers who experienced a low-food environment. Mothers in low-food treatments produced offspring that were 3.4 % larger at birth [$F_{1,35}$ = 4.354; P = 0.044; estimated marginal mean for fish in the low-food treatment (EMM_{low}) = 9.75 mm, estimated marginal mean for fish in the high-food treatment (EMM_{high}) = 9.43 mm; η_p^2 = 0.11] (Figure 3-2A) and were more exploratory [$\chi^2(1)$ = 4.245; P = 0.039; CI: (0.04, 1.42); η_p^2 = 0.18] (Figure 3-3E). For all remaining traits (brood size, age at maturity, growth rate, burst swimming, and feeding accuracy), 'Food Treatment' was not included in the best-supported model (see Table 3-1), suggesting that maternal effects did not affect the expression of those traits.

How do functional traits vary throughout ontogeny, and how do population differences and maternal effects interact with ontogeny?

To determine how traits changed through offspring development, we compared each phenotype across age groups. 'Age' was a significant predictor of growth rate and burst swimming, indicating that these two phenotypes changed through ontogeny, while the other phenotypes did not change with age. Across populations and food treatments, fry grew at a slower relative rate [$\chi^2(1) = 147.91$, $P < 0.001$, CI: (-0.010, -0.007), $\eta_p^2 = 0.50$] and performed better in burst swimming trials [$\chi^2(1) = 30.178$, $P < 0.001$, CI: (0.024, 0.049), $\eta_p^2 = 0.16$] as they got older (Table 3-1). Burst swimming scores were also lower in sulfidic individuals (see above), but the interaction term 'Population × Age' was not significant, demonstrating that both populations exhibited similar changes in burst swimming throughout ontogeny.

To ensure that signals of population differentiation and maternal effects occurring at birth were not obscured by measurements later in life, we also subset our dataset for each phenotype to obtain only the earliest data point for each family. As mentioned above, fry from the sulfidic ecotype and the low-food maternal treatment had significantly larger size at birth (Figure 3-2; Table 3-1). However, there were no population differences or maternal effects on brood size, which was higher in larger mothers [Mother SL; $F_{1,37} = 23.342$, $P < 0.001$, CI: (0.73, 1.79), $\eta_p^2 = 0.38$] and mothers who spent longer in the food treatment [Treatment Length; $F_{1,37} = 8.926$, $P = 0.005$, CI: (0.07, 0.40), $\eta_p^2 = 0.19$]. Growth rate at birth was significantly lower in sulfidic fry [$F_{1,37} = 5.025$, $P = 0.032$, CI: (-0.12, -0.006); $\eta_p^2 = 0.12$] (Table 3-1; Figure 3-3A), but there was no evidence for maternal effects on growth rate at birth ('Food Treatment' was not included in best-supported model). Note that the population difference in growth rate at birth disappeared as

fry developed (Figure 3-3B). All other traits that were measured throughout ontogeny (burst swimming, exploratory behavior, and feeding accuracy) did not exhibit significant population differences or maternal effects at birth. Interestingly, the maternal effect that we detected on exploratory behavior across all ages (see above) was not evident at birth [$F_{1,31} = 2.727$, $P = 0.108$, CI: (-0.29, 2.81)]. These results collectively demonstrate that—contrary to our hypothesis regarding ontogenetic variation of functional traits—population differentiation did not increase with age, and maternal effects were not always observable at birth, nor did they decline with age.

How do population differences interact with maternal effects?

We hypothesized that variation from maternal effects would be additive (i.e., the same direction) to population differences, but that there would be a significant interaction between population differences and maternal effects (i.e., different magnitudes of maternal effects in each population). Contrary to our prediction, no 'Population × Food Treatment' interactions were included in the best-supported model for any of the traits we measured (Table 3-1), suggesting a general lack of support for significant interactions between population differences and maternal effects. Therefore, maternal effects—if present at all—were similar between populations.

To further address our hypothesis, we asked whether population differences and maternal effects were, in fact, additive, and whether they explained a similar proportion of phenotypic variance when they acted in unison. We compared the signs (positive vs. negative coefficient estimates) and effect sizes (η_p^2) of population differences and maternal effects for the two traits with evidence of both effects simultaneously impacting trait expression—size at birth and exploratory behavior. For both traits, population differentiation and maternal effects produced trait shifts in the same direction, indicating that the effects were additive (Figures 3-2A and 3-

3E). Effect size estimates for 'Population' and 'Food Treatment' for both traits indicated that 'Population' had a larger effect on trait expression than 'Food Treatment' (η_p^2 = 0.36 vs. 0.11 for size at birth and 0.28 vs. 018 for exploratory behavior).

In addition to the univariate analyses of trait variation, we were also interested in understanding how the traits covaried with one another, and whether/how multivariate phenotypes were impacted by population differences and maternal effects. We averaged each phenotype across ages for each family and conducted a PCA with a correlation matrix. We then analyzed principal component scores along the first two principal component axes. The first principal component accounted for 28.5 % of the variance in multivariate phenotypes, and scores along the first principal component were positively correlated with brood size, growth rate, and feeding accuracy, and they were negatively correlated with size at birth, age at maturity, burst swimming, and exploratory behavior (see Table B-1). The second principal component explained 21.1 % of the variance and was positively correlated with burst swimming and size at birth and negatively correlated with brood size, growth rate, feeding accuracy, exploratory behavior, and age at maturity (Table B-1). After conducting model selection, the best-supported model for PC1 contained the main effects of 'Population', 'Food Treatment', and 'Treatment Length', and the best-supported model for PC2 contained the main effects of 'Population', 'Mother SL', and 'Treatment Length'.

Sulfidic offspring had significantly lower scores along PC1 [$F_{1,24}$ = 16.266, P < 0.001, CI: (-2.40, -0.78), η_p^2 = 0.40], as did offspring from low-food mothers [$F_{1,24}$ = 4.746, P = 0.039, CI: (-1.59, -0.04)], indicating population differences and maternal effects on PC1 scores. As for the univariate analyses, maternal effects were additive to population differences in PC1 scores (Figure 3-4), 'Population' had a larger effect than 'Food Treatment' (η_p^2 for 'Population' = 0.40;

η_p^2 for 'Food Treatment' = 0.17), and there was no evidence for significant 'Population × Food Treatment' interactions (term not included in best-supported model). We also found that sulfidic individuals exhibited marginally lower PC2 scores [$F_{1,24}$ = 3.321, P = 0.081, CI: (-1.45, 0.09), η_p^2 = 0.12] and detected no evidence for maternal effects (the 'Food Treatment' term was not included in the best-supported model) on PC2.

Discussion

Phenotypic variation can be shaped by numerous genetic and nongenetic factors, but the interplay of genes and environmental effects is rarely disentangled in natural systems even though it fundamentally impacts our inference of adaptation. Therefore, we examined how genetic variation (i.e., population differences), phenotypic plasticity mediated by maternal effects, and their interactions shape trait expression in two populations of *P. mexicana* that are exposed to strong divergent selection in nature. We found significant trait divergence between populations despite years of common garden captivity. In contrast, exposure of mothers to different food treatments impacted relatively few traits in their offspring and had weaker effects than the origin of populations. In addition, maternal effects tended to be additive to population differences and act in the same way across populations. Multivariate analyses supported the overall conclusion that populations are significantly different from one another despite generations of identical rearing conditions and that maternal effects were additive to population differences. We also found no evidence for interactions between populations and food treatments, suggesting that while populations have diverged in phenotypic traits, they have retained similar maternal influences on those same traits. Overall, we found that the stark phenotypic differences between ecotypes of *P. mexicana* that are evident in nature are largely a

consequence of genetic divergence and thus represent local adaptation. Maternal effects in response to resource availability, while present in some traits, appear to accentuate population differences, but do not cause them.

Trait variation across populations and maternal food treatments

There is a rich history integrating field-based studies quantifying trait variation in poeciliid fishes that occur along environmental gradients with lab-based studies that isolate causative factors (Endler 1980, 1995; Reznick and Bryga 1987; Reznick et al. 1990; Langerhans et al. 2004; Brown et al. 2005; Tobler et al. 2008; Plath et al. 2013; Riesch et al. 2014). While most studies focus on single or a few related traits (e.g., Endler and Houde 1995; Tobler 2008a; Riesch et al. 2011b), our study demonstrated that multiple complex trait differences quantified in common garden conditions, across food treatments and across ontogeny, closely mirror trait differences between populations in nature.

First, our study corroborates a genetic basis for population divergence in reproductive life history traits. Among-population variation in size at birth has been documented in many poeciliids exposed to divergent selection, including guppies and *Brachyrhaphis* species living in habitats with different predation regimes (Reznick et al. 1996b; Belk et al. 2020), *Poeciliopsis* inhabiting streams with variable stability (Weeks and Gaggiotti 1993), and least killifish with different levels of matrotrophy and superfetation (Schrader and Travis 2012). We documented significant population divergence in size at birth, with sulfidic mollies giving birth to larger offspring (see Figures 3-2A and 3-4). This finding is consistent with life history studies of sulfide spring populations in *P. mexicana* and other poeciliid species, suggesting that divergence in offspring size between sulfidic and nonsulfidic populations is a repeatable evolutionary

consequence of sulfide spring colonization that is largely driven by natural selection for larger offspring size (Riesch et al. 2014). Our multivariate results (Figure 3-4) also revealed that, like other poeciliids, the nonsulfidic ecotype of *P. mexicana* closely resembles the "opportunistic" life history strategy put forward by Winemiller (1992), which places a premium on earlier maturity and higher fecundity, at the expense of lower juvenile survivorship. The population differences we found in the sulfidic population may represent a shift towards a more "equilibrium" strategy, which is energetically costly to parents but should ultimately benefit offspring competitive ability and thus maximize parental fitness. In natural populations, maternal effects induced by variation in resource availability likely accentuate genetic differences in offspring size across sulfidic and nonsulfidic habitats. This pattern was also evident across a plethora of life history traits in guppies (Felmy et al. 2022). In the 10 life history traits for which they found evidence for significant divergence due to both genetic and maternal effects, divergence due to both effects occurred in the same direction in all of them (Felmy et al. 2022).

Second, we found patterns of population differentiation in age at maturity that match with previous work in livebearing fishes. Prior studies have demonstrated that guppies in streams with high predation pressure on adults mature significantly earlier than other populations (Reznick and Endler 1982). Likewise, several studies have documented genetic differences in age at maturity between populations of *P. mexicana*. For example, cave mollies mature significantly later than surface populations, but unlike our study, there was also significant plasticity in age at maturity due to resource availability (Riesch et al. 2016a). Another study of surface mollies that compared size at maturity in sulfidic and nonsulfidic populations found that sulfidic individuals reached maturity at a significantly smaller size than nonsulfidic individuals, likely because energy limitation caused by reduced resource availability in sulfidic habitats selects for reduced

growth rates and smaller size at maturity (Riesch et al. 2011a). Because poeciliid males exhibit determinate growth that ends at maturity (Stearns and Sage 1980), size at maturity should closely mirror age at maturity, assuming growth rates are the same like we found in our study. However, we found that sulfidic mollies matured at a significantly later age than nonsulfidic individuals, despite identical growth rates across ontogeny (Figure 3-3B). This discrepancy between our findings and those of previous studies suggests that plasticity in age at maturity may be quite strong and that experimental design and rearing conditions matter.

Third, we found significant behavioral differences between populations from sulfidic and nonsulfidic habitats, including exploratory behavior, feeding accuracy, and burst swimming. Previous work has shown that nonsulfidic mollies—and ones that were better fed—were bolder in their natural habitat, but behavioral differences also disappeared in the laboratory (Riesch et al. 2009). While we did not measure boldness *per se* in our experiments, exploratory behavior as measured in our experiment (i.e., activity level in a novel arena) is often characterized as part of a behavioral syndrome that correlates with boldness (Conrad et al. 2011). Unlike previous work, our study found that individuals from the sulfidic population and the low-food treatment were more exploratory (Table 3-1 and Figure 3-3E). These results support that maternal resource environment affects exploratory behavior, but also imply heritable differences in organismal performance. Such heritable population differences were also found for feeding accuracy and burst swimming but without any evidence for maternal effects. At least for burst swimming, the pattern of population differentiation in our experiment again mirrors findings from adult fish in natural habitats (Camarillo et al. 2020).

Our study found evidence for population differences and maternal effects in multiple functional traits that often mirrored findings across the family Poeciliidae from natural

populations living along gradients of predation, ecosystem stability, maternal provisioning, and permanent darkness, in addition to sulfide. In accordance with previous findings, we found that the sulfidic population was born larger, reached maturity at a later age, were more exploratory, and exhibited worse feeding accuracy and burst swimming, while maternal effects caused individuals in the low-food treatment to be born larger and be more exploratory. However, reconciling our findings with predictions from theory requires a detailed understanding of the traits that have diverged between populations and correlated characters that may have coevolved without being directly selected for.

Trait variation and adaptive function

Sulfide springs and proximate nonsulfidic habitats do not only differ in the presence and absence of H_2S, but they also vary in numerous abiotic and biotic factors that are often unaddressed in studies of adaptation. Sulfidic habitats tend to have lower dissolved oxygen concentrations, higher temperature, higher specific conductivity, and lower pH (Tobler et al., 2011), which in turn affect the biotic community (Greenway et al. 2014) and selection associated with resource exploitation, competition, predation, and parasites (Riesch et al. 2010a; Tobler et al. 2014, 2015). Adaptation in sulfide springs is therefore not solely in response to selection from sulfide but is instead in response to a multifarious selective regime that has caused stark multivariate phenotypic differentiation between ecotypes. As previously mentioned, colonization of sulfidic habitats has coincided with divergence between sulfidic and nonsulfidic ecotypes in numerous functional traits that mirror our experimental findings. Specifically, sulfidic mollies tend to be born larger and have smaller broods (Riesch et al. 2014), have poorer burst swimming performance (Camarillo et al., 2020), have larger heads and increased gill ventilation capacity

(Camarillo et al., 2020; Tobler et al., 2008), and suffer from poor body condition due to hypoxia-induced resource limitation (Tobler, 2008b).

Our analyses documented population differences in several functional traits that match with theoretical predictions. Selection from H_2S, lower predation, and lower resource availability should favor the evolution of larger size at birth (Reznick and Endler 1982; Riesch et al. 2014), while reductions in predation pressure and resource availability should also favor the evolution of enhanced exploratory tendencies (Teska et al. 1990; Kaun et al. 2007; Huang et al. 2012), as was found in our study. However, while the major selective agents in sulfide springs (i.e., sulfide, predation, and resource availability) may not directly impact the expression of the other traits for which we found population differences (burst swimming and feeding accuracy), variation in these traits may be the consequence of tradeoffs with other traits. Tradeoffs are common in organisms inhabiting contrasting environments because divergent selection can optimize traits for a given environment, sometimes at the expense of other traits due to constraints (Ghalambor et al. 2004; Kawecki and Ebert 2004; Garland et al. 2022). For example, reductions in burst swimming performance—as documented in our study—may arise as a simple consequence of selection for increased steady swimming efficiency, as different body shapes are associated with optimization of steady vs. unsteady swimming (Langerhans 2007, 2009; Tokić and Yue 2012; see Camarillo et al. 2020). Indeed, a trade-off between burst speed and sustained swimming performance has been documented in adult fish from the very same populations we studied here (Camarillo et al. 2020). Our study found a similar reduction in burst swimming performance in the sulfidic population, even among individuals that had not yet reached maturity, suggesting population differences in burst swimming arise quite early in development, though the presence of a tradeoff with endurance remains to be evaluated.

Our results also matched with *a priori* predictions regarding tradeoffs between respiration and feeding. Habitats rich in H_2S also experience rampant hypoxia, which has selected for the evolution of craniofacial traits (larger heads and jaws and longer gill filaments) that increase ventilation efficiency (Tobler et al. 2008). In addition to changes in morphology, sulfidic individuals also exhibit decreased foraging efficiency compared to nonsulfidic individuals (Tobler et al. 2009a), which was supported by our findings related to feeding accuracy (Figure 3-3F). The decreases in feeding efficiency noted in this and other studies may therefore be an inadvertent consequence of the craniofacial modifications that accompany colonization of sulfidic habitats, though this hypothesis remains to be tested.

Phenotypes in nature are the sum of genetic and environmental components, but surprisingly, we found no evidence for canalization or the evolution of plasticity by genotype × environment interactions, as originally predicted. Our work demonstrated that maternal effects tended to act in the same direction as population differences, furthering trait divergence between populations, and that the trait shifts induced by maternal effects were of a similar magnitude in both populations. For both traits in which we observed significant population differentiation and maternal effects (size at birth and exploratory behavior), the lowest phenotypic scores were found in nonsulfidic fry from high-food mothers, and the highest scores were found in sulfidic fry from low-food mothers (Figures 3-2A and 3-3E). Because sulfidic mollies exhibit reduced foraging efficiency and body condition as a consequence of hypoxia (Tobler, 2008b; Tobler et al., 2009), sulfidic habitats are naturally analogous to our low-food treatment, and nonsulfidic habitats are similar to the high-food treatment. If maternal effects enhance population differences in sulfide springs in nature, as seen in our experiments, then this may help explain why stronger divergence is typically observed in nature than in common garden reared fish (e.g., Tobler et al.

2008; Passow et al. 2015), highlighting the need for studies to address multiple concomitant selective agents simultaneously. Likewise, it is important to note that our experiments only captured a small aspect of plasticity, the portion controlled by the mother. While we attempted to standardize the environments experienced by each family, the offspring were ultimately born into different environments. Future work needs to address genetic effects, maternal effects, plasticity of exposed offspring, and their interactions to better understand the forces driving trait variation in nature. We also caution that the shifts found in our study are only from comparisons of a single population pair. Our understanding of the relative roles of genetic and/or plastic components of phenotypic variation would benefit greatly from studies that expand to other population pairs in *P. mexicana* and other species.

Conclusions

Natural selection can act on phenotypes to cause evolutionary change in populations within a single generation (Grant and Grant 1993). However, selection rarely impacts a single phenotype, since many traits have numerous correlated characters (Lande and Arnold 1983). Likewise, selection can act differently throughout an organism's life (Freda et al. 2019), so ontogenetic patterns of trait variation are crucial for our understanding of how traits evolve (Tinbergen 1963). It is therefore necessary to analyze variation in multiple fitness-relevant traits across ontogeny to truly understand adaptation. By doing so, our study provided strong evidence that the patterns of phenotypic variation found in wild populations of *P. mexicana* inhabiting divergent habitats are primarily driven by population differences indicative of local adaptation. Maternal effects, on the other hand, were present but generally weak, so they likely accentuate population differences, rather than causing them.

Acknowledgements

We would like to thank the members of the community in Tapijulapa, Tabasco for access to field sites and for their hospitality. We also would like to thank the Mexican government for issuing permits for past field collections, as well as our collaborators at the Universidad Juárez Autónoma de Tabasco and the Centro de Investigación E Innovación Para la Enseñanza y El Aprendizaje for support of research expeditions to obtain the specimens that gave rise to the populations used in this study. The assistance provided by J. Onnen and Q. La Fon was quintessential for processing our diverse datasets. Funding for this research was provided by grants to M.T. from the National Science Foundation (IOS-1931657) and the U.S. Army Research Office (W911NF-15-1-0175) and by a Graduate Assistance in Areas of National Need (GAANN) fellowship from the U.S. Department of Education to J.L.C.

Data Availability Statement

All datasets and R scripts used to analyze the data, with associated documentation, have been archived on GitHub (https://github.com/michitobler/common-garden).

Figures

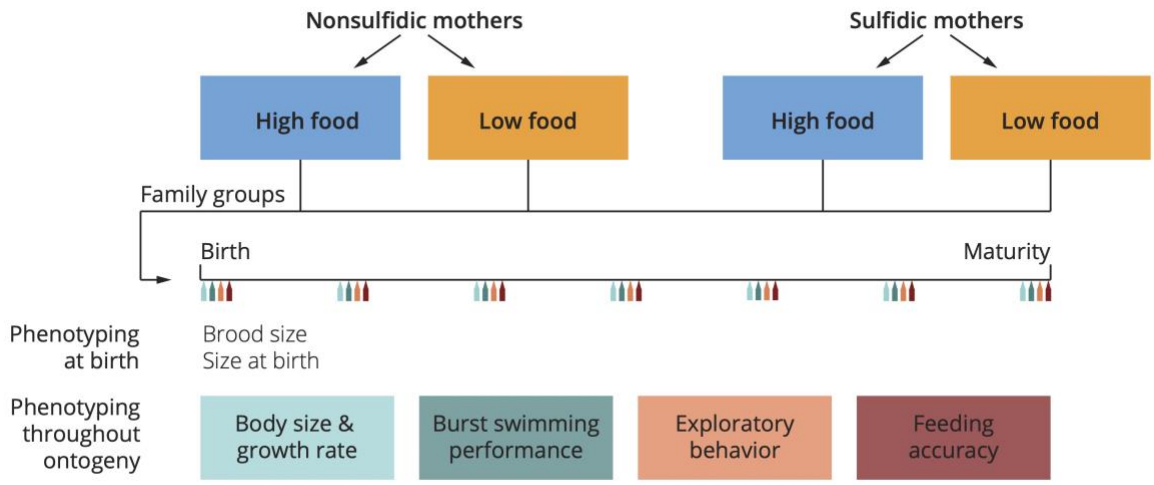

Figure 3-1: Diagram of our experimental design. We subjected pregnant sulfidic and nonsulfidic mothers to either a high- or a low-food treatment and measured seven traits in their offspring throughout their development. These traits include four life history traits (birth size, brood size, growth rate, and age at maturity) and three behavioral traits (burst swimming, exploratory behavior, and feeding accuracy).

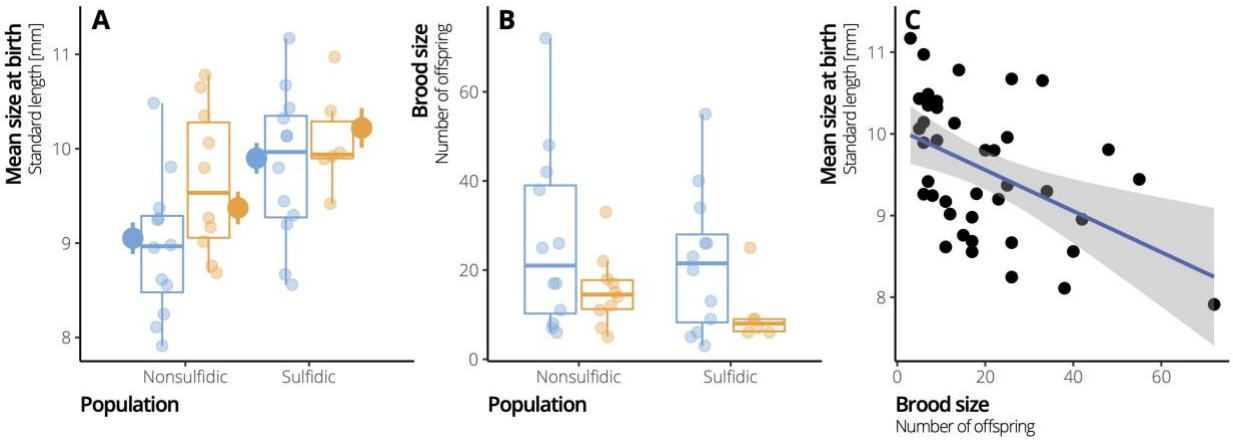

Figure 3-2: Plots of average standard length (A), brood size (B), and their covariance (C) in each tank. Data is portrayed as boxplots with raw data points overlaid. In both populations, the high-food treatment is shown in blue and the low-food treatment is shown in orange. Estimated marginal means for each group are portrayed as large dots.

