Colonisation of toxic environments drives predictable life-history evolution in livebearing fishes (Poeciliidae)

Rüdiger Riesch,1,2* Martin Plath,3 Ingo Schlupp,4 Michael Tobler5 and R. Brian Langerhans1

Abstract
New World livebearing fishes (family Poeciliidae) have repeatedly colonised toxic, hydrogen sulphide-rich waters across their natural distribution. Physiological considerations and life-history theory predict that these adverse conditions should favour the evolution of larger offspring. Here, we examined nine poeciliid species that independently colonised toxic environments, and show that these fishes have indeed repeatedly evolved much larger offspring size at birth in sulphidic waters, thus uncovering a widespread pattern of predictable evolution. However, a second pattern, only indirectly predicted by theory, proved additionally common: a reduction in the number of offspring carried per clutch (i.e. lower fecundity). Our analyses reveal that this secondary pattern represents a mere consequence of a classic life-history trade-off combined with strong selection on offspring size alone. With such strong natural selection in extreme environments, extremophile organisms may commonly exhibit multivariate phenotypic shifts even though not all diverging traits necessarily represent adaptations to the extreme conditions.

Keywords
Divergent natural selection, ecological speciation, extreme environments, Gambusia, life-history evolution, phenotypic convergence, Poecilia, replicated evolution.

INTRODUCTION
The predictability of evolution has captivated the minds of evolutionary ecologists dating back to the publication of Charles Darwin’s (1859) On the Origin of Species. Even though evolutionary biology is often descriptive and reconstructs historical events (e.g. Bowmaker 1998; Whiting et al. 2003), modern evolutionary ecology can be both experimental and predictive (e.g. Schluter 1994; Losos et al. 2006; Stern & Orgogozo 2008; Langerhans 2010). Predictable and convergent differentiation, where independent lineages repeatedly evolve similar phenotypes along replicated environmental gradients, provides some of the strongest evidence of adaptation by natural selection (e.g. Losos 2011; Langerhans & Riesch 2013). Not surprisingly, convergent evolution has been documented for a wide variety of taxa, traits and sources of selection (e.g. Hockstra et al. 2006; Protas et al. 2006; Hulsey et al. 2008; reviewed by Langerhans & Riesch 2013). However, studies demonstrating convergent life-history evolution are relatively rare and mainly focus on convergence as a response to divergent predator regimes or differences in resource availability (e.g. Reznick et al. 1996; Johnson 2001; Duponchelle et al. 2008; Riesch et al. 2013; but see Riesch et al. 2010a). Here, we evaluate the evidence for convergent life histories in several clades across two genera of livebearing fishes (family Poeciliidae) that have independently colonised extreme environments characterised by the presence of toxic hydrogen sulphide (H$_2$S).

H$_2$S is a widespread natural toxicant at deep-sea hydrothermal vents, sulphur springs, coastal salt marshes or mudflats, but can also originate from anthropogenic sources such as pulp mills or tanneries (Bagarinao 1992; Grieshaber & Völk 1998). H$_2$S is acutely toxic to most metazoans, because it interferes with mitochondrial respiration and blood oxygen transport, and simultaneously leads to extreme hypoxia in the environment (Bagarinao 1992; Grieshaber & Völk 1998). These adverse effects notwithstanding, a number of livebearing fishes inhabit sulphide springs exhibiting continuously high concentrations of H$_2$S (e.g. Tobler et al. 2008b, 2011; Riesch et al. 2010a; Plath et al. 2013) – concentrations which are lethal for various freshwater fishes even at relatively short-term exposures (Table S1). Examining the specific adaptations of fishes exposed to natural sources of H$_2$S is thus of great importance also for ecotoxicological research and ecological risk assessment (e.g. Calow & Forbes 2003). Making use of sulphide-adapted lineages in two distantly related genera (Gambusia and Poecilia sensu lato; Hrbeck et al. 2007), we investigate if convergent life-history evolution previously reported for fishes from a very small geographical area in southern Mexico (Riesch et al. 2010a) is in fact representative of (and thus, evolves convergently and predictably in) the various livebearing fishes that have independently colonised sulphur springs in the United States, Caribbean and Neotropics. Within Gambusia, we investigated populations from four different clades, the G. affinis clade, the G. bolivianus clade, the G. babbi clade and the G. euryzona/G. sexradiata clade (Lydeard et al. 1995; Langerhans, unpubl. data). Within
Poecilia (sensu lato), we examined the subgenus Acanthophacelus (P. reticulata), and two clades within the subgenus Mollienesia (shortfin mollies: P. mexicana, P. sulphuraria, sailfin mollies: P. latipinna) (Breden et al. 1999; Hrbek et al. 2007).

