vol. 190, no. 1 the american naturalist july 2017

How Parallel Is Parallel Evolution? A Comparative Analysis in Fishes

Krista B. Oke,[1,*] Gregor Rolshausen,[1,2] Caroline LeBlond,[1] and Andrew P. Hendry[1]

1. Redpath Museum and Department of Biology, McGill University, 859 Sherbrooke Street West, Montreal, Quebec H3A 2K6, Canada;
2. Senckenberg Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany

Submitted September 12, 2016; Accepted February 14, 2017; Electronically published April 26, 2017

Online enhancements: appendix. *Dryad data:* http://dx.doi.org/10.5061/dryad.bb206.

abstract: Evidence of phenotypic parallelism is often used to infer the deterministic role played by natural selection. However, variation in the extent or direction of divergence is often evident among independent evolutionary replicates, raising the following question: just how parallel, overall, is parallel evolution? We answer this question through a comparative analysis of studies of fishes, a taxon where parallel evolution has been much discussed. We first ask how much of the among-population variance in phenotypic traits can be explained by different "environment" types, such as high predation versus low predation or benthic versus limnetic. We then use phenotypic change vector analysis to quantify variation in the direction (vector angles) and magnitude (vector lengths) of environment-associated divergence. All analyses show high variation in the extent of parallelism—from very high to very low, along with everything in between—highlighting the importance of quantifying parallelism rather than just asserting its presence. Interestingly, instances of low extents of parallelism represent important components of divergence in many cases, promising considerable opportunities for inferences about the factors shaping phenotypic divergence.

Keywords: parallel evolution, convergent evolution, nonparallel evolution, repeatability, fishes.

Introduction

The independent evolution of similar traits in similar environments is classically used to support a deterministic role for natural selection in shaping evolution. At the phenotypic level, this pattern has been variously called evolutionary "repeatability," "predictability," "parallelism," or "convergence" (Clarke 1975; Langerhans et al. 2004; Arendt and Reznick 2008; Losos 2011; Wake et al. 2011); we will use the term "parallelism" owing to its predominance in the literature for our focal taxon. Increasingly, however, many cases are being described of considerable variation in the extent of phenotypic similarity among populations from similar environments, both between species (Brinsmead and Fox 2002; Langerhans and DeWitt 2004; Langerhans et al. 2006; Rosenblum and Harmon 2010) and within species (e.g., Hoekstra and Nachman 2003; Landry and Bernatchez 2010; Kaeuffer et al. 2012; Ravinet et al. 2013*a*; Fitzpatrick et al. 2014; Moore et al. 2016). Many potential factors (or any combination of them) could cause this extensive (non)parallelism, including (1) variation in natural selection despite seemingly parallel environments or the use of overly simplified habitat categories (Berner et al. 2008; Kaeuffer et al. 2012; Moore et al. 2016); (2) differences in genetic factors, such as drift, mutation, gene flow, and past selection (Hendry et al. 2001; Schluter et al. 2004; Bolnick and Nosil 2007); (3) variation in sexual selection (Bonduriansky 2011; Maan and Seehausen 2011); and (4) alternative phenotypic solutions to the same functional problem (Gould and Lewontin 1979; Arnold 1983; Alfaro et al. 2004; Wainwright et al. 2005). Thus, deviations from strict phenotypic parallelism provide a useful substrate for exploring the relative importance of these different forces in shaping the diversity of life. The utility of such endeavors will depend on just how variable the extent of parallelism typically is in nature. Here, we consider the extent of parallelism in studies of parallel evolution by means of a comparative analysis across studies in the fish literature.

Our analysis focuses on fishes, which present several relevant benefits (see also Moore et al. 2016). First, fishes are a highly diverse group of well-studied organisms that include several classic examples of parallel evolution, such as high-predation versus low-predation guppies (*Poecilia reticulata*; e.g., Reznick and Endler 1982), benthic versus limnetic threespine stickleback (*Gasterosteus aculeatus*; e.g., Schluter and McPhail 1992), and sulfidic versus nonsulfidic spring livebearing fish (often also comparing cave vs. surface populations; *Poecilia mexicana*; e.g., Tobler et al. 2011; Riesch et al. 2016). Second, fishes have a reasonably common set of phenotypic traits that facilitate among-species comparisons of within-species parallelism. Examples include body shape (e.g., Taylor et al. 1997; Østbye et al. 2005; Adams et al. 2008; Kaeuffer et al. 2012; Ingley et al. 2014*a*), numbers and sizes

---

[*] Corresponding author; e-mail: kristaoke@gmail.com.

Am. Nat. 2017. Vol. 190, pp. 1–16.
DOI: 10.1086/691989

of offspring (e.g., Reznick et al. 1996; Jennions and Telford 2002; Kavanagh et al. 2010; Riesch et al. 2013), and gill raker traits (Taylor and Bentzen 1993; Palkovacs et al. 2008; Matthews et al. 2010; Tobler et al. 2011; Evans et al. 2013; Theis et al. 2014*a*). Third, population divergence into different environments often occurs independently in different watersheds, providing ample evolutionary replication to assess parallelism (e.g., Johnson 2001; Østbye et al. 2006; Tobler et al. 2011; Kaeuffer et al. 2012).