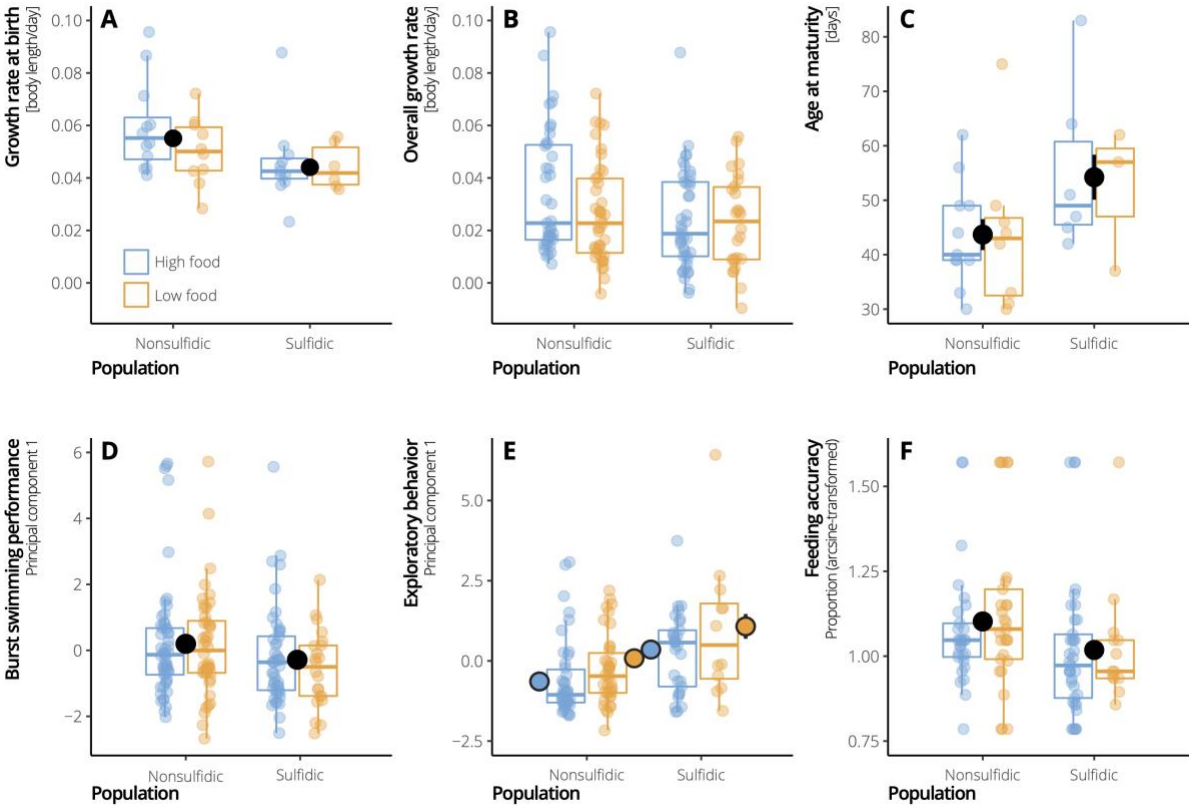

Figure 3-3: Data for each individual phenotype, including (A) growth rate at birth, (B) overall growth rate, (C) age at maturity, (D) burst swimming, (E) exploratory behavior, and (F) feeding accuracy. Data is portrayed as boxplots with raw data points overlaid. In both populations, the high-food treatment is shown in blue and the low-food treatment is shown in orange. When the best-supported model for a phenotype contained the terms 'Population' and/or 'Food Treatment', the estimated marginal means for those effects were visualized as large dots. These dots were colored black to show only 'Population' effects (i.e., if 'Food Treatment' was not included in the best-supported model), and they were colored according to the food treatment if both 'Population' and 'Food Treatment' were included in the best-supported model.

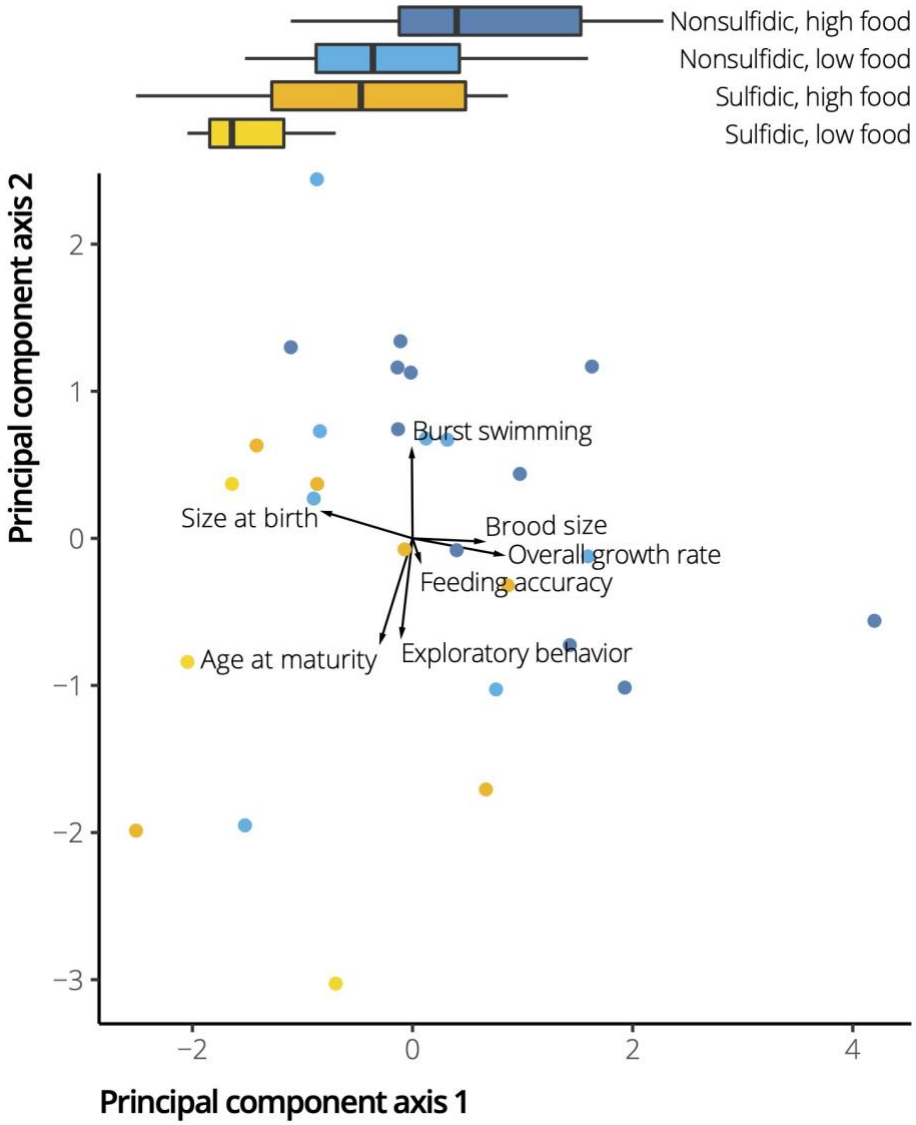

Figure 3-4: Plot of principal component scores representing linear combinations of all phenotypes in multivariate space. Scores were plotted along the first two principal components. Nonsulfidic individuals are shown in shades of blue, while sulfidic individuals are shown in shades of yellow. The high-food treatment within each population is shown in the darker shade (i.e., dark blue or dark yellow). Radiating from the origin are loading arrows that represent the correlation between the principal component scores and each input variable (shown as text at the end of each loading arrow). Loadings were calculated by multiplying the eigenvector for each input variable by the square root of the eigenvalue for that principal component axis. At the top of the figure are boxplots that show the distribution of principal component scores along the first principal component axis.

Tables

Table 3-1: Model terms in the best-supported model for each dependent variable. Empty cells represent terms that were not present in the best-supported model for that dependent variable. Type-III ANOVA or Wald's Chi-square tests were used to determine whether model terms were significantly different from zero. *F*-statistics or χ^2 statistics are presented for each included model term, with the degrees of freedom in a subscript. *P*-values are presented underneath each test statistic.

Dependent Variable	Brood Size	Treatment Length	Mother SL	Age	Population	Food Treatment
Size at Birth	$F_{1,35} = 20.876$ $P < 0.001$		$F_{1,35} = 16.593$ $P < 0.001$		$F_{1,35} = 20.032$ $P < 0.001$	$F_{1,35} = 4.354$ $P = 0.044$
Brood Size		$F_{1,37} = 8.926$ $P = 0.005$	$F_{1,37} = 23.342$ $P < 0.001$			
Age at Maturity					$F_{1,26} = 4.420$ $P = 0.045$	
Growth Rate at Birth					$F_{1,37} = 5.025$ $P = 0.031$	
Overall Growth Rate				$\chi^2_1 = 147.91$ $P < 0.001$		
Burst Swimming at Birth			$F_{1,39} = 5.025$ $P = 0.121$			
Overall Burst Swimming				$\chi^2_1 = 30.178$ $P < 0.001$	$\chi^2_1 = 4.308$ $P = 0.038$	
Exploratory Behavior at Birth						$F_{1,31} = 2.727$ $P = 0.108$
Overall Exploratory Behavior					$\chi^2_1 = 7.513$ $P = 0.006$	$\chi^2_1 = 4.245$ $P = 0.039$
Feeding Accuracy at Birth		$F_{1,38} = 0.775$ $P = 0.384$				
Overall Feeding Accuracy					$\chi^2_1 = 4.645$ $P = 0.031$	
Multivariate Trait Variation		$F_{1,24} = 15.817$ $P < 0.001$			$F_{1,24} = 14.984$ $P < 0.001$	$F_{1,24} = 4.746$ $P = 0.039$

Chapter 4 - Do sperm competition and cryptic female choice impact reproductive isolation in locally adapted populations of a livebearing fish?

John L. Coffin & Michael Tobler

Abstract

The mechanisms generating biodiversity are unknown in many natural systems. Understanding the speciation process requires in-depth knowledge of the evolution of reproductive isolation. Theoretically, reproductive isolation can evolve before mating (i.e., precopulatory), after mating but before fertilization (i.e., postcopulatory prezygotic), and/or after fertilization (i.e., postzygotic), but the relative roles of each process on limiting gene flow in nature are poorly understood. In Southern Mexico, populations of Atlantic mollies (*Poecilia mexicana*) have colonized and adapted to freshwater habitats rich in a respiratory toxin, hydrogen sulfide (H_2S). Across multiple river drainages, sulfidic ecotypes have diverged phenotypically and genetically from the ancestral nonsulfidic ecotypes, indicating that ecological speciation is ongoing. However, the barriers that limit gene flow between ecotypes remain poorly resolved. Previous work found that precopulatory isolation was strong but insufficient to limit gene flow to the observed levels, indicating that isolation may arise after copulation. We therefore measured male and female reproductive traits that could give rise to postcopulatory prezygotic isolation. We quantified sperm competitive ability in two contexts: i) in an H_2S gradient to determine whether sperm competitive ability was locally adapted; and ii) in ovarian fluid from a female in the same (i.e., native) or different (i.e., foreign) population to determine whether cryptic female choice

selectively biases sperm competitive ability. We expected that sperm competitive ability would be locally adapted to the sulfide regime that males were adapted to, and that ovarian fluid would bias sperm competitive ability in favor of native males, both of which could lead to the evolution of postcopulatory prezygotic isolation. Our experiments found little evidence for either mechanism contributing to reproductive isolation, but instead found that male body size was a consistent driver of reproductive success in both ecotypes of *P. mexicana*.

Keywords: ecological speciation, reproductive isolation, precopulatory prezygotic barriers, *Poecilia mexicana*, sperm competition, cryptic female choice, sperm storage, ejaculate size, sperm viability

Introduction

Understanding how reproductive isolation evolves is key for making sense of the speciation process (Turelli et al. 2001). Reproductive isolation can evolve due to many processes, including differential habitat or feeding preferences (Katakura et al. 1989; Nosil et al. 2002), different courtship behaviors (Houde 1997), gamete traits (Aguilar & Reyley, 2005; Gasparini, Marino, Boschetto, & Pilastro, 2010), and fitness reductions in immigrants (Tobler et al. 2009b; Plath et al. 2010) and/or hybrids (Scopece et al. 2008). Many studies on the role of sexual selection in speciation focus on traits prior to mating because these are often the easiest to observe and quantify (Birkhead and Brillard 2007). However sexual selection continues after copulation (Gasparini & Evans, 2018; Gasparini & Pilastro, 2011), and postcopulatory mechanisms can contribute to the formation of species (Birkhead and Møller 1998; Birkhead and Brillard 2007).

While copulation is necessary for sexual reproduction in internal fertilizers, it by no means guarantees successful fertilization (Eberhard 1985). After copulation, fertilization is only successful after intromission of genitalia, insemination, transport of seminal products to the correct storage or fertilization sites, as well as the final act of syngamy (Eberhard 1996). Accordingly, post-mating prezygotic isolation can be mediated by a variety of mechanisms. In polyandrous species, where females (i.e., egg producing individuals) mate with multiple males (i.e., sperm producing individuals), ejaculates from different males enter into competition for fertilization of the ova (Parker 1970). Two of the largest determinants of male fertilization success are ejaculate size (Parker 1990) and sperm velocity during travel through the female reproductive tract (Birkhead et al. 1999). Consequently, population differences in ejaculate or sperm traits can impact fertilization success in among-population matings and contribute to reproductive isolation. Females, however, are not passive participants in these mating games. In

many species, females can influence the outcome of copulation without directly choosing a mate (Eberhard 1996). This can manifest as an active choice by the female, mediated by perception of a male's quality (Pilastro et al. 2004), or through internal changes in the female's physiology (Aguilar and Reyley 2005; Fitzpatrick et al. 2020). Such cryptic female choice gives females control over which males are most likely to fertilize their ova by directly impacting sperm competitive traits (Gasparini et al., 2020), including sperm velocity (Cramer, Alund, McFarlane, Johnsen, & Qvarnström, 2016; Cramer, Stensrud, et al., 2016; Devigili et al., 2018; Fitzpatrick et al., 2020; Gasparini & Pilastro, 2011; Liberti, Baer, & Boomsma, 2016), trajectory (Liberti et al. 2016), chemical sensing ability (Fitzpatrick et al. 2020), viability (den Boer et al. 2009), and longevity (Gasparini & Evans, 2013). Hence, cryptic female choice can act as a reproductive barrier if female traits favor sperm of males from the same population.

Ecological speciation is frequently studied among populations that occur along environmental gradients and adapt to different ecological conditions, which has facilitated the identification of traits that influence reproductive isolation and ultimately rates of gene flow across incipient species (Nosil, 2012; Nosil, Vines, & Funk, 2005; Rundle & Nosil, 2005; Schluter, 2001), although postcopulatory prezygotic barriers have received comparatively little attention (Birkhead and Brillard 2007). For example, populations of Atlantic mollies (*Poecilia mexicana*), a species of livebearing fish, have colonized freshwater habitats rich in toxic hydrogen sulfide (H_2S). Exposure to H_2S is lethal to most animals even at micromolar concentrations (Grieshaber and Völkel 1998), because it inhibits aerobic ATP production by disrupting cytochrome c oxidase (COX) function in the mitochondrial oxidative phosphorylation pathway (OXPHOS; Cooper & Brown, 2008; Evans, 1967). *Poecilia mexicana* have colonized sulfide springs in multiple river drainages in the Mexican states of Tabasco and Chiapas,

adapting to the toxic effects of H_2S through convergent molecular modifications of OXPHOS, differential expression of sulfide detoxification genes, or other mechanisms (Kelley et al. 2016; Tobler et al. 2018; Greenway et al. 2020). Despite a lack of physical barriers separating populations in adjacent sulfidic and nonsulfidic habitats, gene flow between the locally adapted ecotypes is minimal, suggesting ongoing ecological speciation (Plath et al. 2013). Strong natural selection clearly favors H_2S-adapted individuals in the toxic habitats (i.e., strong selection against migrants: Plath et al. 2013). However, one of the unresolved questions in this system is why there is such limited gene flow from H_2S-adapted populations to populations in non-sulfidic habitats (Tobler et al. 2018). Previous work has documented reductions in gene flow due to prezygotic isolation caused by natural selection against migrants and assortative mating in some nonsulfidic ecotypes (Plath et al. 2013; Greenway et al. 2016). However, Plath et al. (2013) also found that the observed levels of gene flow between sulfidic and nonsulfidic ecotypes did not correlate with the prezygotic isolation observed in their experiments, suggesting that postmating barriers are likely in place and contribute to the low rates of gene flow found in nature.

Like other members of the family Poeciliidae, Atlantic mollies are polyandrous (Magurran 2011), have a coercive mating system (Plath et al. 2007a), and can store viable sperm in folds of the ovarian epithelium (i.e., micropockets) and the gonoduct for several weeks and even months at a time (Philippi 1908; Evans et al. 2011; Torres-Martínez et al. 2017). All these attributes facilitate the temporal overlap of ejaculates from multiple males in the same female reproductive tract, potentiating sperm competition and cryptic female choice. Sperm number, velocity, and longevity should be critical for sperm competition success in this species (Gasparini et al. 2010; Boschetto et al. 2011), and there is strong evidence in related species (*Poecilia reticulata*) that cryptic female choice mediated by ovarian fluid composition can lead

to variation in sperm competitive ability (Gasparini and Pilastro 2011). However, it remains to be investigated how traits associated with sperm competition and cryptic female choice vary within and among populations of *P. mexicana*, and whether such traits might contribute to reproductive isolation among populations adapting to different environmental conditions.

Here, we investigated whether postcopulatory prezygotic barriers have evolved between ecotypes of *P. mexicana*. We assessed sperm competitive traits in males from different populations under different environmental conditions and tested for cryptic female choice by measuring sperm performance in ovarian fluid from females belonging to the same or different population as the male. We addressed four specific predictions with our experiments: 1) We expected males from nonsulfidic populations would have larger ejaculates, because wild-caught males from nonsulfidic populations have larger testes and ejaculates (Schlupp et al. 2006; Franssen et al. 2008). 2) We predicted that sperm produced by males from sulfidic habitats would swim faster than sperm produced by males from nonsulfidic habitats in the presence of H_2S. Sperm performance is critically dependent on mitochondrial function (Moraes and Meyers 2018), which is targeted by H_2S (Cooper and Brown 2008), and *P. mexicana* from sulfidic habitats have modified mitochondria that facilitate the maintenance of aerobic ATP production in the presence of H_2S (Pfenninger et al. 2014; Greenway et al. 2020). But we also expected a cost of H_2S adaptation, where sperm from sulfidic males are inferior competitors under nonsulfidic conditions. 3) We predicted that sperm from nonsulfidic males would have higher longevity due to energetic costs associated with H_2S adaptation (Riesch et al. 2010b). 4) We predicted that sperm would swim fastest in ovarian fluid derived from females of the same compared to the opposite population, indicating that cryptic female choice can bias sperm performance. There is evidence for correlated changes in male and female genitalia between diverging populations of

P. mexicana (Greenway et al. 2019), and we hypothesized that there are correlated changes in the characteristics of sperm and ovarian fluid that lead to context-dependent sperm performance.

Methods

Collection and maintenance of focal populations

Our experiments were conducted on two sulfidic/nonsulfidic population pairs of *P. mexicana*, one of which was from the Tacotalpa River drainage, and the other from the Puyacatengo River drainage. Tacotalpa males were used for all experiments, while Puyacatengo males were only used for sperm competition analyses (see below). We used males from both drainages because previous work on physiological adaptation of both populations to H_2S revealed that sulfidic individuals from the Tacotalpa drainage exhibited decreased COX activity at high sulfide concentrations, while sulfidic individuals from the Puyacatengo drainage maintained high COX activity at all sulfide concentrations (Pfenninger et al. 2014; Greenway et al. 2020). We therefore wanted to determine whether the mechanisms by which different lineages tolerated sulfide exposure (i.e., whether lineages had sulfide tolerant vs. intolerant COX enzymes) had functional ramifications for their sperm competitive ability in sulfidic habitats.

All individuals used were lab-reared progeny of wild ancestors that had been living in captivity for at least three generations in 150-gallon stock tanks (Rubbermaid). All tanks were fed *ad libitum* twice daily with commercial dry food (Purina) and received a 50 % water change weekly. From each stock tank, 15 males were removed and isolated in 5-gallon aquaria for one week prior to the experiment with feeding conditions as above. Once the isolation was complete, we subjected the males randomly to either a high- or low-food treatment for at least two weeks. Food availability is significantly different between populations in sulfidic and nonsulfidic

habitats in nature (Tobler et al. 2009a), so we used variation in food availability to reflect natural conditions and potentially induce plastic shifts in individual resource allocation (Riesch et al. 2016a). To make feeding solutions, we thawed 25 g of frozen *Cyclops* copepods (Jehmco) in 250 ml of filtered tap water. This feeding solution was kept refrigerated and was replaced weekly. The high-food treatment received 4 ml of this solution once daily, while the low-food treatment received 1 ml once daily.

Sperm extraction

We followed protocols used in other poeciliids to extract sperm (Franssen et al. 2008; Boschetto et al. 2011; Gasparini and Evans 2013). Poeciliid sperm are organized in discrete sperm bundles, which consist of approximately 22,000 immotile spermatozoa (at least in related guppies: Cattelan et al. 2018) that are held together by a mixture of complex glycoconjugates (Greven 2005) until they are activated. Prior to anesthetization, we measured the standard length of each male. Males (N = 12–14 for each treatment group) were then placed in an anesthetic water bath containing 0.16 g/l MS-222 (Pentair) buffered with 0.16 g/l sodium bicarbonate. Once males lost equilibrium, they were removed from the anesthetic, patted dry with a paper towel, and placed on a plastic weigh boat. We added 100 ml of saline solution (0.9 % NaCl) near the base of the gonopodium (the modified anal fin in male poeciliids that serves as an intromittent organ). The weigh boat was placed under a dissecting microscope (Nikon SMZ1000; Tokyo, Japan) fitted with a digital camera (Canon EOS Rebel T5i; Tokyo, Japan) at low magnification. The gonopodium was swung forward with the tip of a pipet, and gentle pressure was applied to the abdomen with the thumb, resulting in the ejaculation of sperm bundles. Pressure was applied three times to ensure that all sperm bundles were ejaculated. Any sperm bundles that were

attached to the gonopodium or the male's skin were removed by gently pipetting additional saline and applying suction to the area. All sperm bundles and saline were then transferred to a 1.5 ml microcentrifuge tube, where they remained quiescent until activation.

Quantification of ejaculate size

The full ejaculate in saline was placed on a weigh boat, and an image was taken of each male's ejaculate under the dissecting microscope at a magnification of 0.8×. Raw photos were uploaded to ImageJ (version 1.49; Schneider et al. 2012). Thresholds were applied to each image to separate sperm bundles from the background, and we then used the Analyze Particles feature of ImageJ to measure the total area of sperm bundles by counting the number of pixels that passed the thresholding criteria in each image (as in Rudin-Bitterli et al. 2020). We used area as a proxy for absolute ejaculate size because the sperm bundles often stuck to each other, making them difficult to accurately count. This measure of absolute ejaculate size was divided by $\log_{10}$-transformed male body size (SL) to understand variation in relative ejaculate size (per unit of body size).

Measurement of sperm velocity

Preparation of hydrogen sulfide

Hydrogen sulfide solution was prepared daily following Barts (2020). In brief, a 1 mM stock solution was prepared under anoxic conditions in a glovebox filled with gaseous nitrogen. After bubbling 2 l of deionized water under anoxic conditions for 10 min to deoxygenate the solution (Butler et al. 1994), we removed any additional oxygen from the solution by dissolving 10 g of sodium sulfite, which reacts with oxygen to form sodium sulfate. We then added 0.4804 g of

sodium sulfide nonahydrate, which dissociates to form various sulfides (H_2S, HS^-, and S^{2-}) depending on the pH. Realized total sulfide concentrations (H_2S, HS^-, and S^{2-}) were measured with a methylene blue assay on a Hach DR1900 Portable Spectrophotometer.

Sample preparation

From each ejaculate, three sample tubes were prepared, one for each of three sulfide concentrations (0 mM, 50 mM, and 100mM). Each tube received 6 μl of activation solution (150 mM KCl and 2 mg/ml bovine serum albumin; Billard and Cosson 1992; Spagopoulou et al. 2020). The 0 mM tube received 4 μl of saline containing approximately 50 sperm bundles and no H_2S. The 50 mM tube received 3.5 μl of saline containing 50 sperm bundles and 0.5 μl of H_2S. The 100 mM tube received 3.0 μl of saline containing 50 sperm bundles and 1.0 μl of H_2S. All three tubes contained the same total volume (10 μl) and only differed in the concentration of H_2S.