Because the ratio of sulphide influx to oxidation is crucial for the successful maintenance of homeostasis, natural selection should favour larger offspring size at birth in H2S-rich – and thus hypoxic – environments due to a more beneficial body volume-to-surface area ratio that facilitates detoxification in larger neonates (theory: Powell 1989; empirical evidence for the mussel Macoma balthica: Jahn et al. 1997). Furthermore, life-history theory predicts offspring size to become larger in low-quality environments (theory: Rollinson & Hutchings 2013a; empirical evidence for Atlantic salmon, Salmo salar: Rollinson & Hutchings 2013b), as well as in habitats with increased resource competition (theory: Brockelman 1975; Jørgensen et al. 2011; empirical evidence for the guppy: Bashey 2008), as for example, induced by the increased population densities typically found in sulphide springs (J. Jourdan, D. Bierbach, M. Plath & R. Riesch, unpublished data). Congruently, previous studies on fish from sulphide springs in southern Mexico found them to produce significantly larger, but also fewer, offspring than their counterparts from non-toxic waters (Riesch et al. 2010a,b). Thus, we specifically predicted that females from sulphidic environments should produce larger but fewer offspring than their closest relatives from non-sulphidic waters. However, convergent phenotypic evolution can occur even in the absence of direct selection on the apparently convergent trait(s) (Losos 2011), and so reduced fecundity in sulphidic waters could simply reflect a by-product of the prominent life-history trade-off between offspring size and fecundity rather than reflecting a result of direct selection on fecundity per se (e.g. Smith & Fretwell 1974; Riesch et al. 2010a). Therefore, we further predicted that patterns of divergence should show stronger convergence across toxic environments for offspring size than fecundity. The large degree of independent replication of sulphidic vs. non-sulphidic habitats in this study allows us, for the first time, to specifically test these predictions.

MATERIAL AND METHODS

Field collections

We collected life-history data from 623 pregnant females of 22 populations from nine species across the range of poeciliid fishes in the New World: samples were collected in the United States of America (Oklahoma and Florida), the Bahamas, Mexico and Venezuela (Fig. 1; Table S2). Except for P. reticulata, which were originally collected by K. O. Winemiller and obtained from the Texas Natural History Collection (TNHC, University of Texas at Austin), collections were made by the authors. All fishes were sacrificed immediately upon capture and preserved in 10% formalin. For seven of the nine species, we collected populations from both sulphidic and non-sulphidic habitats. However, this was not possible for two of the species, Gambusia eurytoma and Poecilia sulphuraria, because they are endemic to sulphur-spring complexes in Tabasco, southern Mexico (Tobler et al. 2008b).