Most studies of phenotypic parallelism seek only to infer its existence, often by simply considering the significance of a grouping factor, such as "environment" (the term we will use for habitat types) or "ecotype" (the term we will use for population types), in statistical models analyzing multiple populations from multiple environments. This qualitative yes-or-no approach fails to inform just how parallel the inferred parallel evolution is, and it also discourages investigations into the causes of deviations from strict parallelism. Therefore, along with other recent investigators (e.g., Langerhans and DeWitt 2004; Schluter et al. 2004; Berner et al. 2008; Kaeuffer et al. 2012), we feel it is critically important to quantify and emphasize the extent of phenotypic parallelism. Quantification can come in several forms, which we here illustrate by reference to a common study design: two ecotypes (e.g., high predation vs. low predation, benthic vs. limnetic, sulfidic vs. nonsulfidic) from each of multiple independent origins (typically the pairs are from separate watersheds). In this classic design, a researcher can ask several questions, the first two focusing on trait-by-trait analyses and the second two on multidimensional trait analyses. First, how often is the direction of environment-associated trait divergence (e.g., deeper bodies in benthic than limnetic fish, more offspring in high-predation than low-predation fish) the same across multiple evolutionary replicates? Second, how much of the variation among population trait means (or among all individuals) can be explained by the environment (or ecotype) term in a statistical model? Three, to what extent does the direction of environment-associated divergence in multidimensional trait space vary among replicates (i.e., variation in angles between divergence vectors)? Fourth, to what extent does the magnitude of environment-associated divergence in multidimensional trait space vary among replicates (i.e., differences in the lengths of divergence vectors)? The first and third questions consider only variation in the direction of divergence. The fourth question considers only variation in the magnitude of divergence. The second question incorporates variation in both magnitude and direction.

We see considerable value in all of the above questions. For instance, separately considering variation in the direction and magnitude of divergence yields a greater ability to tease apart the drivers of deviations from strict parallelism. For instance, variation in the direction of divergence might be due to variation in the direction of selection (Landry and Bernatchez 2010; Kaeuffer et al. 2012; Fitzpatrick et al. 2014), whereas variation in the magnitude of divergence might be due to variation in constraints, such as gene flow (Hendry and Taylor 2004; Bolnick and Nosil 2007). Although it might be tempting to consider variation only in direction when evaluating parallel evolution, variation in magnitude is also critical. Imagine a scenario where multiple ecotype pairs all diverge in the same direction, but half of the pairs show very slight (presumably biologically unimportant) differences and the other half show very large (presumably biologically important) differences. Despite similar directions of divergence in all pairs, invoking strong parallelism would be unhelpful and misleading. Hence, quantification of parallelism should consider variation in both the magnitude and the direction of divergence, separately and in combination. Hence, when discussing parallelism as a general phenomenon, we consider variation in both the direction and the magnitude of divergence. We use the term "(non)parallelism" to refer, in general, to the potential for deviations from strict parallelism or variability among ecotype pairs in their extent of parallelism. However, when referring specifically to results from the vector analyses that separate these two components, we use the more explicit terms "similarity in direction" and "similarity in magnitude."

The quantitative assessment of the extent of parallelism, which we undertake in the present article, not only informs just how parallel parallel evolution is but also emphasizes the utility of the concept of (non)parallelism for inferring the ecological and evolutionary drivers of diversification. We investigate the extent of parallelism in published studies of parallel evolution in fishes to ask the question, how parallel is parallel evolution?

Methods

We searched for studies of fish that examined conspecific populations of multiple ecotypes in multiple locations, such as lake versus stream, benthic versus limnetic, and high predation versus low predation. The search terms in Web of Science were "parallel*" (or "convergen*") and "evolution" and "fish*" (the final search was conducted on March 7, 2016). Disagreement exists as to optimal use of the terms "parallel" versus "convergent"; from our phenotypic perspective, the difference is immaterial. Additional studies were found by scanning the reference lists of studies identified in our search and by contacting colleagues. For these reasons, the selected studies do not represent a random sampling of studies of evolution but rather studies specifically evaluating repeated or parallel evolution (or related studies from study systems widely considered to have undergone parallel evolution). Studies were retained for analysis if they reported mean trait values for two or more populations of each of two or more environments (or "habitat types") and if they also provided evidence for independent origins of the populations. Examples of this latter evidence include (1) geographical information showing that populations of similar ecotypes were in separate watersheds or otherwise physically isolated or (2) genetic information showing that populations of similar ecotypes were independently derived. The final data set included 618 trait comparisons in 92 studies of 23 species (see table A1, available online, for the list of studies; all data are deposited in the Dryad Digital Repository, http://dx.doi.org/10.5061/dryad.bb206 [Oke et al. 2017]).

During data collection, we also recorded sampling type, trait type, study system, species, study design, and study identity. Sampling types were either "wild caught" or "common garden," depending on whether the fish were captured from the wild or reared for their entire lives in a common laboratory environment. Trait type categories were inspired by Mousseau and Roff (1987), with several modifications. We retained the categories morphology, meristic, life history, and other (e.g., color). However, so few traits fell into the behavior and physiology categories that they were added to the category of other, whereas so many studies measured trophic (diet) traits that it was made into a separate category. To account for the possibility of pseudoreplication, study system was designated according to the species, location, and selective pressure or ecotype contrast. For example, benthic-limnetic, lake-stream, and mud-lava threespine stickleback would all be categorized as separate "systems," as would lake-stream stickleback pairs from different regions (e.g., Vancouver Island, Haida Gwaii, Switzerland, and Ireland). Study design designated either "paired" or "unpaired" designs. In the paired design, two (or more) ecotypes were found in sympatry or parapatric in each of multiple independent locations. An example would be parapatric lake-stream stickleback pairs in each of multiple watersheds on Vancouver Island (Kaeuffer et al. 2012). In the unpaired design, allopatric populations of different ecotypes were found in multiple locations, without any obvious geographic ecotype pairings. An example would be populations of marine versus freshwater ecotypes spread across independent lakes or marine sites, such as the marine-lake-pond ecotypes of ninespine stickleback in Fennoscandia (*Pungitius pungitius*; Herczeg et al. 2010). Study identity, another factor that helps eliminate pseudoreplication, was accounted for by simply assigning a unique identifier to each study.

Analysis: Variance among Means

We started by quantifying, for each study, the percentage of variance among population means that could be explained by the ecotype designation. These percentages were obtained through simple univariate ANOVA on the population means for a given trait, with ecotype as a fixed factor. Other factors were not included because we are interested in the variance explained by ecotype designations before taking other terms into consideration. Other analyses that consider the variation explained by ecotype designations while also accounting for variation due to other factors would be very interesting but are beyond the scope of our study. The reasons for using population means were twofold: (1) variation among means is what studies of parallel evolution seek to explain and (2) data at the individual-fish level were rarely available for published studies. The R^2 values for the ecotype term were then analyzed in two ways. First, we described the distribution of R^2 values so as to consider the dispersion and typical strength of parallelism estimates. This distribution will allow future authors to place their own estimates in the context of past work. Specifically, we fit a beta distribution using maximum likelihood in the fitdistrplus package in R (R Development Core Team 2012; Delignette-Muller and Dutang 2015). Goodness of fit was assessed by visual examination and with the Akaike information criterion (Akaike 1987), the Bayesian information criterion (Schwarz 1978), and log-likelihood ratios (Fisher 1922; Woolf 1957).