Recording sperm motion

Immediately after the sample tubes were prepared, the mixture was vortexed at 2,000 rpm for 10 s to help release sperm from their bundles (Gasparini et al. 2010; Spagopoulou et al. 2020). Then, 3 μl of the activated solution was pipetted into the well of a Leja slide (20 μm depth), which was placed under a light microscope (Olympus CH-2; Tokyo, Japan) fitted with a digital camera at 10 × magnification. Since sperm motion is highly correlated with time after activation (Purchase and Earle 2012), we began recording 20 s after activation to provide enough time to find sufficient motile sperm in the field of view and standardize the time after activation at which

we began recording. The camera recorded unencoded raw footage at 1,280 pixel × 720 pixel resolution and 60 fps for 15 s for each sample.

Measuring sperm swimming performance metrics

Video footage was converted to .avi format using ffmpeg (Tomar 2006). The .avi files were imported into ImageJ, and motion parameters were estimated for two half-second intervals, one from 0.5 s to 1 s and the other from 10.5 s to 11 s, using the CASA plugin (Wilson-Leedy and Ingermann 2007). We generally followed the approach of Sanches et al. (2013) to measure sperm swimming performance, with a few adjustments to the parameters used for CASA (see Table C-1). The output from this plugin provided the total number of cells observed, the percent motility (MOT; the percentage of the observed cells that were moving) and averaged values of straight-line velocity for all motile sperm cells (VSL; mm/s). MOT values were arcsine-square-root transformed prior to analysis.

Assessment of sperm longevity

To test for variation in sperm longevity, we used a LIVE/DEAD Sperm Viability Kit (L7011; Thermo Fisher Scientific, Waltham, MA, USA) to distinguish between live (stained green with SYBR-14, a membrane permeant nucleic acid stain) and dead (stained red with propidium iodide, which is membrane impermeant) sperm cells. This kit has been used to measure sperm viability in other poeciliids (Gasparini et al. 2010; Gasparini and Evans 2013; Cardozo et al. 2020). After sperm extraction, we left the ejaculate sample in its microcentrifuge tube for 1 min so the sperm bundles would settle on the bottom of the tube. Then, we placed 2 μl of the sperm bundles into 28 μl of activation solution and vortexed the entire sample at 2,000 rpm for 1 min.

We left the vortexed sample to rest for 2 min to enable sperm to leave their bundles. We then followed the manufacturer's instructions and took a 6 μl aliquot of the activated sample and immediately stained it with 1 μl of SYBR-14 and 5 μl of propidium iodide, pipetting the entire sample five times to ensure adequate mixing. After staining, we incubated the samples in a water bath at 36 °C for 5 min. Each incubated sample was then mixed by pipetting five times, and 3 μl of the sample was added to each of two replicate wells on a 12-well multitest slide (MP Biomedicals, Irvine, CA, USA) and covered with a coverslip (as in Gasparini et al. 2010). The activated sample was left at room temperature for four hours (Cardozo et al. 2020), at which time a second 6 μl aliquot was taken, and the staining procedure was repeated.

Each replicate well was imaged in two distinct fields of view on a Zeiss AxioImager Z1 epifluorescence microscope with a 20× objective lens and Zeiss ZEN 3.0 software, resulting in four images for each sample. For each replicate, the image with the largest number of cells was chosen, and results were averaged between replicates to obtain a single count of live and dead cells for each slide. Microscopy imaging was conducted following the methods outlined in Chen et al. (2020). Live cells were visualized in a green channel using a GFP filter and dead cells were visualized in a red channel with an mCherry filter. Using the 'Z-stack' option in ZEN 3.0, we took 10 images in each channel that were 1.5 μm apart, providing us with 15 μm of focal resolution.

Raw image stacks were imported as .czi files into ImageJ, and the red and green channels were split into two separate stacks of images. Each stack was analyzed as a separate Z Project in ImageJ with the projection type set to 'Max intensity'. The type of each image was set to '8-bit' and the contrast was increased by 20 units. We then performed thresholding on each image separately and counted the number and measured the area of all identified cells.

Measuring the effect of cryptic female choice on sperm performance

Males and females from both populations in the Tacotalpa drainage were isolated in 5-gallon aquaria for 2 weeks prior to the experiment. Standard length was measured for all individuals. Males (N = 10) and females (N = 10) from each population (total N = 40 fish) were anesthetized, and sperm was extracted from males as above. To test for cryptic female choice, we then measured whether sperm swam faster in ovarian fluid of females from the same or the opposite population, which has been shown to bias sperm performance and fertilization success in other poeciliids (Gasparini et al. 2020).

To extract ovarian fluid, we placed anesthetized females on their back in a slit in a wet sponge under a dissecting microscope with their gonopore easily visible (Gasparini and Pilastro 2011). We gently injected and removed 2.5 μl of saline solution from the female's gonopore three times (total ovarian fluid sample volume of approximately 7.5 μl) and placed this sample in a 1.5 ml microcentrifuge tube. We used a two-by-two block design (N = 10 replicates) to determine whether sperm competition success is mediated by ovarian fluid origin (Devigili et al. 2018). We measured swimming performance of sperm from sulfidic and nonsulfidic males in ovarian fluid of a female from the same and the opposite population. Each ejaculate was thus split into two aliquots of approximately 50 sperm bundles each, and ovarian fluid samples were also split in half. Each ejaculate aliquot was then tested in ovarian fluid of a female from the same population and in ovarian fluid from the opposite population (3 μl of activation solution and 2 μl of ovarian fluid). Once the ejaculate was added to each solution, half of the volume was pipetted 25 times to mix the contents and break up the sperm bundles. We then added 3 μl of this activated sample to a Leja slide and recorded sperm motion with a digital camera as above.

Videos were again analyzed with ImageJ's CASA plugin, using the same input parameters, to obtain data on sperm performance (MOT and VSL).

Statistical analyses

We used a model selection approach to analyze our data and accounted for uncertainty in our model selection choice with model averaging (Burnham and Anderson 2002; Symonds and Moussalli 2011). Absolute and relative ejaculate size were analyzed with general linear models using the lm function in the stats R package, while sperm swimming performance, longevity, and cryptic female choice were analyzed with linear mixed effects models using the lmer function in the lme4 package (Bates et al. 2015). For each of these dependent variables, we constructed a global model and included all possible effects, including the ones of interest for this study (i.e., population, food treatment, sulfide concentration, and timepoint, if applicable) and covariates (i.e., treatment length and body size). All possible two-way interaction terms were also included in the global model. The maximum model size was set to four independent variables to avoid overfitting. Using the dredge function from the MuMIn R package (Bartón 2020), we generated a model selection table with all possible models weighted by AICc. Full model selection tables for each dependent variable can be found in Table C-2.

We averaged all models with a cumulative model weight (i.e., the probability that the chosen model accurately models the true, data-generating mechanism) of 0.95 using the model.avg function in the MuMIn package, creating a 95 % model confidence set (Hansen et al. 2011). This approach calculates parameter estimates for each component model and averages them across all models for which a particular term appears, based on the weight of that model. We calculated *Z*-scores by dividing the weighted, averaged estimates of each model term by the

standard errors for each estimate, and calculated *P*-values based on these *Z*-scores to draw inferences. We also calculated 95 % confidence intervals (CI) to assess whether the parameter estimates in the averaged model were significantly different from zero. The averaged estimates, *Z*-scores, *P*-values, and CIs for all terms for the 95 % confidence model set for each dependent variable can be found in Table 4-1.

Results

Do nonsulfidic mollies have larger ejaculates?

Absolute ejaculate size was measured as the total area of sperm bundles in standardized images of ejaculates. We averaged the top 10 of 19 possible component models for absolute ejaculate size, which collectively analyzed variation in absolute ejaculate size due to the effects of 'Ecotype' (sulfidic vs. nonsulfidic), 'Standard Length', 'Treatment Length', and 'Food Treatment'. As expected, we found that larger males produced larger ejaculates in terms of absolute size in pixels ($Z = 2.635$; $P = 0.008$; CI: 8416, 57258), but we found no evidence for differences in absolute ejaculate size between ecotypes, food treatments, or as a function of treatment length ($Z < 1.157$, $P > 0.247$ for each effect). In terms of relative ejaculate size, which was averaged among the top 13 models for the effects of 'Ecotype', '$\log_{10}$(Standard Length)', 'Treatment Length', 'Food Treatment', and 'Food Treatment × Ecotype', we surprisingly found that larger males also had larger relative ejaculate size per unit of body size (effect of $\log_{10}$(Standard Length): $Z = 2.265$; $P = 0.024$; CI: 185975, 2578017; see Figure 4-1 and Table 4-1). This indicates that larger males invest relatively more in reproduction than smaller males. However, as shown in Figure C-1, we did not detect differences in relative ejaculate size among ecotypes ($Z = 0.593$; $P = 0.553$; CI: -132578, 247493) or food treatments ($Z = 1.086$; $P = 0.277$;

CI: -84157, 293509). We likewise did not detect an effect of either treatment length or a food treatment × ecotype interaction ($Z < 1.070$, $P > 0.284$ for both).

Is there evidence for local adaptation in sperm swimming performance?

Sperm swimming performance was measured with two dependent variables separately: MOT, which provides information on the relative number of sperm that can enter into competition with other ejaculates, and VSL, which provides information on sperm velocity and path straightness. The model set we used to analyze MOT contained 5 of the 71 total component models and calculated averaged estimates for the effects of 'Ecotype', 'Food Treatment', 'Timepoint', and 'Standard Length' (Table 4-1). 'Individual ID' and 'Drainage' were set as random factors. We found that smaller males produced more motile ejaculates ($Z = 2.245$; $P = 0.025$; CI: -0.012, -0.0008; Figure 4-2A) and that motility decreased with time-after-activation ($Z = 1.834$; $P = 0.067$; CI: -0.049, 0.002). We detected no differences in MOT between ecotypes ($Z = 0.633$; $P = 0.507$; CI: -0.056, 0.114) or food treatments ($Z = 0.183$; $P = 0.855$; CI: -0.099, 0.082; Figure C-2A).

Our model set for VSL contained the top 8 models and analyzed variation associated with the fixed effects of 'Ecotype', 'Food Treatment', 'Sulfide', 'Timepoint', 'Treatment Length', 'Food Treatment × Sulfide', and 'Ecotype × Sulfide', and the random effects of 'Individual ID' and 'Drainage' (Table 4-1). VSL decreased marginally with time-after-activation ($Z = 1.888$; $P = 0.059$; CI: -5.165, 0.096) and strongly decreased as the treatment length (i.e., the length of male isolation before measurement) increased ($Z = 2.728$; $P = 0.006$; CI: -0.372, -0.061). The effect of sulfide concentration also depended on the food treatment (effect of 'Food Treatment × Sulfide': $Z = 2.566$; $P = 0.010$; CI: -15.488, -2.073). Specifically, there was little variation associated with

sulfide concentration in the high-food treatment, while there was a significant decrease in VSL as sulfide concentration increased in the low-food treatment (Figure 4-2B). We found no evidence for significant differences in VSL arising from the main effects of food treatment (Z = 1.344; P = 0.179; CI: -2.468, 13.235), the sulfide concentration ($Z \leq 0.652$; $P \geq 0.515$), or ecotype (Z = 0.281; P = 0.779; CI: -6.149, 8.208), indicating that sperm performance does not exhibit patterns of local adaptation (Figure C-2B).

Do nonsulfidic sperm have higher longevity?

Our model selection approach for sperm longevity yielded a single best-supported model that had an Akaike weight of 0.93, making it the only model in our confidence set. This best-supported model contained only the fixed effect of 'Timepoint' and the random effect of 'Individual ID', and we analyzed it with a type-III Wald's Chi-square test using the Anova function implemented in the car package (Fox and Weisberg 2019). We detected a significant decrease in longevity through time (χ^2_1= 63.477; P < 0.001; CI: -0.335, -0.203; Figure C-3). We interpreted the lack of other model terms in our best-supported model as evidence for a lack of significant effects of ecotype, food treatment, body size, and treatment length, since their inclusion in the model did not improve model fit.

Is there evidence for cryptic female choice enhancing sperm performance in native ovarian fluid?

As for our sperm competition analyses above, we analyzed the effect of ovarian fluid origin on two aspects of sperm swimming performance, MOT and VSL. Our model set for MOT contained the top three models (out of 43 total possible models) and analyzed variation in arcsine-square-

root transformed MOT related to the fixed effects of 'Male Ecotype' and 'Timepoint' and the random effects of 'Male ID' and 'Female ID'. We found no evidence for ecotypic differences in MOT (Z = 0.665; P = 0.506; CI: -0.196, 0.097; Figure C-4A), nor for differences in MOT across timepoints (Z = 0.825; P = 0.410; CI: -0.075, 0.030). In fact, the best-supported model (Akaike weight of 0.84) was the null model, indicating that there was little evidence for any effects in our mixed model for MOT (Table 4-1).

VSL, on the other hand, was analyzed with a set of the top 17 models, which collectively contained the fixed effects of 'OF origin', 'Male Standard Length', 'Female Standard Length', 'Male Ecotype', 'Timepoint', 'OF origin × Male Ecotype', 'OF origin × Timepoint', and 'Male Ecotype × Timepoint', and the random effects of 'Male ID' and 'Female ID'. We found that sperm from larger males swam significantly faster, regardless of ecotype and ovarian fluid context (Z = 3.195; P = 0.001; CI: 0.930, 3.882; Figure 4-3). Contrary to our predictions, we did not find evidence for a significant interaction between male ecotype and OF origin (Z = 1.505; P = 0.132; CI: -42.029, 5.516; Figure C-4B) or other effects on VSL (Table 4-1).

Discussion

Ecological speciation generates biodiversity as a consequence of reproductive isolation arising from adaptation to contrasting ecological conditions (Nosil 2012). In populations of *P. mexicana* that inhabit different habitats (some with toxic H_2S and some without), ecological speciation is incipient, but reproductive isolation is strong, with low rates of gene flow across very small spatial scales (Tobler et al. 2018). The relative strength of isolating barriers in this system are still subject to investigation. Precopulatory isolation—in the form of natural selection against migrants from nonsulfidic to sulfidic populations and sexual selection against migrants from

sulfidic to nonsulfidic populations—is strong but not sufficient to explain the low rates of gene flow in this system (Plath et al. 2013). Hence, we devised an experiment to test for the presence and strength of isolation that occurs after copulation but before fertilization (postcopulatory prezygotic isolation) and quantified proxies of sperm competitive ability (ejaculate size, sperm longevity, and sperm swimming performance) and cryptic female choice (sperm swimming performance in ovarian fluid of females from the same or opposite population). Body size impacted several of the traits analyzed in this study; larger males produced larger ejaculates that swam faster in ovarian fluid but were also less motile. However, there was a lack of evidence for population differences in sperm performance or cryptic female choice playing a role in reducing gene flow between ecotypes. Contrary to our predictions, sperm did not swim fastest in their local environment, neither in the context of ambient sulfide concentration nor ovarian fluid origin. Overall, our results suggest that mechanisms other than the postcopulatory prezygotic isolation might determine reproductive isolation between populations of *P. mexicana*.

Sperm competition and cryptic female choice in the family Poeciliidae

Contrary to our results, sperm competition and cryptic female choice are common phenomena in poeciliid fishes. For example, eastern mosquitofish that perceive higher sperm competition risk tend to be more sexually active and produce larger ejaculates (Evans et al. 2003). There is also evidence that ejaculates that are successfully transferred to a female can survive inside her reproductive tract and successfully fertilize eggs at least up to 10 months after the death of the male (López-Sepulcre et al. 2013). Higher sperm viability is associated with enhanced sperm competition success in numerous livebearers, including *Xiphophorus nigrensis* (Smith 2012) and guppies (Fitzpatrick and Evans 2014). However, when sperm are stored before fertilization, as is

common in many poeciliid species (Evans et al. 2011), males with higher sperm velocity counterintuitively have lower fertilization success, suggesting that velocity and viability—two key determinants of sperm competition success—may tradeoff with each other *in vivo* (Smith 2012).

Even though sperm competition is frequently investigated with a focus on male traits, sperm competition in internal fertilizers ultimately occurs in the female reproductive tract, and cryptic female choice can interact with sperm competitive traits to jointly shape postcopulatory prezygotic isolation. Evidence from guppies suggests that cryptic female choice mediated by ovarian fluid composition can lead to contrasting patterns of reproductive isolation, depending on the mating context. In one intrapopulation study, sperm competitive ability was higher when sperm velocity was measured in ovarian fluid from an unrelated female compared to when it was measured in ovarian fluid of a sister, suggesting cryptic female choice for unrelated males serves as a mechanism for inbreeding avoidance (Gasparini and Pilastro 2011). However, when comparing the effect of ovarian fluid origin on sperm competitive ability between populations inhabiting different river drainages, there was no difference in sperm competitive ability between populations within the same river drainage (as in our study). However, sperm activity was stimulated by ovarian fluid from females of the same river drainage compared to ovarian fluids derived from females in another river drainage (Devigili et al. 2018). These findings beg questions about whether more closely related lineages might exhibit different effects of cryptic female choice on sperm performance than do more distantly related lineages. Likewise, it remains unclear whether ecological similarity might play a role in predicting sperm performance and the direction and magnitude of cryptic female choice. Studies comparing the effects of sperm

competition and cryptic female choice in this system would benefit by including additional population pairs in other river drainages and measuring ecological similarity between habitats.

Effects of life history on sperm competition success

While our hypotheses about postcopulatory prezygotic isolation in *P. mexicana* were not supported by our experiments, our results shared several key similarities with other studies. In many animal species, body size is a key driver of reproductive success (molluscs: Downing et al. 1993; insects: Berg et al. 1997; fishes: Dickerson et al. 2002). Generally, larger females can have larger broods and higher lifetime reproductive success (as in Riesch et al. 2014). Males, on the other hand, can exhibit alternative reproductive tactics depending on their body size. In *P. mexicana*, females prefer larger males (Plath et al. 2004), which spend more time chasing off smaller rival males, can monopolize on matings, and impose lower costs associated with sexual harassment to females (Tobler et al. 2011b; Bierbach et al. 2014). In contrast, smaller *P. mexicana* males have a higher sexual activity and sneak past larger dominant males to force copulations regardless of female cooperation (Plath et al. 2007a). Theory on alternative reproductive tactics suggests that smaller sneaker males should produce larger, more competitive ejaculates per unit of body size, likely because they experience stronger sperm competition (Gross 1996; Taborsky 1998; Locatello et al. 2007). This has been corroborated in the swordtail *Xiphophorus nigrensis*, where smaller sneaker males produced ejaculates that were more viable and survived longer after ejaculation (Smith 2011). In contrast, we found that smaller males produced smaller ejaculates per unit of body size, but sperm from smaller males were more motile in control and sulfidic conditions. However, size-dependent effects were different in ovarian fluid; when activated in ovarian fluid—regardless of its origin relative to the male—

body size had no effect on motility, and larger males produced sperm that swam faster, perhaps indicating cryptic female choice for larger males. The chemical environment to which sperm are exposed clearly has important implications for sperm swimming performance that may lead to divergent patterns of sperm competitive ability.

We also detected an effect of treatment length on sperm performance, which corroborates previous work on sperm aging (Gasparini et al. 2010, 2014, 2017). We found that males that were in our feeding treatment for a longer time (i.e., they were isolated from females for longer) suffered a significant decrease in VSL, which would confer a poorer sperm competitive ability to those males. Older male guppies also have slower sperm, which may be caused by morphological changes during sperm aging (Gasparini et al. 2010, 2014). Furthermore, within the same male, older sperm swim slower and fertilize offspring that perform worse in sperm competition in future generations (Gasparini et al. 2017). It is therefore critical to standardize sperm age for any studies of sperm competition or cryptic female choice.

Ecological speciation in extremophile P. mexicana

If precopulatory isolation alone cannot account for the empirical data on reproductive isolation between populations of *P. mexicana* in southern Mexico (Plath et al. 2013), and postcopulatory prezygotic isolation is weak, as found in the present study, then the question remains what else contributes to the observed strong reproductive isolation between proximate populations in contrasting environments? There are three potential mechanisms that could be responsible for limiting gene flow between populations. First, there could be prezygotic barriers that we simply have not yet identified or do not understand, such as population differences in competitive ability (Schluter 2001) or disease susceptibility (Karvonen and Seehausen 2012). For example,

preliminary data suggests that sulfidic individuals are inferior competitors in nonsulfidic conditions (M. Tobler, unpublished data). Second, there could also be postcopulatory prezygotic barriers that were unaddressed in the present study. One promising avenue is divergence in reproductive microbiomes, which could lead to reproductive isolation, e.g., if seminal and ovarian microbes are incompatible in heterotypic matings (Rowe et al. 2020). Finally, postzygotic isolation could play a role in limiting gene flow between ecotypes, though there is a dearth of research on this topic. Genotypic data suggests hybrids are quite rare (Palacios et al. 2013; Plath et al. 2013; Riesch et al. 2016b), though putative hybrids have occasionally been found at the confluence of sulfidic and nonsulfidic habitats (R. Riesch, M. Plath, and M. Tobler, personal observations) and can be generated in the lab (J. Coffin, personal observations). The mechanisms limiting the production and/or success of hybrids remain to be resolved. Two non-mutually exclusive hypotheses have been proposed to understand what may promote postzygotic isolation between ecotypes of *P. mexicana* (Tobler et al. 2018). On one hand, there may be natural and/or sexual selection against hybrids. If hybrids have intermediate phenotypes that are maladapted to either parental habitat, then hybrids should be selected against by extrinsic factors and thus be rare in nature, assuming there is no intermediate habitat in which hybrids may be locally adapted (Gow et al. 2007). On the other hand, the main toxicity target of H_2S is complex IV of the electron transport chain, which has subunits encoded in both the mitochondrial and the nuclear genomes (Evans 1967; Cooper and Brown 2008). Evidence abounds for convergent modifications to this protein in sulfidic *P. mexicana* lineages (Greenway et al. 2020), so selection should lead to the coevolution of the mitochondrial and nuclear genes that code for this protein (Sunnucks et al. 2017). If nuclear and mitochondrial loci have co-evolved, then recombination in hybrid genomes may lead to combinations of nuclear and mitochondrial loci that have never

been present together in the same genetic background, potentially leading to fitness reductions mediated by intrinsic mitonuclear incompatibilities (Wolff et al. 2014; Tobler et al. 2019). One challenge of disentangling and testing these hypotheses is that hybrids are formed rarely in nature, so laboratory studies will be crucial to identify and quantify postzygotic isolation between sulfidic and nonsulfidic ecotypes of *P. mexicana*.

Conclusions

We quantified several metrics of sperm competitive ability (ejaculate size, sperm longevity, motility, and swimming velocity) and cryptic female choice (sperm motility and swimming velocity in ovarian fluid of females from the same or opposite population) in ecotypes of *P. mexicana* inhabiting sulfidic and nonsulfidic habitats to determine whether postcopulatory prezygotic isolation limits gene flow between habitats. Interestingly, we consistently found that male body size—rather than ecotype—was a significant driver of sperm competition success. The prevailing notion among biologists is that, because larger males invest more in somatic growth, they must have relatively less energy for other processes like reproduction. Therefore, a common prediction from theory is that smaller males should have larger relative investment (in terms of quantity and/or quality) in reproduction (Stearns 1989). This has been found in wild-caught *P. mexicana*, which show a positive correlation between body size and absolute gonad mass, but a negative correlation between body size and relative gonad mass (Schlupp et al. 2006). In line with this prediction, smaller males in our experiment produced more motile ejaculates in control and sulfidic conditions. However, our results suggest that these life history tradeoffs are not necessarily omnipresent. Instead, we also found that sperm from smaller males swam slower in ovarian fluid and that smaller males produced relatively smaller ejaculates. One

possible explanation for why our results differ from theoretical predictions is that smaller males may be of lower general quality than larger males. If larger males are better provisioned than smaller males for every aspect of their life history, then they may be able to afford to invest energy into growth and reproduction simultaneously (e.g., Reznick et al. 2000), which would obliterate commonly assumed trade-offs and result in the patterns observed in our study.