Figure 1 Map of sampling localities. Red circles indicate the location of sulphidic habitats, blue circles indicate non-sulphidic sample sites; bi-coloured circles indicate systems for which sulphidic and at least one non-sulphidic sample stem from adjacent populations.
Life-history analyses

We performed standard life-history dissections (for details see Riesch et al. 2010a,b) to estimate average offspring size at birth [mg] and fecundity (i.e. number of offspring per clutch) for each population. During dissections performed on gravid females, all developing offspring were removed and their stage of development determined according to the classification scheme outlined in Riesch et al. (2011). For each population, offspring size at birth was then estimated using the intercept and slope from a population-specific regression of log10-transformed offspring dry weight [mg] on embryonic stage of development (Reznick et al. 2002; Riesch et al. 2011, 2013). Population ‘means’ (rather than individual-level data) were then examined in all subsequent analyses (see below). Data on offspring size at birth and fecundity for P. mexicana, P. sulphuraria and G. sexradiata were re-analysed from a previous study (Riesch et al. 2010a).

Phylogenetic analyses

To allow for a phylogenetically independent comparison of life histories of populations/species from sulphidic and non-sulphidic habitats, we first constructed a phylogeny for the nine species examined in the study based on a 402 bp fragment of the mitochondrial cytochrome b gene (cyt b). Sequence data included new and previously published sequences from Genbank (Table S3). For new Gambusia sequences, we used primers and conditions described in Lydeard et al. (1995), while new Poecilia sequences were obtained following the protocols described in Tobler et al. (2008a). Our final alignment included 37 ingroup samples: nine G. affinis, five G. holbrooki, four G. hubbsi, two G. sexradiata, one G. eurystoma, five P. reticulata, three P. latipinna, three P. sulphuraria, and five P. mexicana. We included one specimen of Xenodezia ctenolepis as an outgroup. For all species for which multiple populations were included in the life-history analysis, we also included multiple populations for the phylogenetic analysis.

We inferred phylogenetic relationships using Bayesian inference with MrBayes 3.2.1 (Ronquist et al. 2012). We determined the optimal maximum-likelihood model of sequence evolution for each codon position of the cyt b gene using jModeltest2 (Guindon & Gascuel 2003; Darriba et al. 2012); the selected model for codon 1 was K80, JC was selected for codon 2, and TrN+Γ for codon 3. MrBayes was run for 5 000 000 generations, sampling trees every 100 generations, with the lower 25% of trees discarded as burn-in (after checking that stationarity was reached) for the computation of a 50% majority-rule consensus tree. We calculated support values for inferred clades using Bayesian posterior probabilities. The resulting phylogeny (Fig. S1) was highly consistent with previously published phylogenies (e.g. Lydeard et al. 1995; Breden et al. 1999; Hrbek et al. 2007), and yielded generally strong support for most nodes. Using the Bayesian 50% majority-rule consensus phylogeny, we then pruned the phylogeny so that the number of tips per species equalled the number of populations for which we had life-history data, and set the branch lengths within species to the average intraspecific branch length observed for each species in the original phylogeny. On the basis of our 402 bp mtDNA fragment, we were unable to resolve the reciprocal monophyly of G. sexradiata and G. eurystoma (Fig. S1), and we effectively treated them as a single species comprising three populations (tips) in all subsequent analyses (Fig. 2). The phylogeny suggested that these nine species form seven distinct clades, with the two sister-species complexes (P. sulphuraria/P. mexicana and G. eurystoma/G. sexradiata) forming just one clade each (Fig. S1). Comparisons between sulphidic and non-sulphidic environments were present within each clade.

Figure 2 Pruned Bayesian phylogeny using the mtDNA cyt b gene (50% majority-rule consensus tree, numbers indicate Bayesian posterior probability for each node), and mean life-history values for each population. Sulphidic habitats are given in red and non-sulphidic habitats in blue. Dashed lines indicate clades used in analyses.
Statistical analyses

Despite the fact that H₂S-concentrations varied substantially among sulphidic habitats (Table S2), all sulphidic sites are acutely toxic to most fish, as even constant exposure to micromolar amounts results in high mortality (Bagarinaz 1992; Grieshaber & Völkel 1998; Table S1). Thus, we treat toxicity as a dichotomous (present vs. absent) rather than a continuous variable in all statistical analyses.