Second, we used the R^2 values as response variables in binomial generalized linear mixed models (GLMMs) in the lme4 package in R (Bates et al. 2015) because R^2 values are bounded between 0 and 1. For this analysis, traits measured under wild-caught and common-garden conditions were analyzed separately because the former were much more common (estimates from 82 studies) than the latter (estimates from 15 studies). The GLMM for wild-caught fish included two fixed factors: trait type and study design. The GLMM for common-garden fish included, owing to data availability, only one fixed factor: study design. The opportunistic nature of our data collection introduced the possibility for pseudoreplication at several different levels. To account for potential pseudoreplication, we included three related random factors: species, study system, and study identity. The significance of each fixed term was assessed using analysis of deviance with type III sums of squares in the car package in R (Fox and Weisberg 2011; R Development Core Team 2012).

Third, R^2 values could be influenced by sample size, and even in the absence of parallelism R^2 might be higher than zero simply by chance. Thus, to provide a baseline level of parallelism that would be expected by chance, we performed permutations for each trait in which each population was randomly assigned a habitat category, and then the ANOVA procedure from above was repeated. Each trait was permuted 100 times (because there are a limited number of reassigned combinations possible, especially for studies of few populations), and we did not allow for replacement in order to hold constant the number of populations of each habitat type in the ANOVA. The results of this analysis (fig. A1; figs. A1, A2 are available online) showed that the vast majority of R^2 values were far higher than would be expected by chance and that sample size does not drive our R^2 results. In addition, we fit a linear regression of actual R^2 results to the mean R^2 from the 100 permutations for that trait and extracted the residuals. As a further test of R^2 values once corrected for the R^2 expected at random, we created mixed models for the residuals from this model with the same structure as the above GLMMs, except that residual R^2 values were normally distributed and did not require a binomial GLMM.

Finally, we included all traits for which data were provided, but some traits might not have been expected to show parallelism. For example, selection might not act on every trait during divergence, or researchers might have included traits for convenience, for historical reasons, or to test other related hypotheses. It is possible that including all traits regardless of whether they were predicted to show parallelism may have lowered our estimates of parallelism. Thus, for each trait we recorded whether the original study had stated an expectation of parallelism for that trait or had called the results for that trait parallel. To determine whether traits expected or claimed to show parallelism differed significantly in observed R^2, we then conducted an ANOVA of our R^2 results with expectation of parallelism/convergence (parallelism, convergence, or neither) and statement of parallelism/convergence (parallel, convergent, or neither) included as fixed factors.

Analysis: Phenotypic Change Vector Analysis (PCVA) and Phenotypic Trajectory Analysis (PTA)

We next analyzed studies of the paired design with PCVA (for studies with two ecotypes) or PTA (for studies with more than two ecotypes; Collyer and Adams 2007; Adams et al. 2009). These methods project population means into multidimensional trait space (fig. 1) and connect the different ecotypes of each pair by vectors (PCVA) or trajectories (PTA). The length and direction of vectors and the length, direction, and shape of trajectories can then inform the extent of similarity across ecotype pairs in the direction and magnitude of between-ecotype divergence.

Our questions and data necessitated some modifications to the usual PCVA/PTA implementation. First, the lack of raw data precluded permutation tests for the significance of specific pairwise differences, which fortunately was not relevant to our general question. Second, vectors based on different traits cannot be directly compared, so we compared the pairwise (among different ecotype pairs) proportional difference in vector lengths (absolute difference in vector length divided by the sum of the two vectors) and vector directions (difference in vector angles divided by 180°). Third, all traits included in a given PCVA/PTA must be measured in units that can be directly compared (Huttegger and Mitteroecker 2011). For each study, we therefore performed separate PCVAs/PTAs for different types of traits, but only studies that included two or more traits of a given type could be included. Fourth, PCVA/PTA requires data for all populations, and so populations with missing data for any trait were excluded from analyses that included that trait. Due to these modifications, we created custom R scripts to perform these analyses, but we recommend the trajectory.analysis function in the geomorph package (Adams and Otarola-Castillo 2013) for PCVA/PTA of individual study systems.

For each pairwise comparison of ecotypes in each study, PCVA yielded relative pairwise differences in vector directions and lengths, and PTA yielded pairwise differences in trajectory shape and relative differences in trajectory directions and lengths. As was the case for R^2 values (above), we fit beta distributions to the PCVA data but not to the PTA data owing to low sample size (24 traits from six studies). PCVA data were analyzed with GLMMs using separate models for relative differences in directions and lengths. Trait type was a fixed factor, and study identity was a random factor. Finally, we asked whether the extent of parallelism differed between relative differences in directions versus lengths, both of which were measured as proportions. Here the GLMM included trait type as a fixed factor, data type (difference in direction or length) as a fixed factor, and pairwise comparison (the identity of the two ecotype pairs that were compared) as a random factor. Including study as an additional random factor did not improve model fit. Sample sizes for PTA data were too low (24 traits from six studies) for similar analyses.