Our results were also contrary to our predictions about reproductive isolation between ecotypes. We expected to see large differences between ecotypes in sperm swimming performance, ejaculate size, and longevity, indicative of ecotypic differences in sperm competitive ability, and a strong positive effect of native ovarian fluid on this competitive ability. However, we found that ejaculate size, longevity, motility, and velocity showed no variation between ecotypes, and that ovarian fluid origin had no impact on sperm competitive ability, suggesting postcopulatory prezygotic isolation between ecotypes (at least that caused by the processes examined in this study) is quite weak. This study resoundingly suggests that isolation likely originates either before copulation or after fertilization. Further studies on precopulatory and postzygotic isolation between ecotypes of *P. mexicana* would help us better elucidate why and how speciation occurs in extreme environments.

Acknowledgements

We would like to thank the Mexican government for collection permits and the community of Tapijulapa, Tabasco, Mexico for their help making previous field expeditions a success. We would also like to thank our collaborators at the Universidad Juárez Autónoma de Tabasco and the Centro de Investigación E Innovación Para la Enseñanza y El Aprendizaje for their support. We benefitted greatly from assistance in the laboratory provided by H. Hoffman-Colburn, Q. La

Fon, and K. Rodriguez. This research was funded by the National Science Foundation (IOS-1931657, provided to M.T.), the U.S. Army Research Office (W911NF-15-1-0175, provided to M.T.) and the U.S. Department of Education (Graduate Assistance in Areas of National Need fellowship, provided to J.L.C.).

Data Accessibility

All datasets and R scripts for data analysis can be found on GitHub (https://github.com/michitobler/poecilia-sperm)

Figures

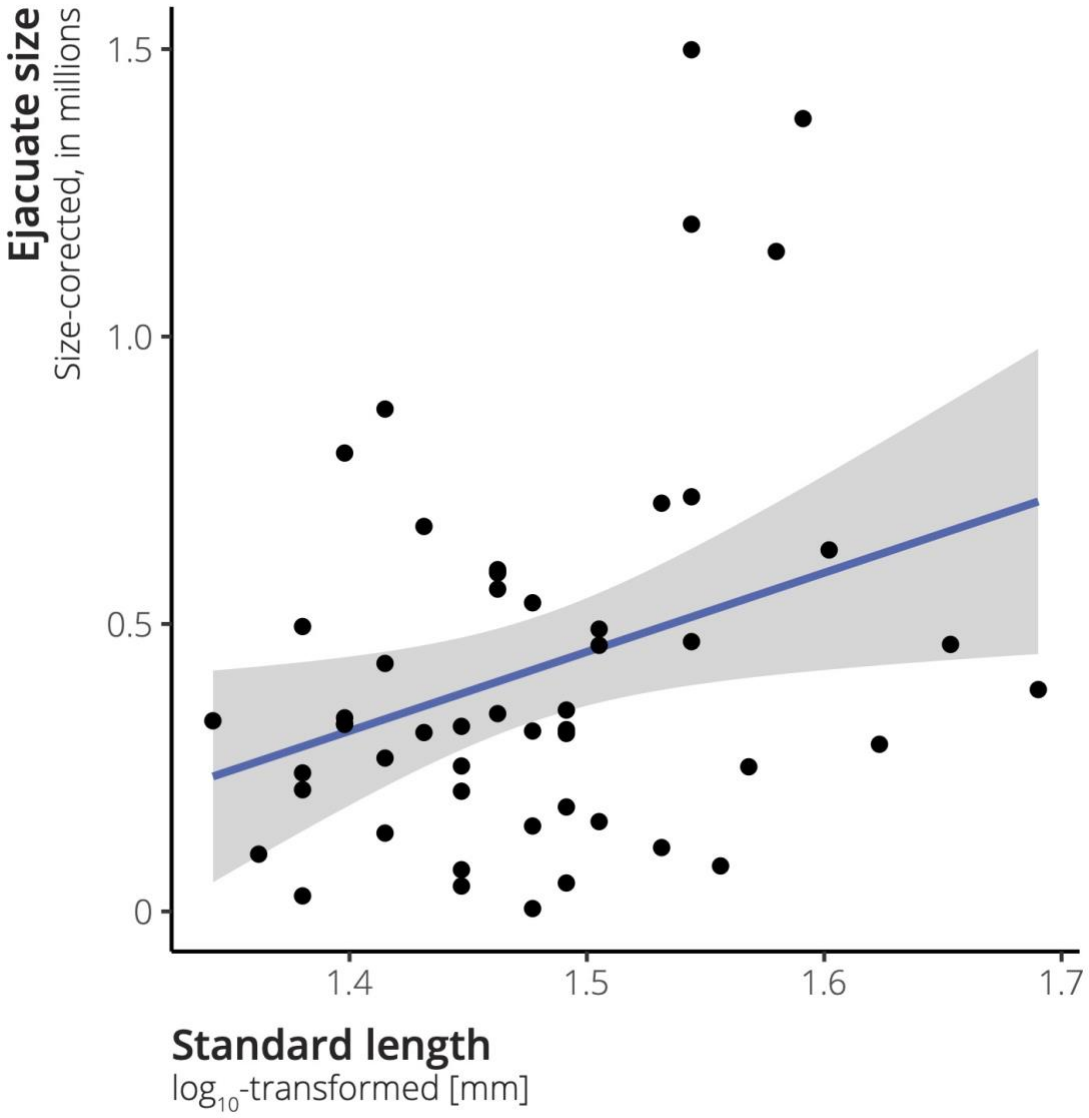

Figure 4-1: Scatterplot documenting variation in ejaculate size in relation to body size. To obtain values for the y-axis, raw ejaculate size was divided by log_{10}-transformed standard length to obtain ejaculate size per unit of body length. A line of best fit is shown in blue, which is surrounded by a prediction interval in grey.

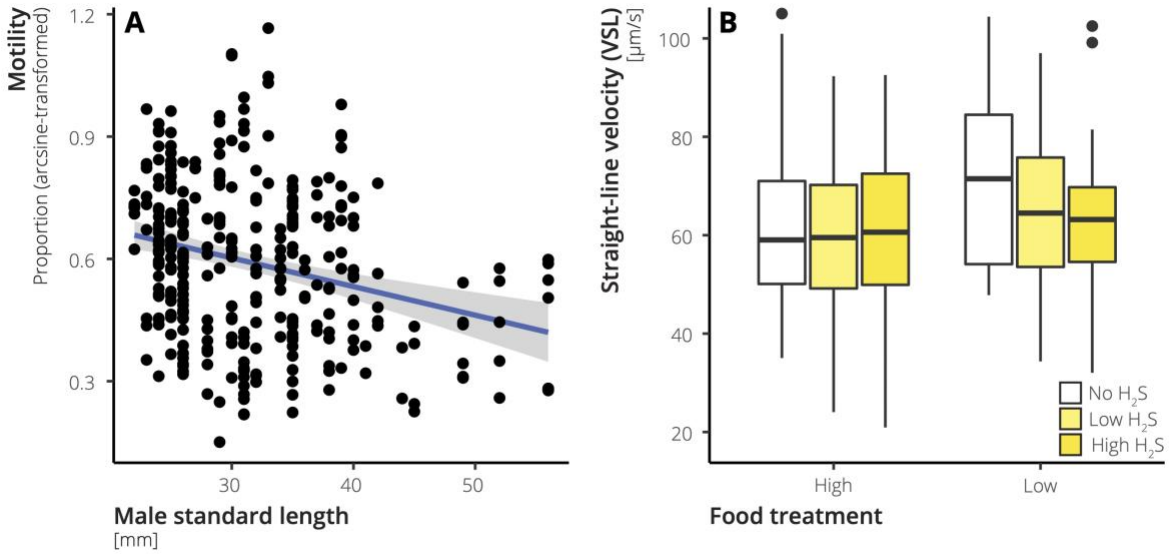

Figure 4-2: Sperm swimming performance data along a sulfide gradient. (A) Scatterplot portraying variation in arsine square root-transformed motility as a function of male body size, including measurements at all sulfide concentrations. A best-fit line is shown in blue, with a prediction interval shown in grey. There is a significant negative correlation between body size and motility, indicating that smaller males produce more motile ejaculates across the entire sulfide gradient. (B) Boxplots of sperm velocity (VSL), showing variation among sulfide concentrations between food treatments. Darker shades of yellow represent higher sulfide concentrations. There is no variation in VSL between sulfide concentrations in the high-food treatment, while VSL decreases at higher sulfide concentrations in the low-food treatment. These different responses to sulfide among the food treatments manifests with a significant food treatment × sulfide interaction.

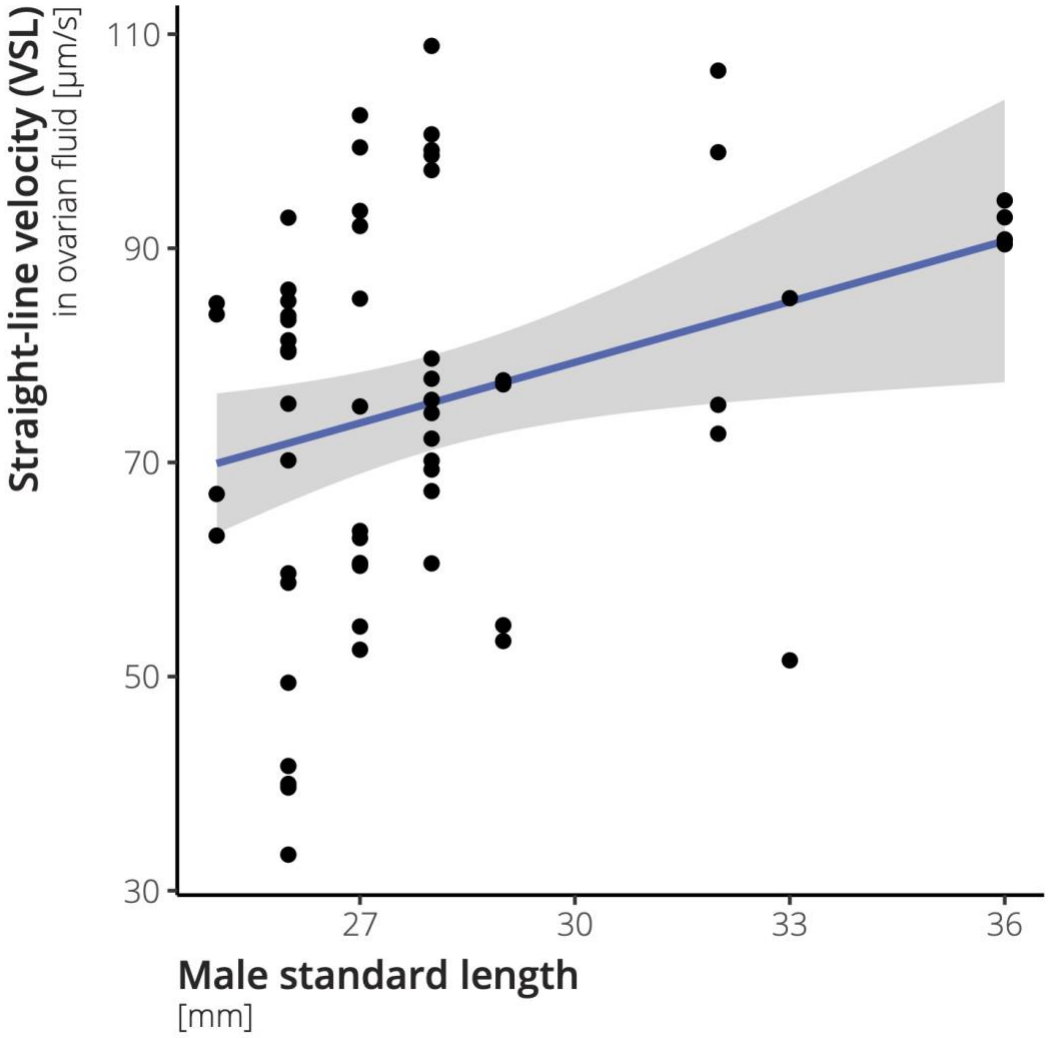

Figure 4-3: Scatterplot demonstrating variation in sperm velocity (VSL) in ovarian fluid, as a function of male body size. A best-fit line is shown in blue, and a prediction interval is shown in grey. There is a significant positive relationship between body size and VSL, indicating that in ovarian fluid, larger males produce faster—and thus more competitive—ejaculates.

Tables

Table 4-1: Descriptive statistics for all model terms included in the 95 % confidence set for each dependent variable. Model terms that are not present in the 'Term' column were not included in the averaged model because their inclusion did not improve model fit. Estimates were averaged among all models in the 95 % confidence set, weighted according to their respective Akaike weights (see Table C-2 for further information). *Z*-scores were calculated by dividing the Est. column by the Adj SE column, and *P*-values were calculated with the *Z*-scores. Longevity was not included in this table because a different statistical test was used to analyze variation in longevity.

DV	Model Term	Est.	CI	Adj SE	*Z*	*P*
Absolute Ejaculate Size	Ecotype	8.1E+4	-1.9E+5, 3.6E+5	1.4E+5	0.575	0.565
	Male SL	**3.3E+4**	**8.4E+3, 5.7E+4**	**1.2E+4**	**2.635**	**0.008**
	Treatment Length	6.7E+3	-4.6E+3, 1.8E+4	5.8E+3	1.158	0.247
	Food Treatment	1.6E+5	-1.1E+5, 4.4E+5	1.4E+5	1.154	0.248
Relative Ejaculate Size	Ecotype	5.7E+4	-1.3E+5, 2.5E+5	9.7E+4	0.593	0.553
	$\log_{10}$(Male SL)	**1.3E+6**	**1.9E+5, 2.6E+6**	**6.1E+5**	**2.265**	**0.024**
	Treatment Length	4.0E+3	-3.4E+3, 2.5E+5	3.8E+3	1.070	0.284
	Food Treatment	1.0E+5	-8.4E+4, 2.5E+5	9.7E+4	1.086	0.277
	Ecotype × Food Treatment	6.8E+4	-3.1E+5, 4.5E+5	1.9E+5	0.353	0.724
MOT (Sulfide)	Timepoint	-0.024	-0.05, 0.001	0.013	1.834	0.067
	Ecotype	0.029	-0.06, 0.11	0.043	0.663	0.507
	Food Treatment	-0.008	-0.10, 0.08	0.046	0.183	0.855
	Male SL	-0.006	-0.01, -0.001	0.003	2.245	0.025
VSL (Sulfide)	Food Treatment	5.384	-2.47, 13.24	4.006	1.344	0.179
	Sulfide (50 μM)	-2.030	-8.14, 4.08	3.115	0.652	0.515
	Sulfide (100 μM)	0.553	-5.62, 6.73	3.151	0.176	0.861
	Timepoint	-2.535	-5.165, 0.10	1.342	1.888	0.059
	Food Treatment × Sulfide (50 μM)	-5.890	-12.85, 1.07	3.551	1.659	0.097
	Food Treatment × Sulfide (100 μM)	**-8.781**	**-15.49, -2.07**	**3.422**	**2.566**	**0.010**
	Ecotype	1.029	-6.15, 8.21	3.663	0.281	0.779
	Treatment Length	**-0.217**	**-0.37, -0.06**	**0.079**	**2.728**	**0.006**
	Ecotype × Sulfide (50 μM)	5.726	-1.12, 12.57	3.492	1.640	0.101
	Ecotype × Sulfide (100 μM)	5.320	-1.17, 11.81	3.312	1.606	0.108
MOT (OF)	Ecotype	-0.050	-0.20, 0.10	0.075	0.665	0.506
	Timepoint	-0.022	-0.07, 0.03	0.027	0.825	0.410
VSL (OF)	OF origin	1.881	-12.91, 16.67	7.547	0.249	0.803
	Male SL	**2.406**	**0.93, 3.88**	**0.753**	**3.195**	**0.001**
	Male Ecotype	4.450	-10.98, 19.88	7.871	0.565	0.572
	OF origin × Male Ecotype	-18.256	-42.03, 5.52	12.129	1.505	0.132
	Timepoint	-4.303	-11.42, 2.81	3.630	1.185	0.236
	OF origin × Timepoint	-0.373	-11.22, 10.48	5.536	0.067	0.946
	Male Ecotype × Timepoint	-4.528	-15.44, 6.38	5.565	0.814	0.416
	Female SL	0.539	-0.52, 1.59	0.538	1.002	0.316

Appendix A - Impacts of heavy metal pollution on the ionomics and transcriptomics of Western mosquitofish (*Gambusia affinis*)

John L. Coffin, Joanna L. Kelley, Punidan D. Jeyasingh, and Michael Tobler

Supplementary analyses and results

Discriminating individuals of each species between sites based on ionomic signatures

To further assess our prediction that all species except mosquitofish would experience elemental shifts in Tar Creek populations, we conducted separate linear discriminant analyses (LDA) for each species to classify individuals into their site of origin based on their ionomic composition. To accomplish this, we divided the overall ionomic dataset into individual species and conducted a principal component analysis separately for each species to reduce data dimensionality. We used the prcomp function with a correlation matrix in R and retained principal components that had an eigenvalue greater than 1. Site was specified as the grouping factor in the LDA, and the principal component scores (from as many axes as were retained for that species) for each individual were input as a formula into the lda function in the MASS package in R (Venables and Ripley 2002). Prior probabilities for Tar and Coal Creek groups were both set to 0.5, and a "hold-one-out" cross-validation procedure was conducted by setting CV = TRUE. To evaluate the accuracy of the linear discriminant function generated by this procedure, we compared the assignment of individuals to sites based on the LDA with their known sampling sites and calculated the frequency of correct assignment for each species (Table A-7). Contrary to our prediction, we were able to correctly classify all species by site with over 75 % accuracy based on the discriminant functions generated from our ionomic dataset; of particular note, four species—*Gambusia affinis*, *Fundulus notatus*, *Pimephales notatus*, and *Lepomis cyanellus*—

were classified with 100 % accuracy. This indicated that mosquitofish—as all other resident species—experience an ionomic shift large enough to distinguish individuals between polluted and unpolluted sites in multivariate space.

Identifying co-expressed genes that correlated with pollution

As a complementary approach to the differential expression analysis, we constructed gene co-expression modules for samples from each tissue separately using weighted gene co-expression network analysis (Zhang and Horvath 2005) to identify modules of co-expressed genes that were associated with the presence of environmental metals. In order to satisfy the assumptions of WGCNA of an approximate scale-free topology of the data, we implemented a variance-stabilizing transformation of the filtered reads using the varianceStabilizingTransformation function from the DESeq2 package (version 1.22.2; Love, Huber, & Anders, 2014), as recommended in the WGCNA package documentation (Langfelder and Horvath 2008, 2012). The filtered, transformed counts for each sample were then normalized based on library size. We then used the transformed counts to calculate the $\log_2$-cpm (counts per million mapped reads) for each gene, using the cpm function from the edgeR package. This scales the read counts of each sample by that sample's library size, which accounts for the natural bias toward higher read counts in libraries with more fragments sequenced.

We clustered all samples based on overall expression using the hclust function from the flashClust package (Langfelder & Horvath, 2012). Then, to construct a weighted correlation network of genes, we constructed an adjacency matrix using the adjacency function from the WGCNA package (version 1.68; Langfelder & Horvath, 2008, 2012). In brief, this calculates the Pearson correlation between each pair of genes and raises this correlation coefficient to a soft

thresholding parameter (β), which we selected by choosing the smallest possible value that maximized both the scale-free topology model fit and mean connectivity between genes in our network. Because we analyzed our data for each tissue separately, we chose a different value of β for each tissue. Based on our data and recommendations from Zhang and Horvath (2005), we used a β of 10 for gill and liver analyses, and 9 for brains.

To obtain modules of co-expressed genes from our correlation network, we computed the dissimilarity between genes with a topological overlap dissimilarity matrix, which was calculated by subtracting the topological overlap measure (TOM) for each gene from 1. This dissimilarity matrix was then used as a metric of distance to hierarchically cluster the genes into a dendrogram. To detect modules within this dendrogram, we used the cutreeDynamic function from the dynamicTreeCut package (Langfelder & Horvath, 2008; Langfelder, Zhang, & Horvath, 2016), and then calculated the eigengene—i.e., the first principal component—for each module to summarize a large proportion of the variation in gene expression for each module, using the moduleEigengenes function from the WGCNA package. Eigengenes with a correlation > 0.55 were merged (Figure A-2), and each remaining eigengene was correlated with data on the presence or absence of environmental metals for each sample. *P*-values were computed for each correlation coefficient using the corPvalueStudent function to determine which modules were significantly associated with environmental conditions. Genes belonging to modules with a *P*-value less than 0.05 were extracted for further analyses.

Co-expression patterns identified through network-based analyses largely mirror gene-by-gene results

Across the three tissues, we found nine modules that were significantly correlated with environmental metals, two in gills, four in livers, and three in brains. In gills, co-expressed genes in the brown module were positively correlated with environmental metals and were enriched for GO processes related to RNA processing, cell adhesion, and mitosis. The blue module, negatively correlated with metals in the gill, was enriched for GO processes related to immune system functions, as was found in all modules that were negatively correlated with metals, across all three tissues. In livers, there were two modules (lavenderblush3 and antiquewhite4) that were positively correlated with environmental metals and two modules (darkolivegreen and darkturquoise) that were negatively correlated with metals. The positively correlated modules in the liver were enriched for metabolism and transport biological processes, while the two negatively correlated modules were enriched in genes involved with immunity, biosynthetic processes, transcription, and cell signaling and movement. In the brain, the white module was positively correlated with metals, but this module was not enriched for any GO processes. In contrast, the two modules in the brain that were negatively correlated with metals (floralwhite and greenyellow) were enriched in genes involved with cell adhesion and motility. Table A-8 contains statistics for module membership, correlation between that gene's expression and the presence of metals in the environment, and relevant *P*-values for each gene in each tissue. None of the modules across the three tissues sampled were enriched with genes involved in metal homeostasis or pathways that we predicted to be modulated in response to heavy metal exposure.

Characterization of genetic differentiation among sites

While gene expression differences between sites are likely in part due to environmental variation, and in this case specifically differences in heavy metal exposure, genetic variation

between populations could also lead to constitutive differences in gene expression not mediated by heavy metal stress (Passow et al. 2017). Using the mapped transcriptomic data from our analyses of gene expression, we quantified genetic variation between sites using ANGSD (version 0.931; Korneliussen, Albrechtsen, & Nielsen, 2014). Genotype likelihoods for the mapped reads were calculated according to the GATK model (by specifying -GL 2). The major and minor alleles were inferred (-doMaf 1), and sites with a minor allele frequency less than 0.05 and a *P*-value greater than 2×10^{-6} were removed from the analysis (-minMaf 0.05 and -SNP_pval 2e-6). The resulting genotype likelihoods were output as a beagle file (-doGlf 2) for further analysis.

To estimate individual admixture proportions among the samples in our analysis, we used NGSadmix (bundled with ANGSD; Skotte, Korneliussen, & Albrechtsen, 2013). The genotype likelihoods from the beagle file were input and the number of genetic clusters expected was specified (setting -K to an integer value). We calculated admixture proportions for K = 1 through K = 5 and repeated each of these calculations ten times. Each replicate run of NGSadmix was given a unique seed to aid in reproducibility. For each value of K, there were ten output files containing the log-likelihood of the estimates of admixture proportions from the run of NGSadmix. These log-likelihoods were then used to calculate ΔK for each value of K, following the procedure established by Evanno et al. (2005). We found that K = 1 had the largest ΔK (Figure A-3), suggesting that there is only one genetic cluster in our dataset.