Because species with larger adult female body size tend to exhibit larger offspring size at birth and greater fecundity (e.g. Jennions & Telford 2002), we conducted phylogenetic size correction following Revell (2009). We calculated residuals from linear regressions of log-transformed offspring weight and square-root-transformed fecundity on log-transformed female standard length, controlling for non-independence due to phylogenetic history and used these values in all subsequent analyses.

To test for life-history differences between toxicity regimes, we used two different approaches. First, we conducted a multivariate analysis of variance (MANOVA) using the size-corrected residuals for estimated offspring size at birth and fecundity as dependent variables, and including ‘clade’ (the seven major clades), ‘toxicity’ (sulphidic vs. non-sulphidic), and their interaction as independent variables. This treats the seven different clades effectively as blocks, testing for effects of toxicity on life-history traits within each clade. We evaluated the relative importance of model terms using the effect size measurement of Wilks’s partial η² (measure of partial variance explained by a particular term; multivariate approximation of SS-effect/[SS-effect + SS-error], see appendix of Langerhans & DeWitt 2004), and used post hoc ANOVAs to evaluate the nature of differences between toxicity regimes (i.e. whether offspring size, fecundity, or both differed between environments). Since the interaction of ‘clade-by-toxicity’ was not significant (P = 0.59), we removed it from the final model.

While this statistical approach successfully controls for differences between clades and offers an intuitive metric for comparing the relative importance of adaptation to extreme environments and clade-level differences (e.g. are life-history differences between environments within clades smaller, as great, or greater than average between-clade differences?), it does not take into account the hierarchical relationships among clades. Therefore, we also conducted a phylogenetic MANOVA using the R package GEIGER (Harmon et al. 2008), in which the F-statistic from ordinary one-way MANOVA testing for toxicity effects was compared to an empirical null distribution of the F-statistic based on simulated dependent variables on the phylogeny under a multivariate Brownian motion model of evolution (10 000 simulations) (Garland et al. 1993). P-values were calculated as the proportion of times the simulated F-statistic was greater than the observed F-statistic.

We were further interested in evaluating whether populations inhabiting sulphidic environments evolved both larger offspring size and reduced fecundity in concert, or whether one of the traits was most responsible for convergent patterns of evolution and simply ‘dragged’ the other along. For this, we again used two separate approaches; in each case, we tested for differences in one life-history trait between sulphidic vs. non-sulphidic environments while controlling for covariation with the other life-history trait. By statistically removing any correlation between the two life-history traits, these analyses thus examined differences between toxic regimes in each life-history trait independent of any indirect effect of the other life-history trait. First, we conducted two separate ANCOVAs, using estimated offspring size at birth as the dependent variable in the first and fecundity as the dependent variable in the second. We further included ‘toxicity’ and ‘clade’ as factors in both models, and included fecundity as a covariate in the ANCOVA on offspring size, and offspring size as a covariate in the ANCOVA on fecundity. As above, we employed a second approach to properly incorporate the hierarchical relationships among lineages within the analysis. For our second approach, we conducted two phylogenetic ANCOVAs (using the ape package in R; R Core Team 2012), one for each life-history trait, while again controlling for the respective other trait by including it as a covariate (otherwise the procedure was similar to the phylogenetic MANOVA described above). We initially also included the interaction between the covariate and toxicity in the models, but subsequently removed them from the final models (P > 0.35 in both cases; d.f.: denominator reduced by number of zero-length branches = 7).