Results

Individual ANOVAs for each trait in each study yielded a mean (±SD) variance explained by ecotype (R^2) of 0.460 ± 0.316, which was best fit by a beta distribution having shape parameters of $\alpha = 0.598 \pm 0.003$ and $\beta = 0.773 \pm 0.004$. A wide range of R^2 values was evident, including some highly parallel traits ($R^2 > 0.90$), some weakly parallel traits ($R^2 < 0.10$), and everything in between (fig. 2). Overall, in the studies included in our analyses, weak parallelism was at least as frequently observed, if not more so, than strong parallelism. For instance, 54.2% of the R^2 values were less than 0.50, meaning that—in more cases than not—the ecotype designation explained less than half of the variation among population means. Moreover, only 10.4% of the estimates indicated very strong parallelism ($R^2 > 0.90$), whereas 18.8% of the estimates indicated very weak parallelism ($R^2 < 0.10$). Importantly, variation in the extent of parallelism was spread across traits within studies and across study systems (table A1), indicating considerable generality to these patterns.

GLMMs of wild-caught fish revealed that trait type ($\chi^2_4 = 15.6$, $P = .004$; fig. 3*A*) influenced the extent of parallelism (R^2), with a very strong interaction between trait type and study design (paired vs. unpaired, $\chi^2_4 = 19.0$, $P < .001$). Most dramatically, parallelism was lowest for trophic traits in unpaired but not paired designs and highest for traits

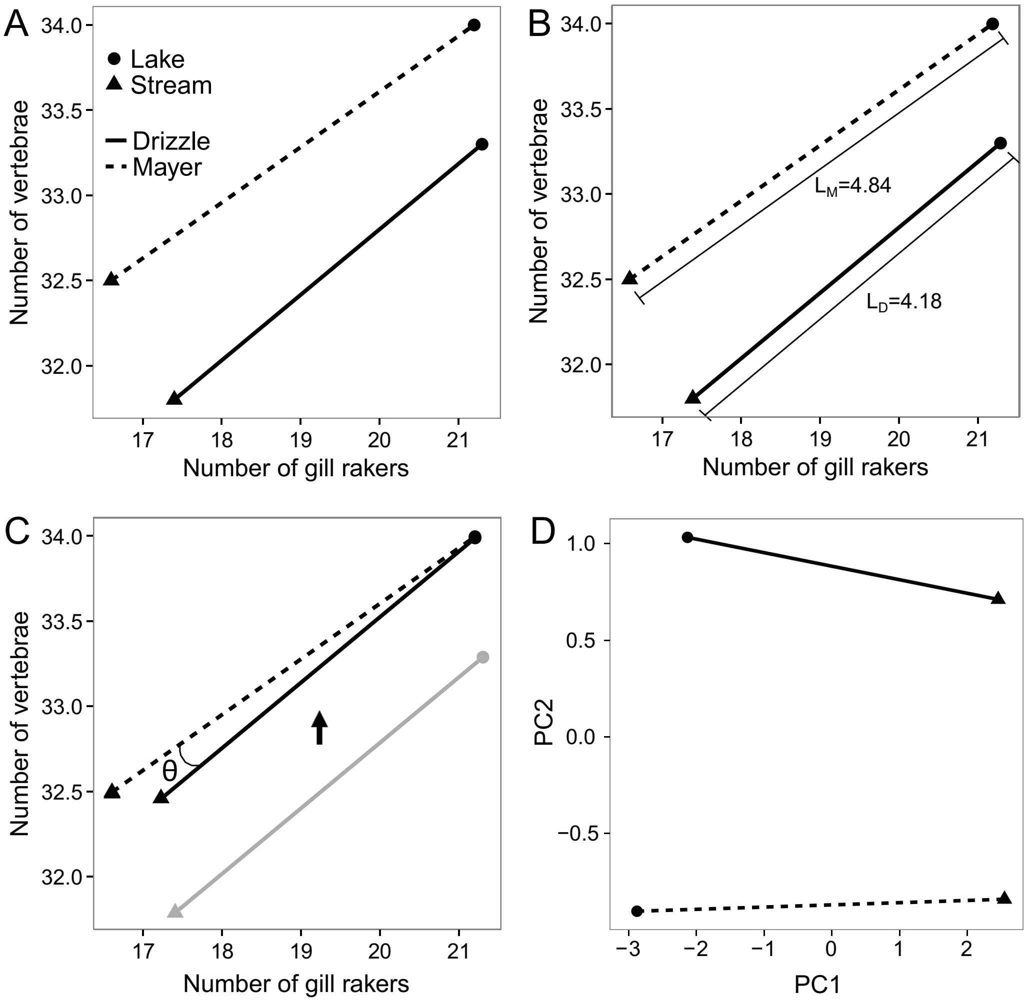


Figure 1: Phenotypic change vector analysis (Collyer and Adams 2007; Adams et al. 2009) is used to quantify the extent to which ecotype pairs show parallel divergence between environments, as demonstrated here using data on lake-stream stickleback from the Drizzle and Mayer watersheds in Haida Gwaii (Reimchen et al. 1985). *A*, Using only two traits for ease of visualization, we project the mean trait value for each population (two ecotypes in each of two watersheds) into trait space and calculate the phenotypic change vectors within each watershed. *B*, The smaller the difference in vector lengths ($L = L_M - L_D$), the more similar the magnitude of the phenotypic divergence. In this case, the magnitude of divergence is very similar in each watershed. *C*, The smaller the angle, θ, between the vectors, the more similar is the direction of phenotypic divergence. In this case, the direction of divergence is very similar. *D*, To visualize higher-dimensional data (i.e., more than two traits) in two dimensions, we plot the phenotypic vectors in principal component (PC) space. Shown here are the vectors based on the first two PCs extracted from five traits (numbers of gill rakers, lateral plates, vertebrae, and anal and dorsal fin rays), which again suggests high parallelism in this system.

in the category of other in unpaired but not paired designs. The main effect of study design did not clearly influence the extent of parallelism for wild-caught ($\chi^2_1 = 2.66$, $P = .103$) or common-garden ($\chi^2_1 = 3.08$, $P = .079$; fig. 3*B*) fish. As noted in "Methods," low sample size for common-garden fish precluded statistical assessment of the potential effects of trait type. Correcting R^2 values for the amount of parallelism expected by chance did not influence our results, except for the (far fewer) common-garden studies. The results of the mixed models on the residual R^2 values did not differ from these results for wild-caught fish (trait type: $\chi^2_4 = 16.9$, $P = .002$; study design: $\chi^2_1 = 0.002$, $P = .964$; interaction: $\chi^2_4 = 25.6$, $P < .001$), but unpaired designs had significantly higher residual R^2 than paired designs ($\chi^2_1 = 9.58$, $P = .002$). Thus, it appears that sample size did not affect our results, except perhaps for traits measured under common-garden conditions, which were themselves relatively rare. Whether the original study included a prediction of parallelism, convergence, or neither did not significantly influence R^2 values ($F_{2,612} = 1.33$, $P = .264$), nor did whether the original study called a given trait parallel, convergent, or neither ($F_{3,612} = 0.342$, $P = .795$).