Identifying patterns of local adaptation with a reciprocal transplant

As a complement to our genetic differentiation data, we also evaluated survival and physiological data to understand whether mosquitofish are adapted—locally or generally—to life

in Tar Creek. Both of these datasets compared mosquitofish from Tar Creek to those from Coal Creek, a nearby unpolluted stream. We first conducted a reciprocal transplant over three days to determine how endogenous concentrations of reactive oxygen species (ROS) varied between fish in their native environment (i.e., residents) and after transplantation in the other habitat type (i.e., transplants). We expected transplants from Coal Creek to have the highest endogenous ROS concentrations, and that Tar Creek residents would have lower endogenous ROS concentrations than Coal Creek transplants, indicating that Tar Creek individuals are locally adapted to the conditions in Tar Creek. Instead, we found similar ROS concentrations between residents and transplants from both habitats (Figure A-4), indicating that Tar Creek individuals are likely not better able to tolerate the extreme conditions in Tar Creek than those from Coal Creek, which suggests they are not locally adapted. This finding was corroborated by another reciprocal transplant that measured survival over a two-day exposure, which found that residents did not have higher survival than transplants in either habitat (Figure A-4). Taken together, this data suggests that evolutionary adaptation does not underly tolerance to heavy metal pollution in mosquitofish.

Appendix A Figures

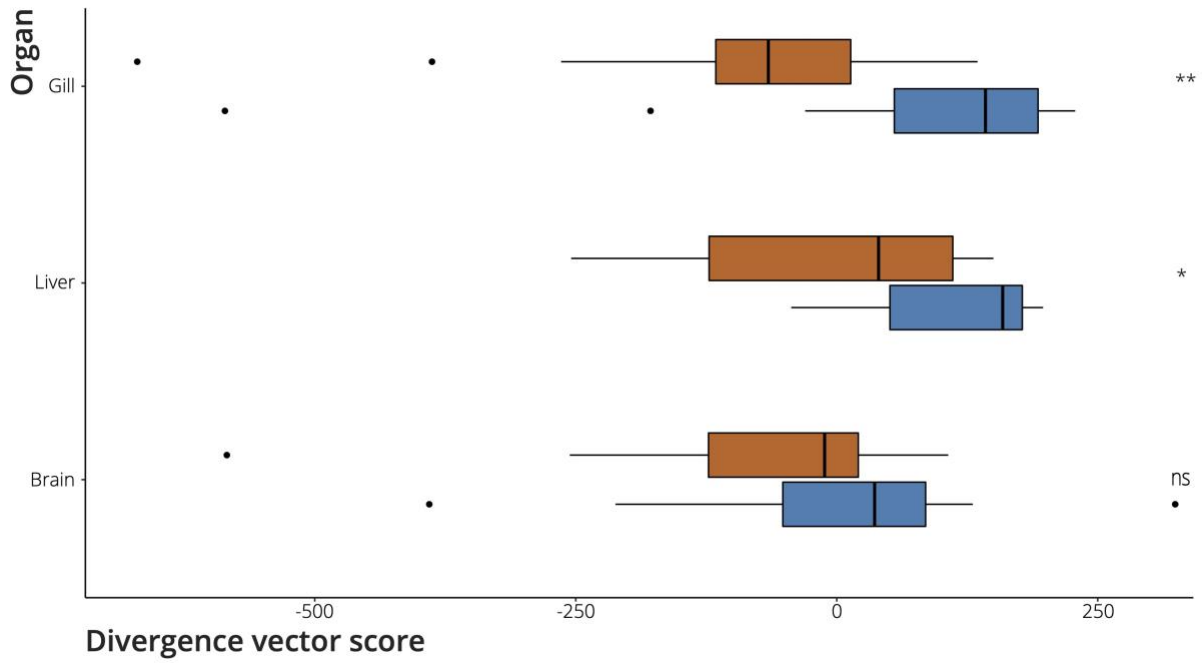

Figure A-1: Boxplots of divergence vector scores for the 'Site' term in the MANOVA comparing ionomic differentiation in gill, liver, and brain tissues in mosquitofish. Blue corresponds to mosquitofish from Coal Creek, while orange corresponds to Tar Creek individuals. Gill and liver tissues were significantly different between sites, while brain tissues were not.

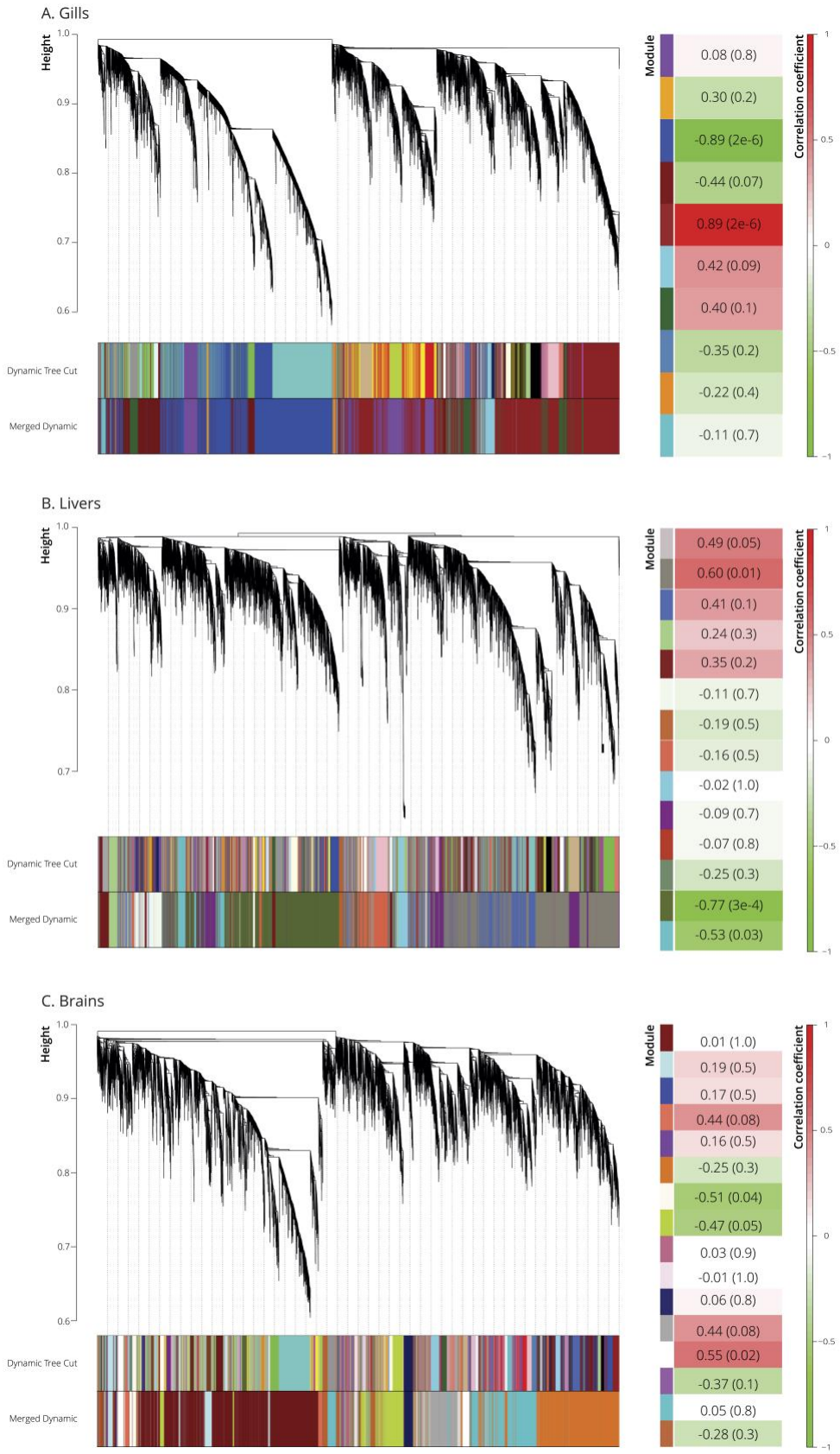
A. Gills
Height
1.0
0.9
0.8
0.7
0.6
Dynamic Tree Cut
Merged Dynamic
Module
0.08 (0.8)
0.30 (0.2)
-0.89 (2e-6)
-0.44 (0.07)
0.89 (2e-6)
0.42 (0.09)
0.40 (0.1)
-0.35 (0.2)
-0.22 (0.4)
-0.11 (0.7)
Correlation coefficient
1
0.5
0
-0.5
-1
B. Livers
Height
1.0
0.9
0.8
0.7
Dynamic Tree Cut
Merged Dynamic
Module
0.49 (0.05)
0.60 (0.01)
0.41 (0.1)
0.24 (0.3)
0.35 (0.2)
-0.11 (0.7)
-0.19 (0.5)
-0.16 (0.5)
-0.02 (1.0)
-0.09 (0.7)
-0.07 (0.8)
-0.25 (0.3)
-0.77 (3e-4)
-0.53 (0.03)
Correlation coefficient
1
0.5
0
-0.5
-1
C. Brains
Height
1.0
0.9
0.8
0.7
0.6
Dynamic Tree Cut
Merged Dynamic
Module
0.01 (1.0)
0.19 (0.5)
0.17 (0.5)
0.44 (0.08)
0.16 (0.5)
-0.25 (0.3)
-0.51 (0.04)
-0.47 (0.05)
0.03 (0.9)
-0.01 (1.0)
0.06 (0.8)
0.44 (0.08)
0.55 (0.02)
-0.37 (0.1)
0.05 (0.8)
-0.28 (0.3)
Correlation coefficient
1
0.5
0
-0.5
-1

Figure A-2: WGCNA results for all 18,693 genes in gill, liver, and brain tissues. Agglomerative hierarchical clustering trees using average linkage, based on topological overlap distance in gill (A), liver (B), and brain (C) samples, respectively. Correlation coefficients between module eigengenes (referred to with colors on the left axis) and habitat type (polluted for Tar Creek or unpolluted for Coal Creek and Little Elm Creek) are found for each tissue to the right of the clustering tree. *P*-values are found in each cell in parentheses.

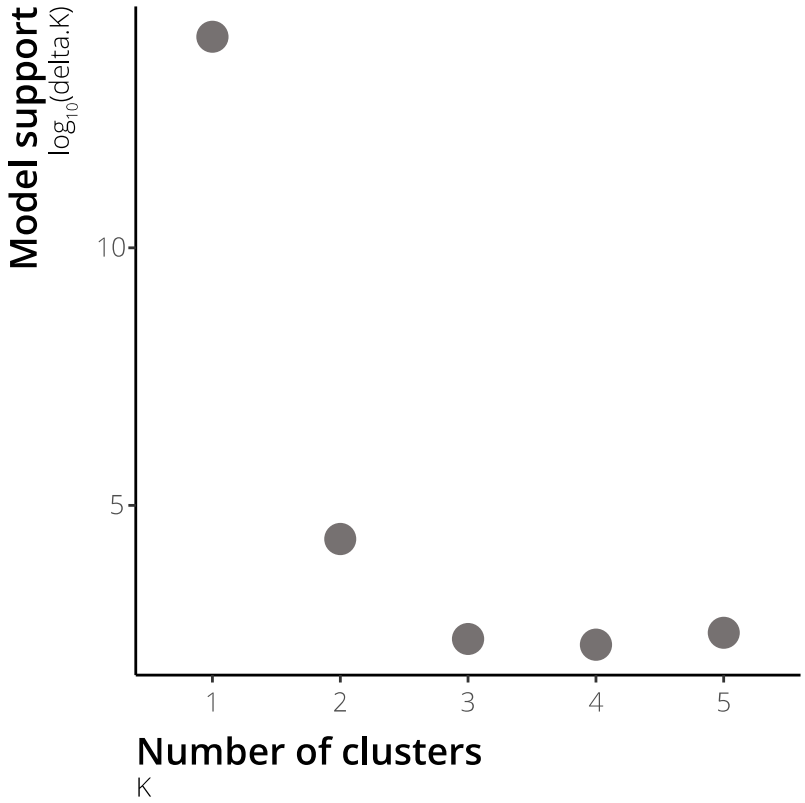

Figure A-3: Dot plot showing evidence for the most likely number of genetic clusters (indicating the most likely number of populations) in our transcriptome dataset. The x-axis is the number of possible populations (K), ranging from one to five. The y-axis shows the log_{10}-transformed value of ΔK, which was calculated according to Evanno et al. (2005). In brief, we ran NGSadmix 10 times for each possible value of K and calculated the average log-likelihood and standard deviation for all ten replicates for each value of K. We then divided the average by the standard deviation to obtain a value of ΔK for each possible value of K. These values were log_{10}-transformed for ease of viewing. Values of K with higher values of ΔK signify stronger evidence for that number of populations. We found that K = 1 had the highest value of ΔK by far, suggesting that there is likely only one panmictic population of mosquitofish that inhabits both polluted and unpolluted habitats.

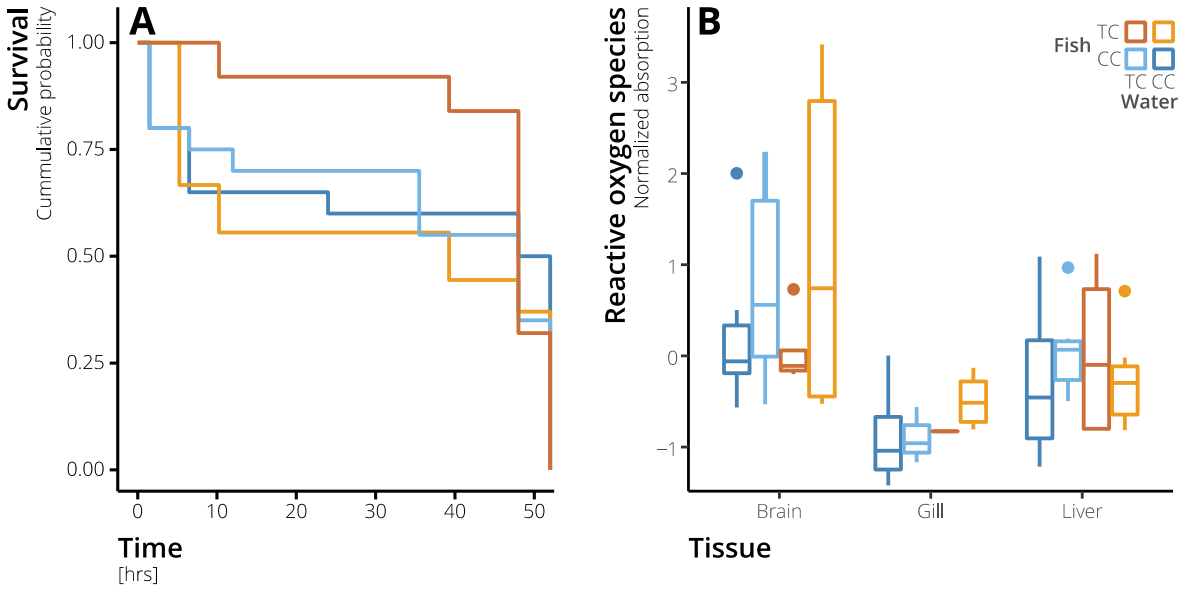

Figure A-4: Plots presenting survival and physiological data after reciprocal transplants of mosquitofish between Tar Creek (shown in shades of orange) and Coal Creek (shown in shades of blue). (A) Plot showing the survival probability of fish caught and placed back into their native environment (i.e., residents, shown in dark orange and dark blue) compared to the survival probability of fish transplanted into the habitat type that is opposite of the one in which they were found (i.e., transplants, shown in light orange and light blue). Survival was tracked over two days, and a log-rank test determined that survival was not significantly different between residents and transplants from either habitat (P = 0.68), indicating that mosquitofish in Tar Creek and Coal Creek are not likely locally adapted. (B) Normalized (Z-transformed) endogenous reactive oxygen species (ROS) concentrations in brains, gills, and livers of mosquitofish from both habitats. While there was significant differentiation in ROS concentrations between tissues, there was no difference within-tissues between residents and transplants, indicating that Tar Creek mosquitofish experience similar endogenous ROS concentrations that Coal Creek individuals.

Appendix A Tables

Table A-1: Loadings of each element quantified along each of the first six principal components. The proportion of variance explained by the addition of each principal component axis is given in parentheses at the base of each column. Loadings were calculated as the eigenvector multiplied by the square root of the eigenvalue for each principal component axis and represent the correlation coefficient between the original element variables and the linearly transformed principal component scores.

	PC1	PC2	PC3	PC4	PC5	PC6
Al	0.057	0.327	-0.315	-0.478	-0.485	-0.34
As	0.494	-0.111	0.41	0.181	-0.258	-0.441
B	0.657	0.549	-0.107	0.032	-0.175	-0.006
Ba	0.236	-0.876	-0.139	0.108	-0.269	0.097
Be	0.85	-0.119	-0.069	0.394	0.067	-0.029
Bi	0.878	-0.119	-0.053	0.383	0.019	-0.026
Ca	0.715	-0.193	0.483	-0.34	0.101	-0.165
Cd	0.639	0.466	0.109	0.08	0	0.268
Co	0.727	-0.107	-0.525	-0.055	0.028	0.018
Cr	0.78	0.435	-0.207	-0.128	-0.055	0.054
Cu	0.573	-0.234	0.265	-0.022	-0.437	-0.012
Fe	0.703	0.495	-0.217	0.146	-0.156	0.128
Mn	0.122	-0.71	-0.115	-0.32	0.02	-0.135
Mo	0.467	0.411	-0.057	-0.11	0.371	-0.473
Ni	0.475	-0.694	-0.239	0.227	-0.318	0.12
P	0.63	-0.45	0.283	-0.476	0.229	0.056
Pb	0.588	0.049	0.27	-0.098	0.288	0.325
S	0.77	-0.313	0.455	-0.098	-0.038	-0.116
Se	0.657	-0.257	-0.284	0.087	0.229	0.001
Si	0.223	0.528	-0.157	-0.588	-0.227	0.26
Sr	0.061	-0.72	-0.32	-0.401	0.102	0.203
Tl	0.935	-0.004	-0.121	-0.014	0.033	0.042
V	0.501	-0.009	-0.625	-0.004	0.27	-0.268
Zn	0.798	0.344	0.359	-0.01	-0.013	0.177
Eigenvalue	9.149	4.41	2.161	1.666	1.203	1.033
% Variance	38.1	18.4	9	6.9	5	4.3

Table A-2: Descriptive statistics for read counts after Illumina HiSeq 2500 paired-end (PE) sequencing. Reads were used for transcriptome profiling of gill, liver, and brain tissues in *G. affinis* populations from polluted Tar Creek and unpolluted Coal and Little Elm Creeks. The number of paired-end reads are provided for each step in our transcriptome analysis pipeline.

Population	Tissue	N	Avg number of raw PE reads	Avg number of trimmed PE reads	Avg number of PE reads after subsampling	Avg number of PE reads mapped
Tar	Gill	5	12,730,165 ± 6,205,316	12,365,478 ± 5,955,697	10,349,872 ± 2,628,120	9,961,866 ± 2,568,214
Tar	Liver	6	10,505,175 ± 4,248,369	10,259,114 ± 4,153,355	9,751,843 ± 3,428,112	9,582,266 ± 3,353,925
Tar	Brain	5	17,838,424 ± 4,871,591	17,386,368 ± 4,836,670	12,979,916 ± 44,909	12,633,116 ± 231,929
Coal	Gill	6	11,237,098 ± 4,082,488	10,934,071 ± 3,960,442	10,245,829 ± 3,135,361	9,971,418 ± 3,057,832
Coal	Liver	6	11,650,728 ± 4,051,906	11,368,373 ± 4,011,393	10,366,247 ± 2,501,218	10,231,575 ± 2,467,142
Coal	Brain	6	10,594,294 ± 4,227,739	10,288,571 ± 4,064,901	9,549,573 ± 2,752,483	9,357,706 ± 2,683,581
Little Elm	Gill	5	15,370,056 ± 7,997,031	14,828,466 ± 7,564,710	11,250,249 ± 2,449,860	10,972,733 ± 2,371,375
Little Elm	Liver	6	16,175,514 ± 3,960,762	15,665,233 ± 3,865,102	12,367,561 ± 1,549,153	12,193,948 ± 1,542,456
Little Elm	Brain	6	11,357,281 ± 4,089,540	11,038,102 ± 4,043,644	10,287,914 ± 3,206,149	10,069,595 ± 3,152,419

Table A-3: Raw gene matrix of non-normalized read counts for all candidate genes identified by Stringtie. Functional information for each gene was obtained from human orthologs in the SwissProt database, shown in the Annotation column, and protein name was extracted from the *Xiphophorus maculatus* reference genome annotation file (.gff). Gill samples are shown in blue, liver samples are shown in yellow, and brain samples are shown in orange. Likelihood ratio tests (LRTs) were used to compare read counts for each gene between each unpolluted site and Tar Creek in each tissue, and summary statistics from each LRT can be found to the right of the gene matrix. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tab "Table A-3".

Table A-4: Results from likelihood ratio tests for significantly differentially expressed genes (FDR < 0.05) for each tissue × site comparison. Results are sorted by logFC and color-coordinated based on the direction of differential expression. Genes colored green were upregulated in comparison to Tar Creek, whereas genes colored in red were downregulated in comparison to Tar Creek. (A) LRT results comparing gene expression between gill samples from Little Elm Creek and Tar Creek. (B) LRT results comparing gene expression between gill samples from Coal Creek and Tar Creek. (C) LRT results comparing gene expression between liver samples from Little Elm Creek and Tar Creek. (D) LRT results comparing gene expression between liver samples from Coal Creek and Tar Creek. (E) LRT results comparing gene expression between brain samples from Little Elm Creek and Tar Creek. (F) LRT results comparing gene expression between brain samples from Coal Creek and Tar Creek. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tabs "Table A-4A", "Table A-4B", "Table A-4C", "Table A-4D", "Table A-4E", and "Table A-4F".

Table A-5: Filtered results from GO enrichment analysis with GOrilla. Column 'N' is the total number of non-duplicate genes with unique GO identifiers present in the background set, Column 'B' is the number of genes in the background set belonging to a particular GO term, Column 'n' is the total number of non-duplicate genes with unique GO identifiers in the target set, and Column 'b' is the number of genes in the target set that belong to a particular GO term. Enrichment is calculated as (b/n)/(B/N). Results were filtered to retain enriched GO terms with FDR <0.05, enrichment score > 2.0, and the number of genes belonging to the GO term (b) > 5. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tab "Table A-5".

Table A-6: Results from a query of the Gene Ontology database of GO terms related to our hypotheses about differential expression in response to heavy metal exposure. The search was restricted to the term "metal" and to biological processes, resulting in a list of 66 biological process GO terms that are related to metals. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tab "Table A-6".

Table A-7: Results from linear discriminant analyses of ionomic data from two populations of *G. affinis.* The number of individuals correctly or incorrectly classified into their known sampling sites based on linear discriminant analyses, as well as the percent of correct assignment, are included. All species were correctly classified > 75 % of the time. Coefficients of the first linear discriminant are shown for all input variables for each species.

	Gambusia affinis	***Fundulus notatus***	***Pimephales notatus***	***Lepomis cyanellus***	***Lepomis gulosus***	***Lepomis macrochirus***	***Lepomis megalotis***
N correct	12	13	12	10	8	10	13
N incorrect	0	0	0	0	2	3	1
Total *N*	12	13	12	10	10	13	14
Assign %	100	100	100	100	80	76.923	92.857
LD1-PC1	0.567	2	1.604	1.792	0.488	0.446	1.127
LD1-PC2	1.187	0.727	-1.046	-2.036	-0.311	0.495	0.462
LD1-PC3	-1.031	-0.035	0.449	-0.226	-0.497	-0.193	0.035
LD1-PC4	-0.285	1.077	0.492	-0.377	-0.079	-0.177	-0.548
LD1-PC5	0.994	N/A	N/A	0.053	N/A	N/A	0.11

Table A-8: Results of WGCNA of all 18,693 genes identified in our RNA-seq analysis. For each gene, the module to which it belongs, the correlation coefficient between the gene's expression and environmental data, and the P-value for this correlation coefficient are provided. Also provided are the correlations and P-values between this gene and all other modules identified. (A) WGCNA results for gills. (B) WGCNA results for livers. (C) WGCNA results for brains. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tabs "Table A-8A", "Table A-8B", and "Table A-8C".