RESULTS

MANOVA revealed highly significant effects of both clade (F₁,₁₂,₂₆ = 5.37, P = 0.0002) and toxicity (F₂,₁₃ = 10.72, P = 0.0018). Thus, while clades differed from one another in life-history traits, consistent differences between toxic and non-toxic environments were evident. Moreover, our measures of effect size (η² p; clade = 71.25, toxicity = 62.26) indicate that differences between sulphidic and non-sulphidic waters were comparable in magnitude to clade-level differences. Similarly, our phylogenetic MANOVA revealed a highly significant effect of toxicity (P = 0.0003), indicating strong and consistent associations between life-history traits and toxicity after correcting for the hierarchical relationships among species. Post hoc tests and inspection of mean trait differences between toxicity regimes demonstrated that populations inhabiting sulphidic environments exhibited a larger offspring size at birth (ANOVA: P = 0.0003) and produced fewer offspring per brood (P = 0.0241) than populations in non-sulphidic habitats, regardless of clade (Fig. 2).

In the ANCOVA for offspring size with ‘clade’ as a blocking factor, and controlling for potential correlation with fecundity, we uncovered a highly significant effect of toxicity on offspring size (toxicity: F₁,₁₂ = 10.64, P = 0.0062; clade: F₀,₁₃ = 3.31, P = 0.033; fecundity: F₁,₁₂ = 1.81, P = 0.20), and this was mirrored in the phylogenetic ANCOVA on offspring size (toxicity: F₁,₁₂ = 14.72, P = 0.0024; fecundity: F₁,₁₂ = 5.72, P = 0.034) (Fig. 3a). In contrast, both the regular ANCOVA for fecundity (toxicity: F₁,₁₂ = 0.40, P = 0.54; clade: F₀,₁₃ = 1.52, P = 0.25; offspring size: F₁,₁₂ = 1.38, P = 0.26) and the phylogenetic ANCOVA (toxicity: F₁,₁₂ = 0.05, P = 0.82; offspring size: F₁,₁₂ = 5.72, P = 0.034) failed to detect a significant influence of toxicity on fecundity when controlling for correlation with offspring size (Fig. 3b).

DISCUSSION

Extreme environments are rare by definition, yet without replication both the causal role of environmental factors in driving adaptive evolutionary change and the generality of those patterns are difficult to discern. Drawing upon life-history data from nine different species of livebearing fishes that have independently colonised toxic...
H2S-toxicity: female poeciliid fishes in sulphidic habitats produce significantly larger offspring compared to females from non-sulphidic habitats. Similar patterns where certain environments seem to drag along due to the fundamental trade-off between offspring size and fecundity have already left descendants behind in the form of the two sulphide endemics, *G. enystoma* and *P. sulphuraria* (Tobler et al. 2008b; Plath et al. 2013).

Previous studies demonstrated convergent phenotypic differentiation and replicated ecological speciation as a result of colonisation of sulphide habitats within a very small geographical area in southern Mexico (e.g. Riesch et al. 2010a; Tobler et al. 2011; Plath et al. 2013). Here, we demonstrate that life-history divergence resulting from colonisation of sulphidic waters is not unique to southern Mexico, but actually represents a widespread and predictable evolutionary phenomenon across multiple phylogenetically and geographically independent lineages of livebearing fishes. Even though divergence might be temporary in some of these systems, and thus might not ultimately result in complete speciation (e.g. Nosil et al. 2009; Langerhans & Riesch 2013), H2S-driven ecological speciation clearly has the potential to result in complete speciation, because it has already left descendants behind in the form of the two sulphide endemics, *G. enystoma* and *P. sulphuraria* (Tobler et al. 2008b; Plath et al. 2013).