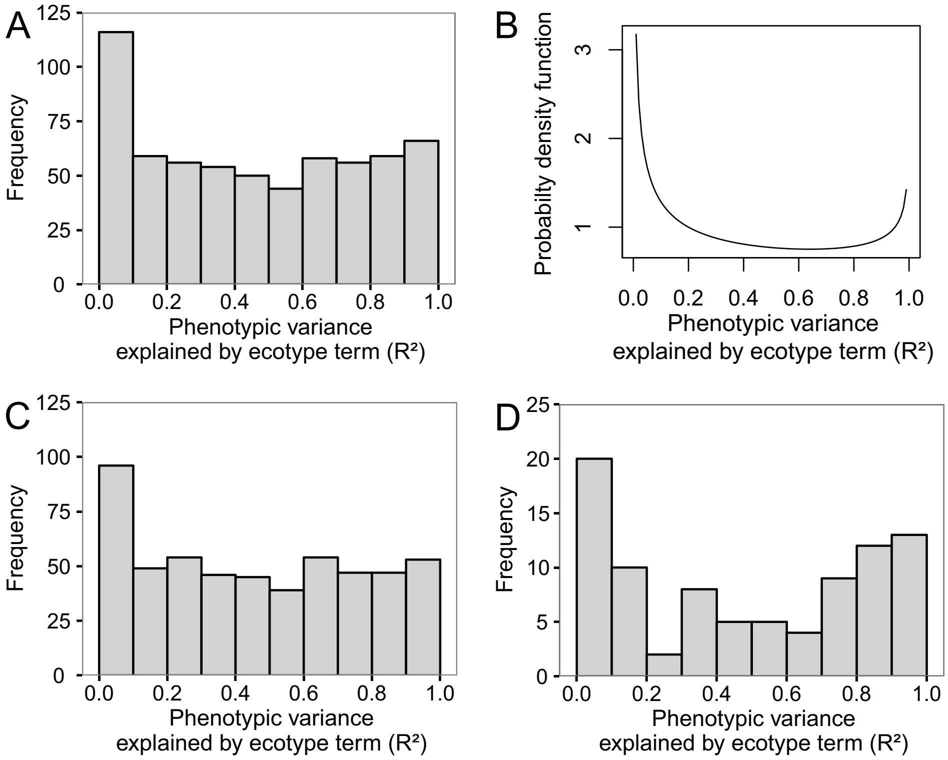


Figure 2: The extent of phenotypic parallelism in fishes is highly variable, as demonstrated by (*A*) a frequency histogram of the proportion of phenotypic variation (R^2) explained by the ecotype term in our ANOVAs of 618 traits and (*B*) a probability density function for the same R^2 values. Broadly similar conclusions emerge when dividing the data into (*C*) wild-caught fish and (*D*) common-garden fish.

PCVA revealed highly variable parallelism across ecotype pairs, with variation resulting from both the direction (angles) and the magnitude (difference in lengths) of divergence (fig. 4). The mean relative (proportional) difference in vector lengths was 0.299 ± 0.240, with a fitted beta distribution having shape parameters of $\alpha = 0.876 \pm 0.0766$ and $\beta = 2.01 \pm 0.203$. The mean difference in directions between ecotype pairs was 44.2° ± 44.8°, corresponding to a relative (proportional) difference in direction of 0.245 ± 0.249, with a fitted beta distribution having shape parameters of $\alpha = 0.527 \pm 0.0441$ and $\beta = 1.42 \pm 0.148$. Thus, the phenotypic change vector for the average ecotype pair differed from other ecotype pairs in the same study system by about 25% of their combined total length and by about 44° (a difference of 90° indicates orthogonal change vectors, and a difference of 180° indicates vectors pointing in opposite directions, sometimes referred to as antiparallel evolution). Both distributions were biased toward lower values (fig. 5), indicating a mixture of many instance of strong parallelism combined with a few instances of very weak parallelism (fig. 4). The latter were striking, with 36.9% of the differences in directions >45° and 16.7% in directions >90°. The extent of parallelism did not differ across trait types in either magnitude or direction (length: $\chi^2_4 = 3.10$, $P = .541$; direction: $\chi^2_4 = 6.96$, $P = .138$). Neither did the extent of parallelism differ when comparing results based on vector length with those based on direction ($\chi^2_1 = 0.819$, $P = .366$).

PTA results for the (fewer) studies with sufficient data generally paralleled the above PCVA results, although parallelism was somewhat lower. The mean relative pairwise difference in trajectory lengths was 0.237 ± 0.218. The mean relative difference in direction was 88.8° ± 77.2°, corresponding to a relative (proportional) difference in direction of 0.493 ± 0.429. The mean pairwise shape difference was 0.778 ± 0.415. Deviations from parallelism were particularly frequent for directions (fig. 5), with no differences falling between 45° and 90° while 54.2% of the differences were >90°.