Appendix B - Genetics and resource availability shape life history and behavioral divergence between ecotypes of Atlantic mollies (*Poecilia mexicana*)

Appendix B Tables

Table B-1: Descriptive information for principal component (PC) analyses for multiple traits. Loadings and eigenvalues for each PC axis for each multivariate phenotype. The cumulative proportion of variance explained by each PC axis is given at the base of each column for each phenotype. Loadings were calculated as the eigenvector multiplied by the square root of the eigenvalue for each PC axis and represent the correlation coefficient between the original phenotypic data and the linearly transformed principal component scores. The table contains information for PC analyses that were conducted for (A) Burst Swimming, (B) Exploratory Behavior, and (C) Multivariate Trait Variation.

Phenotype	Input Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
(A) Burst swimming	V_{max}	0.963	-0.029	-0.270				
	A_{max}	0.911	-0.380	0.159				
	D_{net}	0.901	0.416	0.127				
	Eigenvalue	2.570	0.318	0.114				
	Variance Explained	0.856	0.962	1.000				
(B) Exploratory Behavior	V_{avg}	0.833	0.469	-0.252	-0.151	0.006		
	V_{max}	0.809	-0.581	0.041	0.015	0.073		
	A_{max}	0.804	-0.587	0.055	-0.012	-0.073		
	D_{total}	0.819	0.486	-0.265	0.149	-0.007		
	D_{center}	0.420	0.365	0.831	0.002	0.001		
	Eigenvalue	2.842	1.273	0.829	0.045	0.011		
	Variance Explained	0.569	0.823	0.989	0.998	1.000		
(C) Multivariate Trait Variation	Size at Birth	-0.844	0.188	-0.167	-0.111	-0.365	-0.181	0.215
	Brood Size	0.672	-0.023	-0.304	0.300	-0.174	-0.578	0.011
	Age at Maturity	-0.302	-0.728	-0.096	0.129	0.562	-0.151	0.121
	Growth Rate	0.850	-0.118	-0.006	-0.418	-0.038	0.194	0.225
	Burst Swimming	-0.007	0.628	0.430	-0.374	0.390	-0.357	0.015
	Exploratory Behavior	-0.106	-0.688	0.240	-0.563	-0.276	-0.226	-0.116
	Feeding Accuracy	0.080	-0.178	0.847	0.448	-0.191	-0.014	0.087
	Eigenvalue	1.995	1.480	1.090	0.951	0.746	0.607	0.133
	Variance Explained	0.285	0.496	0.652	0.788	0.894	0.981	1.000

Table B-2: Full model selection tables for each phenotype measured. Models were weighted by AICc, and the best-supported model (with a delta value of 0) was used in analyses. Continuous variables that are included in models are denoted with the regression coefficient for that model term, and categorical variables that were included are denoted with a '+'. Terms that were not included in a model are denoted with 'NA'. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tab "Table B-2".

Appendix C - Do sperm competition and cryptic female choice impact reproductive isolation in locally adapted populations of a livebearing fish?

Appendix C Figures

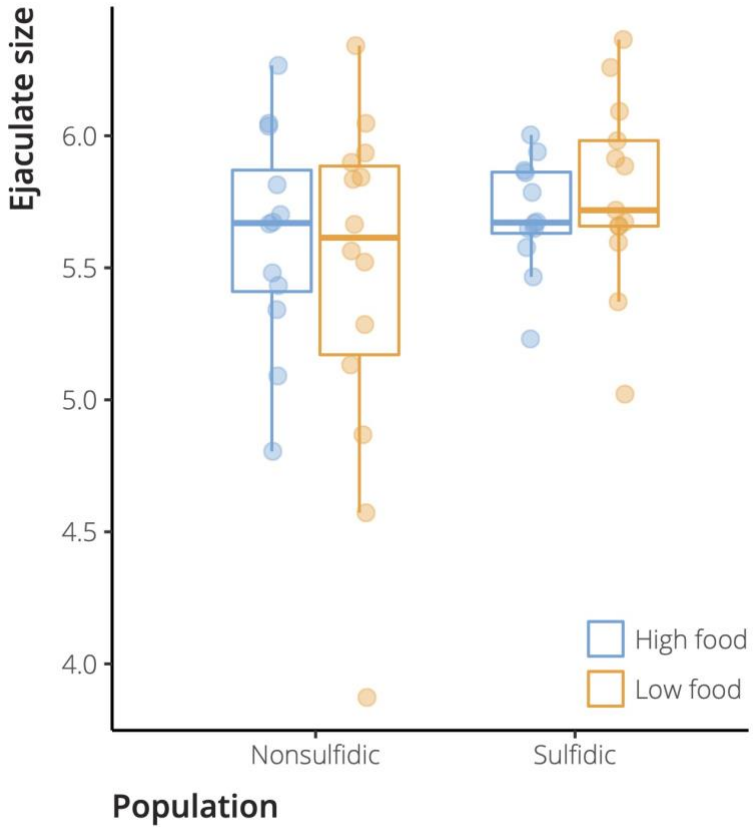

Figure C-1: Boxplot of log_{10}-transformed ejaculate size across ecotypes and food treatments in the Tacotalpa drainage. Nonsulfidic males are portrayed on the left, and sulfidic males are portrayed on the right. The fill of each boxplot denotes the food treatment, with males in the high food treatment colored red and males in the low food treatment colored blue. We did not detect significant differences between ecotypes or food treatments.

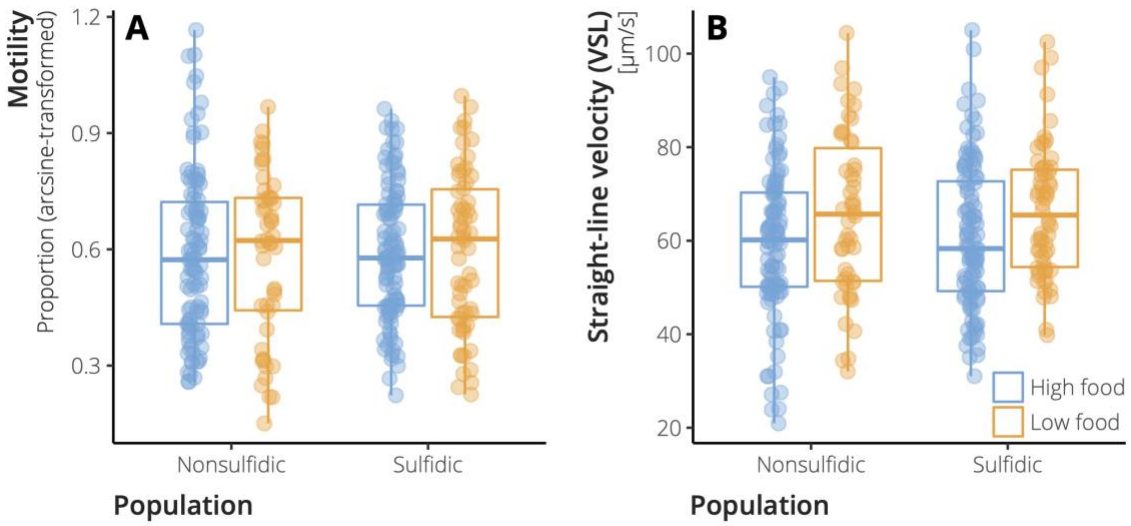

Figure C-2: Boxplots of sperm swimming performance metrics across ecotypes and food treatments. Ecotypes represent sulfidic and nonsulfidic individuals from both the Tacotalpa and Puyacatengo drainages. Nonsulfidic males are portrayed on the left, while sulfidic males are portrayed on the right. (A) arcsine-square-root transformed percent motility (MOT), which describes the percentage of tracked sperm that were moving. (B) straight-line velocity (VSL), which measures the averaged velocity along a straight path (i.e., the displacement) from the beginning to end of sperm tracking. This provides information regarding velocity and path straightness. We did not detect significant differences between ecotypes or food treatments in either MOT or VSL.

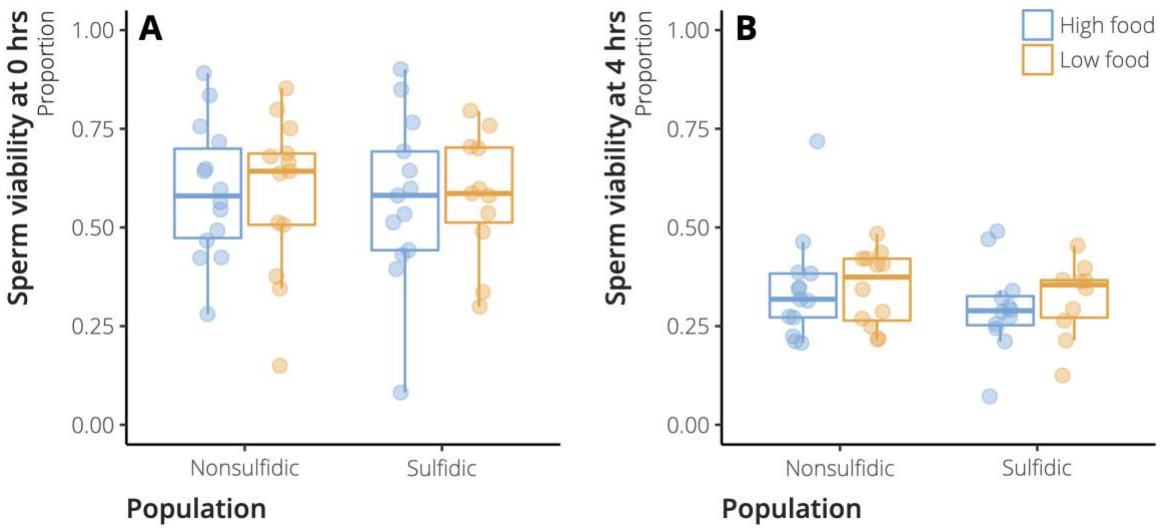

Figure C-3: Boxplots of sperm viability (A) immediately after sperm stripping and activation and (B) 4 hr later. We found that viability decreased significantly from 0 hr to 4 hr after activation, but we did not find evidence for effects of ecotype or food treatment at either time point.

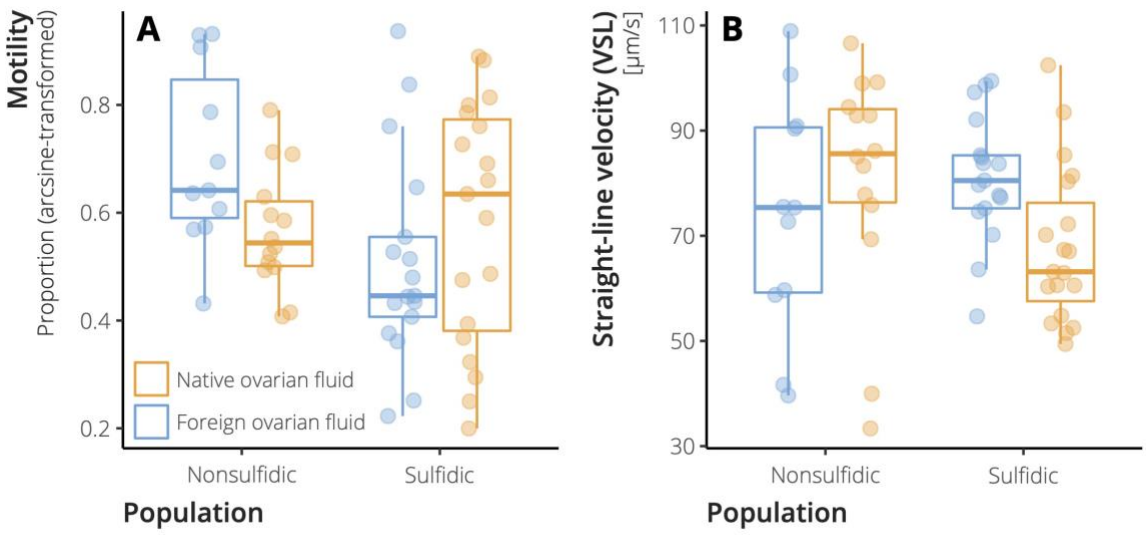

Figure C-4: Boxplots measuring sperm swimming performance in foreign and native ovarian fluid. (A) The effect of ovarian fluid origin (foreign or native) on arcsine-square-root transformed percent motility (MOT), compared between male origins (sulfidic or nonsulfidic). (B) The effect of ovarian fluid origin (foreign or native) on straight-line velocity (VSL), compared between male origins (sulfidic or nonsulfidic). We did not find evidence for effects of male origin, female origin, or their interaction for either VSL or MOT.

Appendix C Tables

Table C-1: Table of input parameters for the ImageJ CASA plugin. We used the same input parameters for each run of CASA, which was done on two 0.5 s intervals for each video, one from 0.5-1.0 s and another from 10.5-11.0 s. We generally used the defaults for the program, but manually validated the input values for minimum and maximum sperm size and maximum sperm velocity between frames, which are species-specific. The technical attributes of the camera we used also caused us to change the frame rate and microns per 1000 pixels.

Parameter	Input
Minimum sperm size (pixels)	1
Maximum sperm size (pixels)	80
Minimum track length (frames)	30
Maximum sperm velocity between frames (pixels)	12
Minimum VSL for motile (μm/s)	3
Minimum VAP for motile (μm/s)	20
Minimum VCL for motile (μm/s)	25
Low VAP speed (μm/s)	5
Maximum percentage of path with zero VAP	1
Maximum percentage of path with low VAP	25
Low VAP speed 2 (μm/s)	25
Low VCL speed (μm/s)	35
High WOB (percent VAP/VCL)	80
High LIN (percent VSL/VAP)	80
High WOB 2 (percent VAP/VCL)	50
High LIN 2 (percent VSL/VAP)	60
Frame rate (frames per second)	60
Microns per 1000 pixels	861
Print xy coordinates for all tracked sperm?	0
Print motion characteristics for all motile sperm?	1
Print median values for motion characteristics?	0

Table C-2: Full model selection tables for each dependent variable. Models were weighted by AICc, and the top models with a cumulative Akaike weight of 0.95 were averaged prior to analysis. Continuous variables that are included in models are denoted with the regression coefficient for that model term, and categorical variables that were included are denoted with a '+'. Terms that were not included in a model are denoted with 'NA'. Due to large size, this table is provided in a Microsoft Excel spreadsheet called "JohnCoffin2022_Appendix.xlsx", under the tab "Table C-2".

References

Aguilar, J., and M. Reyley. 2005. The uterine tubal fluid: secretion, composition and biological effects. Animal Reproduction 2:91–105.

Altmann, J., and S. C. Alberts. 2005. Growth rates in a wild primate population: ecological influences and maternal effects. Behav Ecol Sociobiol 57:490–501.

Aluru, N., and M. M. Vijayan. 2009. Stress transcriptomics in fish: A role for genomic cortisol signaling. General and Comparative Endocrinology 164:142–150.

Andreyev, A. Y., Y. E. Kushnareva, and A. A. Starkov. 2005. Mitochondrial metabolism of reactive oxygen species. 70:15.

Bartón, K. 2020. MuMIn: Multi-Model Inference.

Barts, N. 2020. Comparative physiology and the predictability of evolution in extreme environments. Kansas State University, Manhattan.

Barts, N., R. Greenway, C. Passow, L. Arias-Rodriguez, J. Kelley, and M. Tobler. 2018. Molecular evolution and expression of oxygen transport genes in livebearing fishes (Poeciliidae) from hydrogen sulfide rich springs. Genome 61:273–286.

Basha, P. S., and A. U. Rani. 2003. Cadmium-induced antioxidant defense mechanism in freshwater teleost *Oreochromis mossambicus* (Tilapia). Ecotoxicol. Environ. Saf. 56:218–221.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Baxter, I. 2015. Should we treat the ionome as a combination of individual elements, or should we be deriving novel combined traits? J Exp Bot 66:2127–2131.

Beckerman, A. P., T. G. Benton, C. T. Lapsley, and N. Koesters. 2006. How effective are maternal effects at having effects? Proc Biol Sci 273:485–493.

Belk, M. C., S. J. Ingley, and J. B. Johnson. 2020. Life History Divergence in Livebearing Fishes in Response to Predation: Is There a Microevolution to Macroevolution Barrier? Diversity 12:179. Multidisciplinary Digital Publishing Institute.

Belyaeva, E. A., T. V. Sokolova, L. V. Emelyanova, and I. O. Zakharova. 2012. Mitochondrial Electron Transport Chain in Heavy Metal-Induced Neurotoxicity: Effects of Cadmium, Mercury, and Copper. The Scientific World Journal 2012:1–14.

Benjamini, Y., and Y. Hochberg. 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. [Royal Statistical Society, Wiley].

Ben-Shachar, M., D. Makowski, and D. Lüdecke. 2020. effectsize: Estimation of effect size indices and standardized parameters. Journal of Open Source Software 5:2815.

Berg, S., N. Koeniger, G. Koeniger, and S. Fuchs. 1997. Body size and reproductive success of drones (*Apis mellifera*). Apidologie 28:449–460.

Bierbach, D., S. Oster, J. Jourdan, L. Arias-Rodriguez, J. Krause, A. D. M. Wilson, and M. Plath. 2014. Social network analysis resolves temporal dynamics of male dominance relationships. Behav Ecol Sociobiol 68:935–945.

Billard, R., and M. Cosson. 1992. Some problems related to the assessment of sperm motility in freshwater fish. Journal of Experimental Zoology 261:122–121.

Birben, E., U. M. Sahiner, C. Sackesen, S. Erzurum, and O. Kalayci. 2012. Oxidative stress and antioxidant defense. World Allergy Organ J 5:9–19.

Birkhead, T., and J.-P. Brillard. 2007. Reproductive isolation in birds: postcopulatory prezygotic barriers. Trends in Ecology & Evolution 22:266–272.

Birkhead, T., J. Martinez, T. Burke, and D. Froman. 1999. Sperm mobility determines the outcome of sperm competition in the domestic fowl. Proceedings of the Royal Society B 266:1759–1764.

Birkhead, T., and A. Møller. 1998. Sperm Competition and Sexual Selection. Academic Press, San Diego.

Birkhead, T., and T. Pizzari. 2002. Postcopulatory sexual selection. Nature Reviews Genetics 3:262–273.

Bomblies, K. 2010. Doomed Lovers: Mechanisms of Isolation and Incompatibility in Plants. Annual Review of Plant Biology 61:109–124.

Boots, M., and K. E. Roberts. 2012. Maternal effects in disease resistance: poor maternal environment increases offspring resistance to an insect virus. Proceedings of the Royal Society B: Biological Sciences 279:4009–4014. Royal Society.

Boschetto, C., C. Gasparini, and A. Pilastro. 2011. Sperm number and velocity affect sperm competition success in the guppy (*Poecilia reticulata*). Behavioral Ecology and Sociobiology 65:813–821.

Branson, B. 1967. Fishes of the Neosho River system in Oklahoma. The American Midland Naturalist 78:126–154.

Brown, A. P., L. Arias-Rodriguez, M.-C. Yee, M. Tobler, and J. L. Kelley. 2018. Concordant Changes in Gene Expression and Nucleotides Underlie Independent Adaptation to Hydrogen-Sulfide-Rich Environments. Genome Biology and Evolution 10:2867–2881.

Brown, C., F. Jones, and V. Braithwaite. 2005. In situ examination of boldness-shyness traits in the tropical poeciliid, <Brachyraphis episcopi</i>. Animal Behaviour 70:1003–1009.

Burnham, K., and D. Anderson. 2002. Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach. Springer, New York.

Butler, I., M. Schoonen, and D. Rickard. 1994. Removal of dissolved oxygen from water: A comparison of four common techniques. Talanta 41:211–215.

Camarillo, H., L. Arias Rodriguez, and M. Tobler. 2020. Functional consequences of phenotypic variation between locally adapted populations: Swimming performance and ventilation in extremophile fish. Journal of Evolutionary Biology 33:512–523.

Cardozo, G., A. Devigili, P. Antonelli, and A. Pilastro. 2020. Female sperm storage mediates post-copulatory costs and benefits of ejaculate anticipatory plasticity in the guppy. Journal of Evolutionary Biology 33:1294–1305.

Cattelan, S., A. Di Nisio, and A. Pilastro. 2018. Stabilizing selection on sperm number revealed by artificial selection and experimental evolution. Evolution 72:698–706.

Chambers, J. 1987. The cyprinodontiform gonopodium, with an atlas of the gonopodia of the fishes of the genus Limia. Journal of Fish Biology 30:389–418.

Chapin, F. S., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper, S. Lavorel, O. E. Sala, S. E. Hobbie, M. C. Mack, and S. Díaz. 2000. Consequences of changing biodiversity. Nature 405:234–242.

Chen, J. W., C. Dodia, S. I. Feinstein, M. K. Jain, and A. B. Fisher. 2000. 1-Cys peroxiredoxin, a bifunctional enzyme with glutathione peroxidase and phospholipase A2 activities. J. Biol. Chem. 275:28421–28427.

Chen, K., and J. Morris. 1972. Kinetics of oxidation of aqueous sulfide by O2. Environmental Science & Technology 6:529–537.

Chen, Y., N. Kotian, G. Aranjuez, L. Chen, C. L. Messer, A. Burtscher, K. Sawant, D. Ramel, X. Wang, and J. A. McDonald. 2020. Protein phosphatase 1 activity controls a balance

between collective and single cell modes of migration. eLife 9:e52979. eLife Sciences Publications, Ltd.

Cline, J., and F. Richards. 1969. Oxygenation of hydrogen sulfide in seawater at constant salinity, temperature, and pH. Environmental Science & Technology 3:838–843.

Conrad, J. L., K. L. Weinersmith, T. Brodin, J. B. Saltz, and A. Sih. 2011. Behavioural syndromes in fishes: a review with implications for ecology and fisheries management. Journal of Fish Biology 78:395–435.

Cooke, M. S., M. D. Evans, M. Dizdaroglu, and J. Lunec. 2003. Oxidative DNA damage: mechanisms, mutation, and disease. The FASEB Journal 17:1195–1214. Federation of American Societies for Experimental Biology.

Cooper, C., and G. Brown. 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.

Corona, M., and G. E. Robinson. 2006. Genes of the antioxidant system of the honey bee: annotation and phylogeny. Insect Mol Biol 15:687–701.

Craig, P. M., C. Hogstrand, C. M. Wood, and G. B. McClelland. 2009. Gene expression endpoints following chronic waterborne copper exposure in a genomic model organism, the zebrafish, Danio rerio. Physiological Genomics 40:23–33.

Cramer, E., M. Alund, S. McFarlane, A. Johnsen, and A. Qvarnström. 2016a. Females discriminate against heterospecific sperm in a natural hybrid zone. Evolution 70:1844–1855.

Cramer, E., E. Stensrud, G. Marthinsen, S. Hogner, L. Johannessen, T. Laskemoen, M. Eybert, T. Slagsvold, J. Lifjeld, and A. Johnsen. 2016b. Sperm performance in conspecific and heterospecific female fluid. Ecology and Evolution 6:1363–1377.

Cremeans, M., J. F. Devlin, T. Osorno, and R. Nairn. 2019. Assessment of Bed Hydraulics and Metal Loadings in a Passive Vertical Flow Bioreactor in Commerce, Oklahoma. Groundwater Monitoring & Remediation 39.

Darwin, C. 1859. On the Origin of Species by Means of Natural Selection. John Murray, London.

Datin, D. L., and D. A. Cates. 2002. Sampling and Metal Analysis of Chat Piles in The Tar Creek Superfund Site. Oklahoma Department of Environmental Quality.

Davies, M. J. 2012. Oxidative Damage to Proteins. P. *in* Encyclopedia of Radicals in Chemistry, Biology and Materials. American Cancer Society.

den Boer, S., J. Boomsma, and B. Baer. 2009. Honey bee males and queens use glandular secretions to enhance sperm viability before and after storage. Journal of Insect Physiology 55:538–543.