Losos (2011) cautioned that convergent phenotypic evolution could occur even in the absence of direct selection on the apparently convergent trait – convergence in one trait can result from a correlated response to selection on another character, e.g. via pleiotropy or constraints on certain trait combinations. Our previous findings in extremophile fishes from southern Mexico could be interpreted as suggesting that H2S-toxicity simultaneously imposes selection on both offspring size and fecundity. Nonetheless, based on fitness considerations (Brockelman 1975; Powell 1989; Jahn et al. 1997; Bashey 2008; Jorgensen et al. 2011; Rollinson & Hutchings 2013a,b), our second prediction was that only offspring size – not fecundity – should experience direct selection in extreme environments (see also Riesch et al. 2010a). Our univariate analyses support this prediction and strongly suggest that only offspring size diverged as a direct response to selection from toxicity, while fecundity was dragged along due to the fundamental trade-off between offspring size and fecundity in fishes and other animals (e.g. Smith & Fretwell 1974). In our study, populations in toxic habitats consistently exhibited a larger offspring size at birth even when controlling for covariation with fecundity, but fish in these toxic habitats did not differ in fecundity when controlling for covariation with offspring size. Thus, while toxic waters directly altered the evolution of offspring size, fecundity was only indirectly affected; hence, observed differences in fecundity between sulphidic and non-sulphidic environments actually only reflected divergence in the correlated trait, offspring size.

This was further exemplified by the somewhat irregular patterns uncovered for fecundity. While populations from sulphidic waters had the largest offspring size within each clade, they did not always have the lowest fecundity. For example, guppies from a sulphidic spring complex in Venezuela produced the largest but also the most offspring compared to the two control populations from non-sulphidic waters. Similar patterns where certain environments seemingly select for ‘super’ organisms that defy life-history theory are well documented from other systems; however, closer investigation...
often points to previously undetected selective agents that impact how classical life-history trade-offs are resolved (e.g. Reznick et al. 2000). In the Venezuelan guppies examined here, increased body-cavity volume for embryo development could explain the observed pattern, resulting from scenarios such as genetic drift or differences in selection via predation or resource availability. Future studies will have to investigate this further.

In conclusion, the apparent joint evolution of low fecundity and large offspring size in toxic habitats is the result of an inter-dependency rather than direct selection on both traits. This finding strongly supports Losos’ (2011) cautionary note not to immediately proclaim a specific trait as a direct target of selection simply because of a strong correlation between that particular trait and a certain environmental variable across replicated environmental gradients. Extremophile organisms in particular may commonly exhibit multivariate phenotypic shifts even though not all diverging traits necessarily represent adaptations to the extreme conditions.

ACKNOWLEDGEMENTS

We thank L. Arias-Rodriguez and F. J. García de León for help during the collections in Mexico. The Mexican government (CONAPESCA: DGOPA/16986/191205/8101, DGOPA/02232/230706/1079, DGOPA/06192. 240608-1562, and SGPA/DGVS/04751/08), the Municipal of Tacotalpa (SM/1133/208) and the NAPESCA: DGOPA/16986/191205/8101, DGOPA/02232/230706/1079, DGOPA/06192. 240608-1562, and SGPA/DGVS/04751/08), the German Science Foundation (DFG, PL 470/1-2, to MP), the Alexander von Humboldt Foundation (to IS), and the W. M. Keck Center for Behavioral Biology (to RR and RBL). Financial support came from the National Science Foundation of America (DEB-0743406 to IS; IOS-1121832 to MT, DEB-0842364 to RBL), the Mexican government (CONACyT: DGAPA-230706/1079, DGOPA/16986/191205/8101, DGOPA/02232/230706/1079, DGOPA/06192. 240608-1562, and SGPA/DGVS/04751/08), the Municipal of Tacotalpa (SM/1133/208) and the NAPESCA: DGOPA/16986/191205/8101, DGOPA/02232/230706/1079, DGOPA/06192. 240608-1562, and SGPA/DGVS/04751/08).

STATEMENT OF AUTHORSHIP

RR designed the study and collected the data, all authors contributed specimens, RBL performed analyses, RR created figures and wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

REFERENCES


**SUPPORTING INFORMATION**

Additional Supporting Information may be downloaded via the online version of this article at Wiley Online Library (www.ecologyletters.com).

Editor, Jean-Michel Gaillard
Manuscript received 12 June 2013
First decision made 16 September 2013
Manuscript accepted 7 October 2013