Discussion

Our aim was to determine, quantitatively, just how parallel parallel evolution is. We investigated this question in studies of parallel evolution from the fish literature. Overall, our results suggest that the extent of parallelism was highly var-

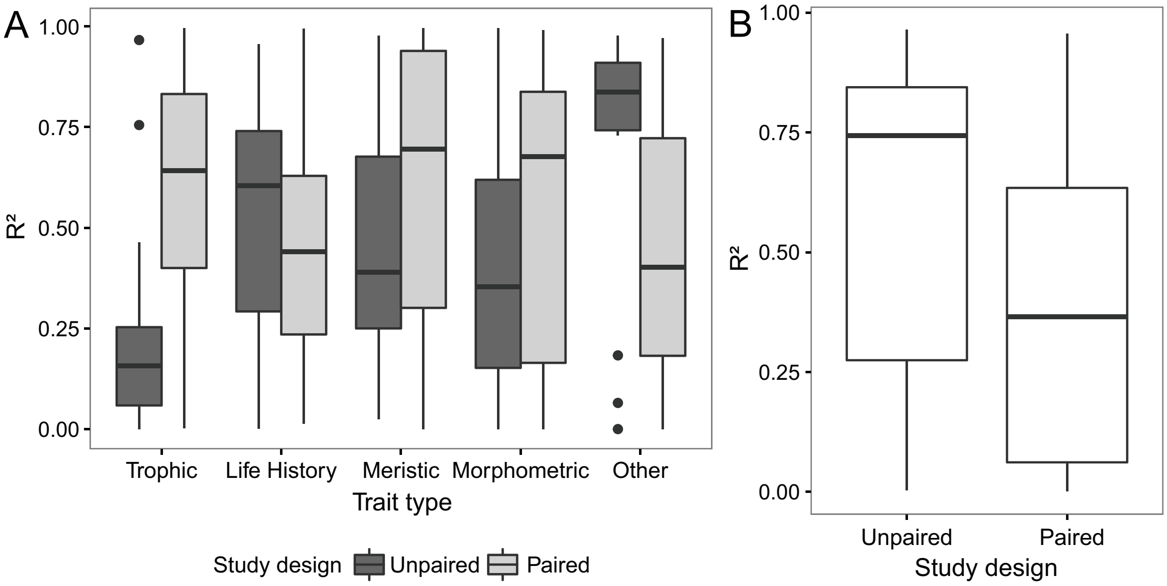


Figure 3: *A*, In wild-caught fish, the effect of the study design (paired or unpaired) on parallelism (R^2) depends on the trait type. *B*, In common-garden fish, parallelism was higher for unpaired designs, although sample sizes were too low to assess interactions with trait type.

iable, and instances of weak parallelism were frequently observed. In the "Discussion" sections of the 92 articles we analyzed, 38 described the observed patterns as parallel, and eight described them as convergent. Yet these assertions were only rarely accompanied by any formal quantification of parallelism, and indeed our analyses revealed that the extent of parallelism ranged from very low to very high (figs. 2–5). Surprisingly, traits for which original studies did or did not predict parallelism (or convergence) or that did or did not call the observed patterns parallel did not differ in the amount of phenotypic variation explained by habitat categories. We argue that all studies seeking to infer phenotypic parallelism (inclusive of the related terms "convergence," "repeatability," and "predictability") should explicitly quantify the extent of parallelism, which then can be compared to the distribution of outcomes reported here. Doing so will allow investigators to report where on the parallelism continuum their results lie in comparison to other systems. Of particular interest from our compilation, many studies inferring the existence of phenotypic parallelism actually included substantial, sometimes predominant, weakly parallel components. Variation in the extent of parallelism should therefore provide a good substrate for inferences about a diversity of evolutionary forces (see the introduction).

Methodological Considerations

We explored several different ways of quantifying and comparing the extent of parallelism. In general, vector comparisons (directions and lengths) in PCVA/PTA suggested higher degrees of parallelism than did the percentage of variance among population means explained by the ecotype term in ANOVA. We postulate several potential explanations for this difference. First, only studies with paired designs (typically sympatric/parapatric population pairs) could be included in PCVA/PTA, and divergence in such cases could be especially parallel. For example, sympatric/parapatric populations might be more parallel for methodological reasons (e.g., more likely to be selected on the basis of prior expectations of parallelism) or biological reasons (e.g., character displacement could enhance divergence). However, we did not find statistical evidence that parallelism was lower in unpaired than paired designs, except perhaps for some trait types (fig. 3). Second, divergence in multidimensional trait space (PCVA/PTA) could be more parallel than divergence in univariate trait space (ANOVAs). Interestingly, PTA (more than two ecotypes per analysis) suggested less parallelism than did PCVA (two ecotypes per analysis). This difference might arise if more opportunities for deviations from strict parallelism exist as the number of ecotype categories increases or if a finer parsing of ecotype categories means that they become less discrete (although the opposite is also possible). In addition, although our results revealed higher parallelism in higher-dimensional trait spaces, it was not necessarily higher-dimensional traits that were considered, as opposed to several related but separate traits. It would be interesting to investigate further whether the dimensionality of traits influences their parallelism, as might be especially true if alternative solutions to similar selective challenges exist (Gould and Lewontin 1979; Arnold 1983; Alfaro et al. 2004; Wainwright et al. 2005; Pfen-