Devigili, A., J. Fitzpatrick, C. Gasparini, I. Ramnarine, A. Pilastro, and J. Evans. 2018. Possible glimpses into early speciation: The effect of ovarian fluid on sperm velocity accords with post-copulatory isolation between two guppy populations. Journal of Evolutionary Biology 31:66–74.

Dickerson, B. R., T. P. Quinn, and M. F. Willson. 2002. Body size, arrival date, and reproductive success of pink salmon, Oncorhynchus gorbuscha. Ethology Ecology & Evolution 14:29–44. Taylor & Francis.

Domencini, P., and R. Blake. 1997. The kinematics and performance of fish fast-start swimming. Journal of Experimental Biology 200:1165–1178.

Downing, J., Y. Rochon, M. Pérusse, and H. Harvey. 1993. Spatial aggregation, body size, and reproductive success in the freshwater mussel *Elliptio complanata*. Journal of the North American Benthological Society 12:148–156.

Eaton, R., R. Bombardieri, and D. Meyer. 1977. The Mauthner-initiated startle response in teleost fish. Journal of Experimental Biology 66:65–81.

Eberhard, W. 1996. Female Control: Sexual Selection by Cryptic Female Choice. Princeton University Press, Princeton.

Eberhard, W. 1985. Sexual Selection and Animal Genitalia. Harvard University Press, Cambridge.

Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10:48.

Endler, J. 1995. Multiple-trait coevolution and environmental gradients in guppies. Trends in Ecology & Evolution 10:22–29.

Endler, J. 1980. Natural selection on color patterns in *Poecilia reticulata*. Evolution 34:76–91.

Endler, J., and A. Houde. 1995. Geographic-variation in female preferences for male traits in Poecilia-reticulata. Evolution 49:456–468.

Escher, B., H. Stapleton, and E. Schymanski. 2020. Tracking complex mixtures of chemicals in our changing environment. Science 367:388–392.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620.

Evans, J. P., M. Pierotti, and A. Pilastro. 2003. Male mating behavior and ejaculate expenditure under sperm competition risk in the eastern mosquitofish. Behavioral Ecology 14:268–273.

Evans, J., A. Pilastro, and I. Schlupp. 2011. Ecology and Evolution of Poeciliid Fishes. The University of Chicago Press, Chicago.

Evans, L. 1967. The toxicity of hydrogen sulphide and other sulphides. Quarterly Journal of Experimental Physiology 52:231–248.

Falkowski, P., R. J. Scholes, E. Boyle, J. Canadell, D. Canfield, J. Elser, N. Gruber, K. Hibbard, P. Högberg, S. Linder, F. T. Mackenzie, B. M. Iii, T. Pedersen, Y. Rosenthal, S. Seitzinger, V. Smetacek, and W. Steffen. 2000. The Global Carbon Cycle: A Test of Our Knowledge of Earth as a System. Science 290:291–296. American Association for the Advancement of Science.

Felmy, A., D. Reznick, J. Travis, T. Potter, and T. Coulson. 2022. Life histories as mosaics: Plastic and genetic components differ among traits that underpin life-history strategies. Evolution 76:585–604.

Fenton, H., and H. Jackson. 1899. I. The oxidation of polyhydric alcohols in presence of iron. Journal of the Chemical Society 75:1–11.

Fernandez-Marcos, P. J., and S. Nóbrega-Pereira. 2016. NADPH: new oxygen for the ROS theory of aging. Oncotarget 7.

Fitzpatrick, J., and J. Evans. 2014. Postcopulatory inbreeding avoidance in guppies. Journal of Evolutionary Biology 27:2585–2594.

Fitzpatrick, J., C. Willis, A. Devigili, A. Young, M. Carroll, H. Hunter, and D. Brison. 2020. Chemical signals from eggs facilitate cryptic female choice in humans. Proceedings of the Royal Society B 287:20200805.

Fox, J., M. Friendly, and G. Monette. 2018. heplots: Visualizing tests in multivariate linear models.

Fox, J., and S. Weisberg. 2019a. An R Companion to Applied Regression. 3rd ed. SAGE Publications, Inc., Thousand Oaks, CA.

Fox, J., and S. Weisberg. 2018. Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. Journal of Statistical Software 87:1–27.

Fox, R., and S. Weisberg. 2019b. An R Companion to Applied Regression.

Franssen, C. M. 2009. The Effects of Heavy Metal Mine Drainage on Population Size Structure, Reproduction, and Condition of Western Mosquitofish, Gambusia affinis. Arch Environ Contam Toxicol 57:145–156.

Franssen, C. M., M. Brooks, R. Parham, K. Sutherland, and W. Matthews. 2006. Small-bodied Fishes of Tar Creek and Other Small Streams in Ottawa County, Oklahoma. Proceedings of the Oklahoma Academy of Science 86:9–16.

Franssen, C., M. Tobler, R. Riesch, F. García de León, R. Tiedemann, I. Schlupp, and M. Plath. 2008. Sperm production in an extremophile fish, the cave molly (*Poecilia mexicana*, Poeciliidae, Teleostei). Aquatic Ecology 42:685–692.

Freda, P. J., Z. M. Ali, N. Heter, G. J. Ragland, and T. J. Morgan. 2019. Stage-specific genotype-by-environment interactions for cold and heat hardiness in Drosophila melanogaster. Heredity (Edinb) 123:479–491.

Friard, O., and M. Gamba. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution 7:1325–1330.

Gall, J. E., R. S. Boyd, and N. Rajakaruna. 2015. Transfer of heavy metals through terrestrial food webs: a review. Environ Monit Assess 187:201.

García-González, F., and L. Simmons. 2005. Sperm viability matters in insect sperm competition. Current Biology 15:271–275.

Garland, T., C. J. Downs, and A. R. Ives. 2022. Trade-Offs (and Constraints) in Organismal Biology. Physiological and Biochemical Zoology 95:82–112. The University of Chicago Press.

Gasparini, C., R. Dosselli, and J. Evans. 2017. Sperm storage by males causes changes in sperm phenotype and influences the reproductive fitness of males and their sons: FITNESS EFFECTS OF MALE SPERM STORAGE. Evolution Letters 1:16–25.

Gasparini, C., and J. Evans. 2018. Female control over multiple matings increases the opportunity for postcopulatory sexual selection. Proceedings of the Royal Society B 285:20181505.

Gasparini, C., and J. Evans. 2013. Ovarian fluid mediates the temporal decline in sperm viability in a fish with sperm storage. PLoS ONE 8:e64431.

Gasparini, C., J. Kelley, and J. Evans. 2014. Male sperm storage compromises sperm motility in guppies. Biol. Lett. 10:20140681.

Gasparini, C., I. Marino, C. Boschetto, and A. Pilastro. 2010. Effect of male age on sperm traits and sperm competition success in the guppy (*Poecilia reticulata*). Journal of Evolutionary Biology 23:124–135.

Gasparini, C., and A. Pilastro. 2011. Cryptic female preference for genetically unrelated males is mediated by ovarian fluid in the guppy. Proceedings of the Royal Society B: Biological Sciences 278:2495–2501.

Gasparini, C., A. Pilastro, and J. Evans. 2020. The role of female reproductive fluid in sperm competition. Philosophical Transactions of the Royal Society B 375:2020077.

Ghalambor, C., K. Hoke, E. Ruell, E. Fischer, D. Reznick, and K. Hughes. 2015. Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. Nature 525:372–375.

Ghalambor, C. K., D. N. Reznick, and J. A. Walker. 2004. Constraints on Adaptive Evolution: The Functional Trade-Off between Reproduction and Fast-Start Swimming Performance in the Trinidadian Guppy (Poecilia reticulata). The American Naturalist 164:38–50. The University of Chicago Press.

Gienapp, P., C. Teplitsky, J. S. Alho, J. A. Mills, and J. Merilä. 2008. Climate change and evolution: disentangling environmental and genetic responses. Molecular Ecology 17:167–178.

Goffredi, S. K., J. J. Childress, N. T. Desaulniers, and F. J. Lallier. 1997. Sulfide acquisition by the vent worm Riftia pachyptila appears to be via uptake of HS-, rather than H_2S. Journal of Experimental Biology 200:2609–2616. The Company of Biologists Ltd.

Gow, J. L., C. L. Peichel, and E. B. Taylor. 2007. Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. J Evol Biol 20:2173–2180.

Grafen, A. 1988. On the uses of data on lifetime reproductive success. Pp. 454–471 *in* T. Clutton-Brock, ed. Reproductive Success. University of Chicago Press, Chicago, IL.

Graham, A., P. Lavretsky, V. Muñoz-Fuentes, A. Green, R. Wilson, and K. McCracken. 2018. Migration-selection balance drives genetic differentiation in genes associated with high-altitude function in the Speckled Teal (*Anas flavirostris*) in the Andes. Genome Biology and Evolution 10:14–32.

Grant, B., and P. Grant. 1993. Evolution of Darwin's finches caused by a rare climatic event. Proceedings of the Royal Society B 251:111–117.

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

Greenway, R., N. Barts, C. Henpita, A. Brown, L. Arias-Rodriguez, C. Peña, S. Arndt, G. Lau, M. Murphy, L. Wu, D. Lin, J. Shaw, J. Kelley, and M. Tobler. 2020. Convergent evolution of conserved mitochondrial pathways underlies repeated adaptation to extreme environments. Proceedings of the National Academy of Sciences of the United States of America 117:16424–16430. Cold Spring Harbor Laboratory.

Greenway, R., S. Drexler, L. Arias-Rodriguez, and M. Tobler. 2016. Adaptive, but not condition-dependent, body shape differences contribute to assortative mating preferences during ecological speciation. Evolution 70:2809–2822.

Greenway, R., R. McNemee, A. Okamoto, M. Plath, L. Arias-Rodriguez, and M. Tobler. 2019. Correlated divergence of female and male genitalia in replicated lineages with ongoing ecological speciation. Evolution 73:1200–1212.

Greven, H. 2011. Gonads, genitals, and reproductive biology. Pp. 3–17 *in* Ecology and Evolution of Poeciliid Fishes. The University of Chicago Press, Chicago.

Greven, H. 2005. Structural and behavioral traits associated with sperm transfer in Poeciliinae. Pp. 145–163 *in* M. Uribe and H. Grier, eds. Viviparous Fishes. New Life Publications, Homestead, FL.

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

Gross, M. 1996. Alternative reproductive strategies and tactics: diversity within sexes. Trends in Ecology & Evolution 11:92–98.

Hansen, P., A. Lunde, and J. Nason. 2011. The model confidence set. Econometrica 79:453–497.

Harmon, S. M. 2009. Effects of Pollution on Freshwater Organisms. Water Environment Research 81:2030–2069.

Hedrick, T. 2008. Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems. Bioinspiration and Biomimetics 3:034001.

Hegyi, G., J. Török, L. Tóth, L. Z. Garamszegi, and B. Rosivall. 2006. Rapid temporal change in the expression and age-related information content of a sexually selected trait. Journal of Evolutionary Biology 19:228–238.

Hildebrandt, T. M., and M. K. Grieshaber. 2008. Three enzymatic activities catalyze the oxidation of sulfide to thiosulfate in mammalian and invertebrate mitochondria. The FEBS Journal 275:3352–3361.

Hoffberg, S. L., N. J. Troendle, T. C. Glenn, O. Mahmud, S. Louha, D. Chalopin, J. L. Bennetzen, and R. Mauricio. 2018. A High-Quality Reference Genome for the Invasive Mosquitofish *Gambusia affinis* Using a Chicago Library. G3 8:1855–1861.

Hogstrand, C. 2011. Zinc. P. *in* C. Wood, A. Farrell, and C. Brauner, eds. Homeostasis and Toxicology of Essential Metals. Elsevier, London, UK.

Houde, A. 1997. Sex, Color, and Mate Choice in Guppies. Princeton University Press, Princeton.

Huang, P., K. E. Sieving, and C. M. St. Mary. 2012. Heterospecific information about predation risk influences exploratory behavior. Behavioral Ecology 23:463–472.

Hüttemann, M., P. Pecina, M. Rainbolt, T. H. Sanderson, V. E. Kagan, L. Samavati, J. W. Doan, and I. Lee. 2011. The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: from respiration to apoptosis. Mitochondrion 11:369–381.

Jacquemyn, H., H. D. Kort, A. V. Broeck, and R. Brys. 2018. Immigrant and extrinsic hybrid seed inviability contribute to reproductive isolation between forest and dune ecotypes of Epipactis helleborine (Orchidaceae). Oikos 127:73–84.

Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405.

Kaiser, H. 1960. The application of electronic-computers to factor-analysis. Educational and Psychological Measurement 20:141–151.

Karvonen, A., and O. Seehausen. 2012. The role of parasitism in adaptive radiations-When might parasites promote and when might they constrain ecological speciation? International Journal of Ecology 2012:1–20.

Katakura, H., M. Shioi, and Y. Kira. 1989. Reproductive isolation by host specificity in a pair of phytophagous ladybird beetles. Evolution 43:1045–1053.

Kaun, K., C. Riedl, M. Chakaborty-Chatterjee, A. Belay, S. Douglas, A. Gibbs, and M. Sokolowski. 2007. Natural variation in food acquisition mediated by a *Drosophila* cGMP-dependent protein kinase. Journal of Experimental Biology 210:3547–3558.

Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. Ecol Letters 7:1225–1241.

Kelley, J., L. Arias-Rodriguez, D. Patacsil Martin, M. Yee, C. Bustamante, and M. Tobler. 2016. Mechanisms Underlying Adaptation to Life in Hydrogen Sulfide–Rich Environments. Molecular Biology and Evolution 33:1419–1434.

Kenderešová, L., A. Staňová, J. Pavlovkin, E. Ďurišová, M. Nadubinská, M. Čiamporová, and M. Ovečka. 2012. Early Zn2+-induced effects on membrane potential account for primary heavy metal susceptibility in tolerant and sensitive Arabidopsis species. Ann Bot 110:445–459. Oxford Academic.

Korneliussen, T., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: Analysis of next generation sequencing data. BMC Bioinformatics 15:356.

Krueger, F. 2014. Trim Galore! Babraham Bioinformatics.

Lahner, B., J. Gong, M. Mahmoudian, E. L. Smith, K. B. Abid, E. E. Rogers, M. L. Guerinot, J. F. Harper, J. M. Ward, L. McIntyre, J. I. Schroeder, and D. E. Salt. 2003. Genomic scale profiling of nutrient and trace elements in Arabidopsis thaliana. Nature Biotechnology 21:1215–1221. Nature Publishing Group.

Lande, R., and S. Arnold. 1983. The Measurement of Selection on Correlated Characters. Evolution 37:1210–1226.

Langerhans, R. 2007. Evolutionary consequences of predation: avoidance, escape, reproduction, and diversification. Pp. 177–220 *in* A. Elewa, ed. Predation in Organisms, a Distinct Phenomenon. Springer, Berlin.

Langerhans, R. B. 2009. Trade-off between steady and unsteady swimming underlies predator-driven divergence in Gambusia affinis. Journal of Evolutionary Biology 22:1057–1075.

Langerhans, R., C. Layman, A. Shokrollahi, and T. DeWitt. 2004. Predator-driven phenotypic diversification in *Gambusia affinis*. Evolution 58:2305–2318.

Langfelder, P., and S. Horvath. 2007. Eigengene networks for studying the relationships between co-expression modules. BMC Systems Biology 1:54.

Langfelder, P., and S. Horvath. 2012. Fast R Functions for Robust Correlations and Hierarchical Clustering. J Stat Softw 46.

Langfelder, P., and S. Horvath. 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559.

Langfelder, P., B. Zhang, and S. Horvath. 2016. dynamicTreeCut: Methods for detection of clusters in hierarchical clustering dendrograms.

Laske, S., A. Rosenberger, M. Wipfli, and C. Zimmerman. 2018. Generalist feeding strategies in Arctic freshwater fish: A mechanism for dealing with extreme environments. Ecology of Freshwater Fish 27:767–784.

Lenihan, H. S., C. H. Peterson, J. E. Byers, J. H. Grabowski, G. W. Thayer, and D. R. Colby. 2001. Cascading of Habitat Degradation: Oyster Reefs Invaded by Refugee Fishes Escaping Stress. Ecological Applications 11:764–782.

Li, H. 2013a. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Li, H. 2013b. Toolkit for processing sequences in FASTA/Q formats.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. Oxford Academic.

Liberti, J., B. Baer, and J. Boomsma. 2016. Queen reproductive tract secretions enhance sperm motility in ants. Biology Letters 12:20160722.

Lindholm, A. K., J. Hunt, and R. Brooks. 2006. Where do all the maternal effects go? Variation in offspring body size through ontogeny in the live-bearing fish *Poecilia parae*. Biol. Lett. 2:586–589.

Locatello, L., A. Pilastro, R. Deana, A. Zarpellon, and M. Rasotto. 2007. Variation pattern of sperm quality traits in two gobies with alternative mating tactics. Functional Ecology 21:975–981.

Lopez, G. A., B. M. Potts, and P. A. Tilyard. 2000. F1 hybrid inviability in Eucalyptus: the case of E. ovata × E. globulus. Heredity 85:242–250. Nature Publishing Group.

López-Sepulcre, A., S. P. Gordon, I. G. Paterson, P. Bentzen, and D. N. Reznick. 2013. Beyond lifetime reproductive success: the posthumous reproductive dynamics of male Trinidadian guppies. Proceedings of the Royal Society B: Biological Sciences 280:20131116. Royal Society.

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.

Lukas, J., F. Auer, T. Goldhammer, J. Krause, P. Romanczuk, P. Klamser, L. Arias-Rodriguez, and D. Bierbach. 2021. Diurnal Changes in Hypoxia Shape Predator-Prey Interaction in a Bird-Fish System. Frontiers in Ecology and Evolution 9.

MacLeod, K. J., G. M. While, and T. Uller. 2021. Viviparous mothers impose stronger glucocorticoid-mediated maternal stress effects on their offspring than oviparous mothers. Ecology and Evolution 11:17238–17259.

Madsen, P. B., A. Morabowen, P. Andino, R. Espinosa, S. Cauvy-Fraunié, O. Dangles, and D. Jacobsen. 2015. Altitudinal distribution limits of aquatic macroinvertebrates: an experimental test in a tropical alpine stream. Ecological Entomology 40:629–638.

Mager, E. 2011. Lead. Pp. 186–225 *in* C. Wood, A. Farrell, and C. Brauner, eds. Homeostasis and Toxicology of Non-Essential Metals. Elsevier, London, UK.

Magurran, A. 2011. Sexual coercion. Pp. 209–217 *in* J. Evans, A. Pilastro, and I. Schlupp, eds. Ecology and Evolution of Poeciliid Fishes. The University of Chicago Press, Chicago.

Marshall, D., and T. Uller. 2007. When is a maternal effect adaptive? Oikos 116:1957–1963.

McGeer, J., S. Niyogi, and S. Smith. 2011. Cadmium. Pp. 126–168 *in* C. Wood, A. Farrell, and C. Brauner, eds. Homeostasis and Toxicology of Non-Essential Metals. Elsevier, London, UK.

Mendel, G. 1865. Versuche über planzen-hybriden. Verhandlungen des naturforschenden Vereines in Brünn Bd.4:3–47.

Meyer, J. N., M. C. K. Leung, J. P. Rooney, A. Sendoel, M. O. Hengartner, G. E. Kisby, and A. S. Bess. 2013. Mitochondria as a Target of Environmental Toxicants. Toxicol Sci 134:1–17.

Monaghan, P. 2008. Early growth conditions, phenotypic development and environmental change. Phil. Trans. R. Soc. B 363:1635–1645.

Moraes, C. R., and S. Meyers. 2018. The sperm mitochondrion: Organelle of many functions. Animal Reproduction Science 194:71–80.

Mousseau, T., and C. Fox (eds). 1998. Maternal effects as adaptations. Oxford University Press, New York.

Naisbit, R., C. Jiggins, and J. Mallet. 2001. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. Proceedings of the Royal Society B 268:1849–1854.

Nemmiche, S. 2016. Oxidative Signaling Response to Cadmium Exposure. Toxicol. Sci. kfw222.

Neuberger, J. S., S. C. Hu, K. D. Drake, and R. Jim. 2009. Potential health impacts of heavy-metal exposure at the Tar Creek Superfund site, Ottawa County, Oklahoma. Environ Geochem Health 13.

Nosil, P. 2012. Ecological Speciation. Oxford University Press, New York.

Nosil, P. 2004. Reproductive isolation caused by visual predation on migrants between divergent environments. Proceedings of the Royal Society B: Biological Sciences 271:1521–1528.

Nosil, P., B. Crespi, and C. Sandoval. 2002. Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 417:440–443.

Nosil, P., S. P. Egan, and D. J. Funk. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: “isolation by adaptation” and multiple roles for divergent selection. Evolution 62:316–336.

Nosil, P., T. Vines, and D. Funk. 2005. Perspective: Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution 59:705–719.

Olson, K., J. Donald, R. Dombkowski, and S. Perry. 2012. Evolutionary and comparative aspects of nitric oxide, carbon monoxide and hydrogen sulfide. Respiratory Physiology & Neurobiology 184:117–129.

O’Neill, J. G. 1981. The humoral immune response of Salmo trutta L. and Cyprinus carpio L. exposed to heavy metals. Journal of Fish Biology 19:297–306.

OWRB. 1983. Water quality characteristics of seepage and runoff at two tailings piles in the Picher Field, Ottawa County, Oklahoma. Oklahoma Water Resources Board.

Oziolor, E., and C. Matson. 2015. Evolutionary Toxicology: Population Adaptation in Response to Anthropogenic Pollution. Pp. 247–278 *in* R. Riesch, M. Tobler, and M. Plath, eds. Extremophile Fishes: Ecology, Evolution, and Physiology of Teleosts in Extreme Environments. Springer International Publishing, Switzerland.

Pacyna, J. M., and E. G. Pacyna. 2001. An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide. Environ. Rev. 9:269–298.

Palacios, M., L. Arias-Rodriguez, M. M, C. C, H. Lerp, A. Lamboj, G. Voelker, and M. Tobler. 2013. The Rediscovery of a Long Described Species Reveals Additional Complexity in Speciation Patterns of Poeciliid Fishes in Sulfide Springs. PLOS ONE 8:e71069. Public Library of Science.

Parker, G. 1970. Sperm competition and its evolutionary consequences in the insects. Biological Reviews 45:525–567.

Parker, G. 1990. Sperm competition games: raffles and roles. Proceedings of the Royal Society B 242:120–126.

Passow, C., R. Greenway, L. Arias-Rodriguez, P. Jeyasingh, and M. Tobler. 2015. Reduction of Energetic Demands through Modification of Body Size and Routine Metabolic Rates in Extremophile Fish. Physiological and Biochemical Zoology 88:371–383.

Passow, C., C. Henpita, J. Shaw, C. Quackenbush, W. Warren, M. Schartl, L. Arias-Rodriguez, J. Kelley, and M. Tobler. 2017. The roles of plasticity and evolutionary change in shaping gene expression variation in natural populations of extremophile fish. Molecular Ecology 26:6384–6399.

Peace, A., P. C. Frost, N. D. Wagner, M. Danger, C. Accolla, P. Antczak, B. W. Brooks, D. M. Costello, R. A. Everett, K. B. Flores, C. M. Heggerud, R. Karimi, Y. Kang, Y. Kuang, J. H. Larson, T. Mathews, G. D. Mayer, J. N. Murdock, C. A. Murphy, R. M. Nisbet, L. Pecquerie, N. Pollesch, E. M. Rutter, K. L. Schulz, J. T. Scott, L. Stevenson, and H. Wang. 2021. Stoichiometric Ecotoxicology for a Multisubstance World. BioScience 71:132–147.

Pérez-Escudero, A., J. Vicente-Page, R. Hinz, S. Arganda, and G. de Polavieja. 2014. idTracker: Tracking individuals in a group by automatic identification of unmarked animals. Nature Methods 11:743–748.

Pertea, M., G. M. Pertea, C. M. Antonescu, T.-C. Chang, J. T. Mendell, and S. L. Salzberg. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33:290–295.

Pfenninger, M., H. Lerp, M. Tobler, C. Passow, J. Kelley, E. Funke, B. Greshake, U. Erkoc, T. Berberich, and M. Plath. 2014. Parallel evolution of cox genes in H_2S-tolerant fish as key adaptation to a toxic environment. Nature Communications 5:3873.

Philippi, E. 1908. Fortpflanzungsgeschichte der viviparen Teleosteer *Glaridichthhys januarius* und *G. decem-maculatus* in ihrem Einfluß auf Lebensweise, makroskopische und mikroskopische Anatomie. Zoologische Jahrbücher, Abteilung Anatomie 27:1–94.

Pigliucci, M. 2001. Phenotypic Plasticity: Beyond Nature and Nurture. Johns Hopkins University Press, Baltimore.

Pilastro, A., M. Simonato, A. Bisazza, and J. Evans. 2004. Cryptic female preferences for colorful males in guppies. Evolution 58:665–669.

Pinto, E., T. C. S. Sigaud-kutner, M. A. S. Leitão, O. K. Okamoto, D. Morse, and P. Colepicolo. 2003. Heavy Metal–Induced Oxidative Stress in Algae1. Journal of Phycology 39:1008–1018.

Pinu, F. R., D. J. Beale, A. M. Paten, K. Kouremenos, S. Swarup, H. J. Schirra, and D. Wishart. 2019. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. Metabolites 9:76. Multidisciplinary Digital Publishing Institute.

Plath, M., A. M. Makowicz, I. Schlupp, and M. Tobler. 2007a. Sexual harassment in live-bearing fishes (Poeciliidae): comparing courting and noncourting species. Behavioral Ecology 18:680–688.

Plath, M., J. Parzefall, K. Körner, and I. Schlupp. 2004. Sexual selection in darkness? Female mating preferences in suface- and cave-dwelling Atlantic mollies, *Poecilia mexicana* (Poeciliidae, Teleostei). Behavioral Ecology and Sociobiology 55:596–601.

Plath, M., M. Pfenninger, H. Lerp, R. Riesch, C. Eschenbrenner, P. Slattery, D. Bierbach, N. Herrmann, M. Schulte, L. Arias-Rodriguez, J. Indy, C. Passow, and M. Tobler. 2013. Genetic differentiation and selection against migrants in evolutionarily replicated extreme environments. Evolution 67:2647–2661.

Plath, M., R. Riesch, A. Oranth, J. Dzienko, N. Karau, A. Schiessl, S. Stadler, A. Wigh, C. Zimmer, L. Arias-Rodriguez, I. Schlupp, and M. Tobler. 2010. Complementary effects of

natural and sexual selection against immigrants maintains differentiation between locally adapted fish. Naturwissenschaften 97:769–774.

Plath, M., M. Tobler, R. Riesch, F. García de León, O. Giere, and I. Schlupp. 2007b. Survival in an extreme habitat: the roles of behaviour and energy limitation. Naturwissenschaften 94:991–996.

Poeggeler, B., R. J. Reiter, R. Hardeland, D.-X. Tan, and L. R. Barlow-Walden. 1996. Melatonin and structurally-related, endogenous indoles act as potent electron donors and radical scavengers in vitro. Redox Report 2:179–184. Taylor & Francis.

Purchase, C., and P. Earle. 2012. Modification to the ImageJ computer assisted sperm analysis plugin greatly improve efficiency and fundamentally alter the scope of attainable data. Journal of Applied Ichthyology 28:1013–1016.

Pyke, G. H. 2005. A Review of the Biology of Gambusia affinis and G. holbrooki. Rev Fish Biol Fisheries 15:339–365.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rajkov, J., A. Weber, W. Salzburger, and B. Egger. 2018. Immigrant and extrinsic hybrid inviability contribute to reproductive isolation between lake and river cichlid ecotypes. Evolution 72:2553–2564.

Ransberry, V. E., A. J. Morash, T. A. Blewett, C. M. Wood, and G. B. McClelland. 2015. Oxidative stress and metabolic responses to copper in freshwater- and seawater-acclimated killifish, Fundulus heteroclitus. Aquatic Toxicology 161:242–252.

Regoli, F., and M. E. Giuliani. 2014. Oxidative pathways of chemical toxicity and oxidative stress biomarkers in marine organisms. Marine Environmental Research 93:106–117.

Reiffenstein, R., W. Hulbert, and S. Roth. 1992. Toxicology of hydrogen sulfide. Annual Reviews of Pharmacology and Toxicology 32:109–134.

Reznick, D., and H. Bryga. 1987. Life-history evolution in guppies (*Poecilia reticulata*): 1. Phenotypic and genetic changes in an introduction experiment. Evolution 41:1370–1385.

Reznick, D., H. Bryga, and J. Endler. 1990. Experimentally induced life-history evolution in a natural population. Nature 346:357–359.

Reznick, D., H. Callahan, and R. Llauredo. 1996a. Maternal Effects on Offspring Quality in Poeciliid Fishes. Am Zool 36:147–156.

Reznick, D., and J. A. Endler. 1982. The Impact of Predation on Life History Evolution in Trinidadian Guppies (poecilia Reticulata). Evolution 36:160–177.

Reznick, D., L. Nunney, and A. Tessier. 2000. Big houses, big cars, superfleas and the costs of reproduction. Trends in Ecology & Evolution 15:421–425.

Reznick, D., H. Rodd, and M. Cardenas. 1996b. Life-History Evolution in Guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in Life-History Phenotypes. The American Naturalist 147:319–338.

Riesch, R., V. Duwe, N. Herrmann, L. Padur, A. Ramm, K. Scharnweber, M. Schulte, T. Schulz-Mirbach, M. Ziege, and M. Plath. 2009. Variation along the shy-bold continuum in extremophile fishes (*Poecilia mexicana, Poecilia sulphuraria*). Behavioral Ecology and Sociobiology 63:1515–1526.

Riesch, R., A. Oranth, J. Dzienko, N. Karau, A. Schiessl, S. Stadler, A. Wigh, C. Zimmer, L. Arias-Rodriguez, I. Schlupp, and M. Plath. 2010a. Extreme habitats are not refuges: Poeciliids suffer from increased aerial predation risk in sulphidic southern Mexican habitats. Biological Journal of the Linnean Society 101:417–426.

Riesch, R., M. Plath, and I. Schlupp. 2010b. Toxic hydrogen sulfide and dark caves: Life-history adaptations in a livebearing fish (*Poecilia mexicana*, Poeciliidae). Ecology 91:1494–1505.

Riesch, R., M. Plath, and I. Schlupp. 2011a. Toxic hydrogen sulphide and dark caves: pronounced male life-history divergence among locally adapted *Poecilia mexicana* (Poeciliidae): Extremophile life histories. Journal of Evolutionary Biology 24:596–606.

Riesch, R., M. Plath, I. Schlupp, M. Tobler, and R. Langerhans. 2014. Colonization of toxic springs drives predictable life-history shift in livebearing fishes (Poeciliidae). Ecology Letters 17:65–71.

Riesch, R., D. Reznick, M. Plath, and I. Schlupp. 2016a. Sex-specific local life-history adaptation in surface- and cave-dwelling Atlantic mollies (*Poecilia mexicana*). Scientific Reports 6:22968.

Riesch, R., I. Schlupp, R. B. Langerhans, and M. Plath. 2011b. Shared and Unique Patterns of Embryo Development in Extremophile Poeciliids. PLoS ONE 6:e27377.

Riesch, R., M. Tobler, H. Lerp, J. Jourdan, L. Doumas, P. Nosil, R. Langerhans, and M. Plath. 2016b. Extremophile Poeciliidae: multivariate insights into the complexity of speciation along replicated ecological gradients. BMC Evolutionary Biology 16.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

Rossiter, M. 1996. Incidence and Consequences of Inherited Environmental Effects. Annual Review of Ecology and Systematics 27:451–476.

Rowe, M., L. Veerus, P. Trosvik, A. Buckling, and T. Pizzari. 2020. The Reproductive Microbiome: An Emerging Driver of Sexual Selection, Sexual Conflict, Mating Systems, and Reproductive Isolation. Trends in Ecology & Evolution 35:220–234.

Rudin-Bitterli, T. S., N. J. Mitchell, and J. P. Evans. 2020. Extensive geographical variation in testes size and ejaculate traits in a terrestrial-breeding frog. Biology Letters 16:20200411. Royal Society.

Rudman, S. M., J. M. Goos, J. B. Burant, K. V. Brix, T. C. Gibbons, C. J. Brauner, and P. D. Jeyasingh. 2019. Ionome and elemental transport kinetics shaped by parallel evolution in threespine stickleback. Ecol Lett ele.13225.

Rundle, H. D., and P. Nosil. 2005. Ecological speciation. Ecology Letters 8:336–352.

Sanches, E. A., R. M. Marcos, R. Y. Okawara, D. Caneppele, R. A. Bombardelli, and E. Romagosa. 2013. Sperm motility parameters for Steindachneridion parahybae based on open-source software. Journal of Applied Ichthyology 29:1114–1122.

Sánchez-Bayo, F., K. Goka, and D. Hayasaka. 2016. Contamination of the Aquatic Environment with Neonicotinoids and its Implication for Ecosystems. Front. Environ. Sci. 4. Frontiers.

Sanchez-Dardon, J., I. Voccia, A. Hontela, S. Chilmonczyk, M. Dunier, H. Boermans, B. Blakley, and M. Fournier. 1999. Immunomodulation by heavy metals tested individually or in mixtures in rainbow trout (Oncorhynchus mykiss) exposed in vivo. Environmental Toxicology and Chemistry 18:1492–1497.

Schartl, M., R. B. Walter, Y. Shen, T. Garcia, J. Catchen, A. Amores, I. Braasch, D. Chalopin, J.-N. Volff, K.-P. Lesch, A. Bisazza, P. Minx, L. Hillier, R. K. Wilson, S. Fuerstenberg, J. Boore, S. Searle, J. H. Postlethwait, and W. C. Warren. 2013. The genome of the platyfish, Xiphophorus maculatus, provides insights into evolutionary adaptation and several complex traits. Nature Genetics 45:567–572.

Scheiner, S. M. 2002. Selection experiments and the study of phenotypic plasticity. Journal of Evolutionary Biology 15:889–898.

Schlupp, I., J. Poschadel, M. Tobler, and M. Plath. 2006. Male size polymorphism and testis weight in two species of mollies (*Poecilia latipinna*, *P. mexicana*, Poeciliidae, Teleostei). Zeitschrift für Fischkunde 8:9–16.

Schluter, D. 1996. Ecological speciation in postglacial fishes. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 351:807–814. Royal Society.

Schluter, D. 2001. Ecology and the origin of species. Trends in Ecology & Evolution 16:372–380.

Schneider, C., W. Rasband, and K. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9:671–675.

Schrader, M., and J. Travis. 2012. Variation in Offspring Size with Birth Order in Placental Fish: A Role for Asymmetric Sibling Competition? Evolution 66:272–279.

Scialò, F., A. Sriram, D. Fernández-Ayala, N. Gubina, M. Lõhmus, G. Nelson, A. Logan, H. M. Cooper, P. Navas, J. A. Enríquez, M. P. Murphy, and A. Sanz. 2016. Mitochondrial ROS Produced via Reverse Electron Transport Extend Animal Lifespan. Cell Metabolism 23:725–734.

Scopece, G., A. Widmer, and S. Cozzolino. 2008. Evolution of postzygotic reproductive isolation in a guild of deceptive orchids. The American Naturalist 171:315–326.

Servedio, M., and M. Noor. 2003. The role of reinforcement in speciation: Theory and data. Annual Review of Ecology, Evolution, and Systematics 34:339–364.

Sharon, G., D. Segal, J. Ringo, A. Hefetz, I. Zilber-Rosenberg, and E. Rosenberg. 2010. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. Proceedings of the National Academy of Sciences of the United States of America 107:20051–20056.

Silva, D. C. V. R., C. V. M. Araújo, J. C. López-Doval, M. B. Neto, F. T. Silva, T. C. B. Paiva, and M. L. M. Pompêo. 2017. Potential effects of triclosan on spatial displacement and local population decline of the fish *Poecilia reticulata* using a non-forced system. Chemosphere 184:329–336.

Skotte, L., T. Korneliussen, and A. Albrechtsen. 2013. Estimating Individual Admixture Proportions from Next Generation Sequencing Data. Genetics 195:693–702.

Skulachev, V. P. 1998. Cytochrome c in the apoptotic and antioxidant cascades. FEBS Lett. 423:275–280.

Smith, C. 2011. Sperm competition and the evolution of alternative reproductive tactics in the swordtail *Xiphophorus nigrensis* (Poeciliidae). University of Texas, Austin, TX.

Smith, C. C. 2012. Opposing effects of sperm viability and velocity on the outcome of sperm competition. Behavioral Ecology 23:820–826.

Spagopoulou, F., R. Vega-Trejo, M. Head, and M. Jennions. 2020. Shifts in reproductive investment in response to competitors lower male reproductive success. The American Naturalist 196:355–368.

Stearns, S. 1992. The evolution of life histories. Oxford University Press, New York.

Stearns, S. 1989. Trade-offs in life-history evolution. Functional Ecology 3:259–268.

Stearns, S., and R. Sage. 1980. Maladaptation in a marginal population of the mosquito fish, *Gambusia affinis*. Evolution 34:65–75.

Steinhilber, D., M. Brungs, O. Werz, I. Wiesenberg, C. Danielsson, J. P. Kahlen, S. Nayeri, M. Schräder, and C. Carlberg. 1995. The nuclear receptor for melatonin represses 5-lipoxygenase gene expression in human B lymphocytes. J. Biol. Chem. 270:7037–7040.

Stohs, S. J., and D. Bagchi. 1995. Oxidative mechanisms in the toxicity of metal ions. Free Radic. Biol. Med. 18:321–336.

Sunnucks, P., H. Morales, A. Lamb, A. Pavlova, and C. Greening. 2017. Integrative approaches for studying mitochondrial and nuclear genome co-evolution in oxidative phosphorylation. Frontiers in Genetics 8:1–12.

Symonds, M. R. E., and A. Moussalli. 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav Ecol Sociobiol 65:13–21.

Taborsky, M. 1998. Sperm competition in fish: “bourgeois” males and parasitic spawning. Trends in Ecology & Evolution 13:222–227.

Teska, W. R., M. H. Smith, and J. M. Novak. 1990. Food Quality, Heterozygosity, and Fitness Correlates in Peromyscus Polionotus. Evolution 44:1318–1325.

The Gene Ontology Consortium. 2004. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 32:D258–D261. Oxford Academic.

Thomas, L. 1984. EPA Superfund Record of Decision: Tar Creek/Picher Mine Field, Ottawa County, Oklahoma, and Cherokee County, Kansas.

Tinbergen, N. 1963. On aims and methods of ethology. Zeitschrift für Tierpsychologie 20:410–433.

Tobler, M. 2008a. Divergence in trophic ecology characterizes colonization of extreme environments. Biological Journal of the Linnean Society 95:517–528.

Tobler, M. 2008b. Divergence in trophic ecology characterizes colonization of extreme habitats. Biological Journal of the Linnean Society 95:517–528.

Tobler, M., D. Alba, L. Arias-Rodríguez, and P. Jeyasingh. 2016. Using replicated evolution in extremophile fish to understand diversification in elemental composition and nutrient excretion. Freshw Biol 61:158–171.

Tobler, M., N. Barts, and R. Greenway. 2019. Mitochondria and the origin of species: Bridging genetic and ecological perspectives on speciation processes. Integrative and Comparative Biology 59:900–911.

Tobler, M., T. DeWitt, I. Schlupp, F. García de León, R. Herrmann, P. Feulner, R. Tiedmann, and M. Plath. 2008. Toxic hydrogen sulfide and dark caves: phenotypic and genetic divergence across two abiotic environmental gradients in *Poecilia mexicana*. Evolution 62:2643–2659.

Tobler, M., and L. Hastings. 2011. Convergent patterns of body shape differentiation in four different clades of poeciliid fishes inhabiting sulfide springs. Evolutionary Biology 38:412–421.

Tobler, M., J. Kelley, M. Plath, and R. Riesch. 2018. Extreme environments and the origins of biodiversity: Adaptation and speciation in sulphide spring fishes. Molecular Ecology 27:843–859.

Tobler, M., M. Palacios, L. Chapman, I. Mitrofanov, D. Bierbach, M. Plath, L. Arias-Rodriguez, F. García de León, and M. Mateos. 2011a. Evolution in extreme environments: replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. Evolution 65:2213–2228.

Tobler, M., and M. Plath. 2011. Chapter 11. Living in extreme environments. Pp. 120–127 *in* Ecology and Evolution of Poeciliid Fishes. The University of Chicago Press, Chicago.

Tobler, M., M. Plath, R. Riesch, I. Schlupp, A. Grasse, G. K. Munimanda, C. Setzer, D. J. Penn, and Y. Moodley. 2014. Selection from parasites favours immunogenetic diversity but not divergence among locally adapted host populations. J. Evol. Biol. 27:960–974.

Tobler, M., R. Riesch, C. Tobler, and M. Plath. 2009a. Compensatory behavior in response to sulfide-induced hypoxia affects time budgets, feeding efficiency, and predation risk. Evolutionary Ecology Research 11:935–948.

Tobler, M., R. Riesch, C. Tobler, T. Schulz-Mirbach, and M. Plath. 2009b. Natural and sexual selection against immigrants maintains differentiation among micro-allopatric populations. Journal of Evolutionary Biology 22:2298–2304.

Tobler, M., K. Scharnweber, R. Greenway, C. Passow, L. Arias-Rodriguez, and F. García de León. 2015. Convergent changes in the trophic ecology of extremophile fish along replicated environmental gradients. Freshwater Biology 60:768–780.

Tobler, M., I. Schlupp, F. García de León, M. Glaubrecht, and M. Plath. 2007. Extreme habitats as refuge from parasite infections? Oecologia-International Journal of Ecology 31:270–275.

Tobler, M., I. Schlupp, K. U. Heubel, R. Riesch, F. J. G. de León, O. Giere, and M. Plath. 2006. Life on the edge: hydrogen sulfide and the fish communities of a Mexican cave and surrounding waters. Extremophiles 10:577–585.

Tobler, M., I. Schlupp, and M. Plath. 2011b. Costly interactions between the sexes: Combined effects of male sexual harassment and female choice? Behavioral Ecology 22:723–729.

Tokić, G., and D. K. P. Yue. 2012. Optimal shape and motion of undulatory swimming organisms. Proceedings of the Royal Society B: Biological Sciences 279:3065–3074. Royal Society.

Tomar, S. 2006. Converting video formats with FFmpeg. Linux Journal 146:10.

Torres-Martínez, A., A. Hernández-Franyutti, M. C. Uribe, and W. M. Contreras-Sánchez. 2017. Ovarian structure and oogenesis of the extremophile viviparous teleost Poecilia mexicana (Poeciliidae) from an active sulfur spring cave in Southern Mexico. Journal of Morphology 278:1667–1681.

Turelli, M., N. Barton, and J. Coyne. 2001. Theory and speciation. Trends in Ecology & Evolution 16:330–343.

Uller, T. 2008. Developmental plasticity and the evolution of parental effects. Trends in Ecology & Evolution 23:432–438.

United States Environmental Protection Agency. 2019. Aquatic Life Criteria Table, Appendix B - Parameters for Calculating Freshwater Dissolved Metals Criteria That Are Hardness-Dependent.

United States Environmental Protection Agency. 2016. Remedial Investigation Data Gap Summary Report Version 1.0.

Urabe, J., S. Naeem, D. Raubenheimer, and J. J. Elser. 2010. The evolution of biological stoichiometry under global change. Oikos 119:737–740.

Uren Webster, T. M., N. Bury, R. van Aerle, and E. M. Santos. 2013. Global transcriptome profiling reveals molecular mechanisms of metal tolerance in a chronically exposed wild population of brown trout. Environ. Sci. Technol. 47:8869–8877.

Van Allen, B., N. Jones, B. Gilbert, K. Carscadden, and R. Germain. 2021. Maternal effects and the outcome of interspecific competition. Ecology and Evolution 11:7544–7556.

Venables, W., and B. Ripley. 2002. Modern Applied Statistics with S. 4th ed. Springer, New York.

Via, S., A. Bouck, and S. Skillman. 2000. Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. Evolution 54:1626–1637.

Viale, G., and D. Calamari. 1984. Immune response in rainbow trout Salmo gairdneri after long-term treatment with low levels of Cr, Cd and Cu. Environmental Pollution Series A, Ecological and Biological 35:247–257.

Wagner, G., G. Booth, and H. Bagheri-Chaichian. 1997. A population genetic theory of canalization. Evolution 51:329–347.

Wallace, J., and R. Wang. 2015. Hydrogen-sulfide based therapeutics: Exploiting a unique but ubiquitous gasotransmitter. Nature Reviews Drug Discovery 14:329–345.

Weeks, S., and O. Gaggiotti. 1993. Patterns of offspring size at birth in clonal and sexual strains of *Poeciliopsis* (Poeciliidae). Copeia 4:1005–1009.

Weiss, L. C. 2019. Sensory Ecology of Predator-Induced Phenotypic Plasticity. Frontiers in Behavioral Neuroscience 12.

Whitehead, A., B. W. Clark, N. M. Reid, M. E. Hahn, and D. Nacci. 2017. When evolution is the solution to pollution: Key principles, and lessons from rapid repeated adaptation of killifish (*Fundulus heteroclitus*) populations. Evol Appl 10:762–783.

Wilson, A., D. Coltman, J. Pemberton, D. Overall, K. Byrne, and L. Kruuk. 2004. Maternal genetic effects set the potential for evolution in a free-living vertebrate population: Maternal genetic effects in Soay sheep. Journal of Evolutionary Biology 18:405–414.

Wilson-Leedy, J., and R. Ingermann. 2007. Development of a novel CASA system based on open source software for characterization of zebrafish sperm motility parameters. Theriogenology 67:661–672.

Wineland, S. M., S. M. Welch, T. K. Pauley, J. J. Apodaca, M. Olszack, J. J. Mosher, J. N. Holmes, and J. L. Waldron. 2019. Using environmental DNA and occupancy modelling to identify drivers of eastern hellbender (Cryptobranchus alleganiensis alleganiensis) extirpation. Freshwater Biology 64:208–221.

Winemiller, K. O. 1992. Life-History Strategies and the Effectiveness of Sexual Selection. Oikos 63:318–327. [Nordic Society Oikos, Wiley].

Winemiller, K., and K. Rose. 1992. Patterns of life-history diversification in North American fishes: implications for population regulation. Canadian Journal of Fisheries and Aquatic Sciences 49:2196–2218.

Wolf, J. B., and M. J. Wade. 2009. What are maternal effects (and what are they not)? Philosophical Transactions of the Royal Society B: Biological Sciences 364:1107–1115. Royal Society.

Wolff, J., E. Ladoukakis, J. Enríquez, and D. Dowling. 2014. Mitonuclear interactions: Evolutionary consequences over multiple biological scales. Philosophical Transactions of the Royal Society B 369:20130443.

Woo, S., S. Yum, H.-S. Park, T.-K. Lee, and J.-C. Ryu. 2009. Effects of heavy metals on antioxidants and stress-responsive gene expression in Javanese medaka (Oryzias javanicus). Comp. Biochem. Physiol. C Toxicol. Pharmacol. 149:289–299.

Wood, C. 2011. An introduction to metals in fish physiology and toxicology: basic principles. Pp. 1–51 *in* Fish Physiology. Elsevier, London, UK.

Wood, C., A. Farrell, and C. Brauner (eds). 2011a. Homeostasis and Toxicology of Essential Metals. Elsevier, London, UK.

Wood, C., A. Farrell, and C. Brauner (eds). 2011b. Homeostasis and Toxicology of Non-Essential Metals. Elsevier, London, UK.

Zhang, B., and S. Horvath. 2005. A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical Applications in Genetics and Molecular Biology 4.

Zhang, H., O. Vedder, P. H. Becker, and S. Bouwhuis. 2015. Age-dependent trait variation: the relative contribution of within-individual change, selective appearance and disappearance in a long-lived seabird. Journal of Animal Ecology 84:797–807.