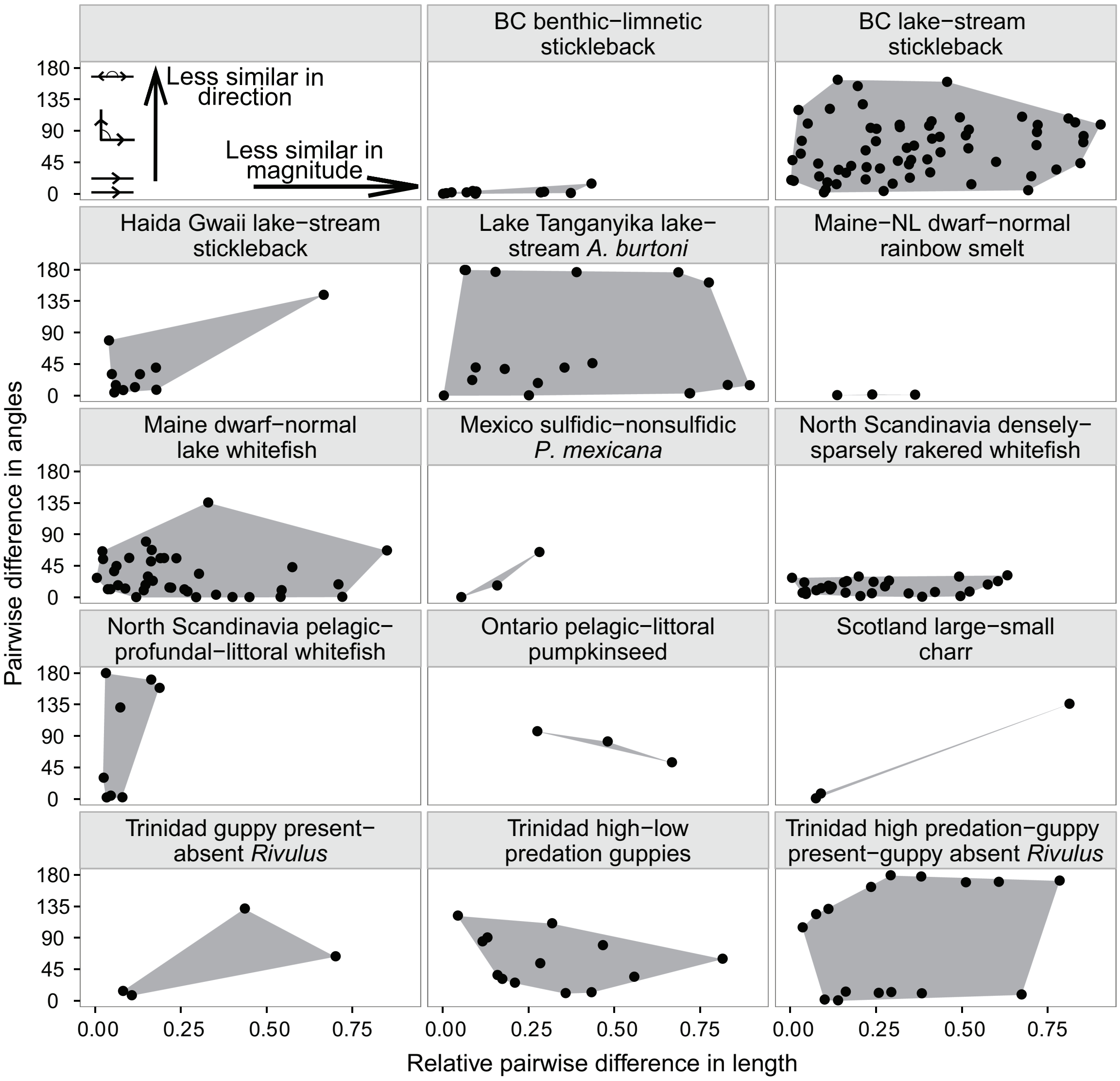


Figure 4: Pairwise estimates of differences in vector directions and relative differences in vectors lengths between ecotype pairs (from study systems with more than two ecotype pairs) reveal that the extent of parallelism varies greatly among both ecotype pairs and study systems. Each data point in this figure is a different pairwise comparison between two ecotype pairs. All panels are based on phenotypic change vector analysis (PCVA) except for three (Trinidad high-predation–guppy-present–guppy-absent *Rivulus*, Maine dwarf-normal lake whitefish, and North Scandinavia pelagic-profundal-littoral whitefish) that are based on phenotypic trajectory analysis (PTA). Only paired study systems could be included in PCVA/PTA (see "Methods"). BC = British Columbia; NL = Newfoundland.

ninger et al. 2014). Third, analyses based on variance explained in ANOVA necessarily combine directions and magnitudes of divergence, which can be analyzed separately using PCVA/PTA. Hence, variation in the magnitude of divergence could lower R^2-based estimates of parallelism, despite similarity in the direction of divergence. Alternatively, large effect sizes for the environment term might be driven by large magnitudes of divergence even in the absence of similarity in the direction of divergence (i.e., a large effect size for an environment term for a wide range of phenotypic change vector angles).

Comparing results from those study systems for which both analysis types were possible (fig. A2), some support emerges for the above possibilities. Direct comparisons are difficult because R^2 results are obtained for each trait (combining all populations), whereas differences in vector angles and lengths are obtained for each pairwise comparison of ecotype pairs (combining multiple traits). However, comparing frequency

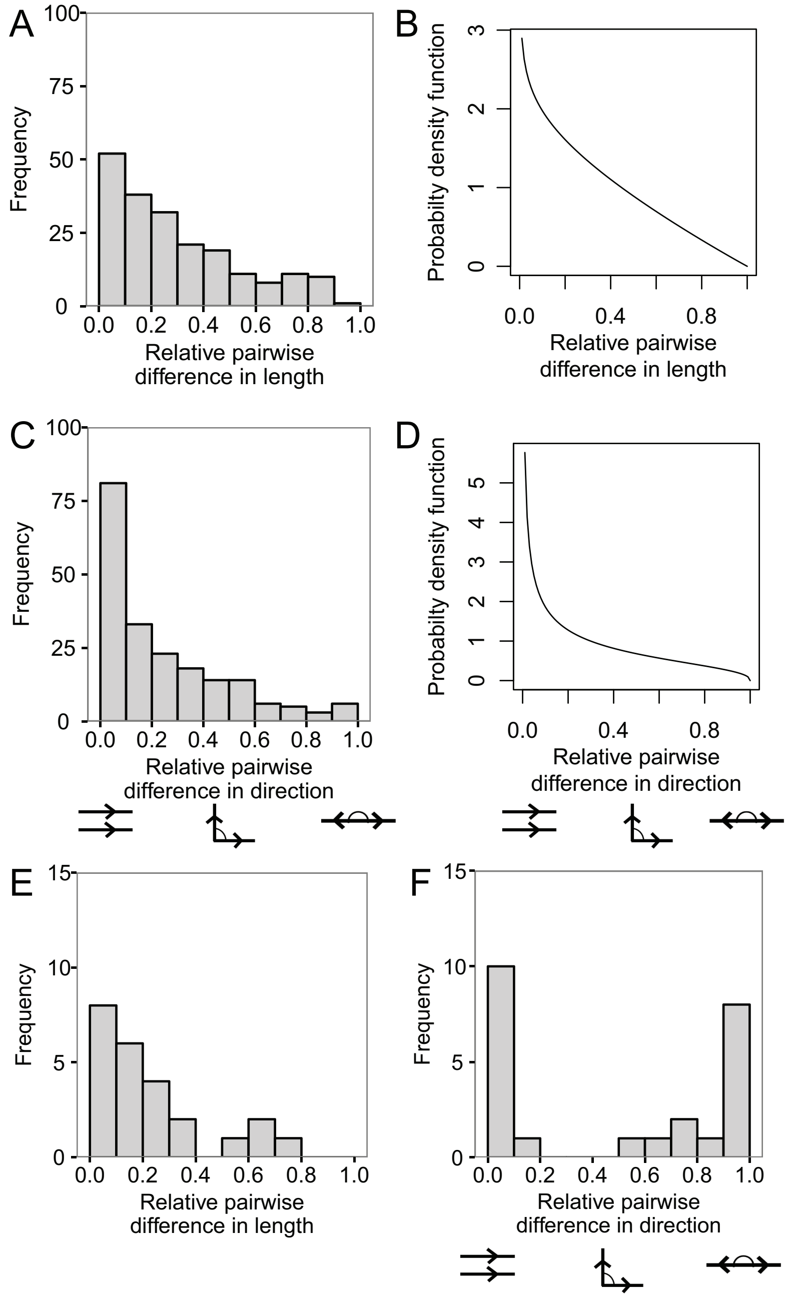


Figure 5: The extent of phenotypic parallelism in fishes is highly variable, as demonstrated by the frequency of relative pairwise difference in (*A*, *E*) vector length and (*C*, *F*) direction results from (*A*, *C*) phenotypic change vector analysis (PCVA) and (*E*, *F*) phenotypic trajectory analysis (PTA; note the difference in axes between PCVA and PTA due to the lower sample size for PTA). The probability density functions for results for pairwise differences in (*B*) lengths and (*D*) directions from PCVA show broadly similar conclusions.

histograms of results from each method within study systems yields some insights (fig. A2). For some study systems, such as Mexico sulfidic-nonsulfidic *Poecilia mexicana*, reasonably strong agreement exists between variance partitioning and vector analyses that parallelism is high: most R^2 values are high, as is the similarity between vectors. In other cases, such as Haida Gwaii lake-stream stickleback and Maine dwarf-normal lake whitefish, agreement is less strong: although most divergence vectors were highly similar, R^2 values for most traits were low. Hence, the different methods are not redundant and instead yield complementary insights. As more data accumulate, more explicit analyses of similarities and differences between the different analysis approaches and the inferences they suggest will likely become more valuable.

Future work could improve on the analyses performed here. First, we did not use formal meta-analytical methods owing to limited availability of the necessary data. Second, the inaccessibility of individual-level data prior to online databases precluded other types of analyses, such as those listed in the paragraph below, and it also prevented statistical tests of specific differences in vector directions and lengths (details in Collyer and Adams 2007; Adams et al. 2009). Third, we considered only fishes, and so the generality of our results remains to be determined. Although we might expect many of the processes shaping the extent of parallelism to be similar across taxa, some differences also seem possible. As just one example, parallelism might be lower for freshwater fishes or island populations than for mainland terrestrial or marine populations because greater isolation of the former could promote site-specific drift and founder effects.

We advocate an integrated approach to quantifying and comparing parallelism. Variance explained—both among population means and among all individuals (e.g., Kaeuffer et al. 2012; McCairns and Bernatchez 2012*a*)—provides an intuitive effect size measure that is directly comparable to analyses of all sorts of phenomena. Similarly, partitioning variation into components that are shared across lineages or unique to individual lineages (and their interaction), as advocated by Langerhans and DeWitt (2004), can provide valuable insights (Langerhans and DeWitt 2004; Langerhans et al. 2006; Franssen et al. 2013). In addition, analyses of individual exchangeability (e.g., discriminant analysis) can strengthen inferences about parallelism by considering the entire trait distribution (e.g., Franssen et al. 2013; Hendry et al. 2013*a*, 2013*b*; Oke et al. 2016). Finally, PCVA/PTA allows further quantification of the extent of parallelism by partitioning parallelism into similarity in the direction and magnitude of divergence (Collyer and Adams 2007; Berner et al. 2008; Adams et al. 2009; Adams 2010). For study systems with more than two habitat categories, PTA also allows comparison of trajectory shape (Collyer and Adams 2007; Adams et al. 2009; Ingley et al. 2014*b*). Although vector directions and length yielded similar overall estimates of the extent of parallelism in our data set (PTA sample size was too low for meaningful shape comparisons), the causal drivers of variation in change vector direction could well be different from the causal drivers of variation in magnitude (or shape).

In the current study, we have included both variance explained and PCVA/PTA methods, in part because of more limited data availability for PCVA/PTA methods and in part because there are no developed PCVA/PTA methods for unpaired study designs and the vast majority of studies make inferences of parallelism on the basis of variance-explained methods. However, we strongly advocate that when evidence of parallelism is detected using the variance-explained method, this method be supported by further analyses, such as PCVA/PTA, that can distinguish between similarity in the direction and magnitude of divergence. Both of these sources of variation are informative and worthy of further investigation.

Explaining (Non)parallelism

Why was (non)parallelism so prevalent in many studies ostensibly documenting parallelism? Here we can return to the suggestions made in the introduction and cast them more directly in the context of our results. First, we analyzed all traits reported in a given study when some of those traits might have been measured for convenience, not because they were expected to experience parallel divergent selection. A general lack of clear predictions of whether individual traits were expected to show parallelism in the original studies precluded our ability to remove traits that might be unrelated to divergence or not expected to be under selection. Our inability to exclude potentially unrelated traits could have led to artificially low estimates of parallelism, although this scenario seems unlikely because we did not detect a significant effect on R^2 values of whether the original study predicted a trait to be parallel or called the trait parallel. Nonetheless, future studies would benefit from careful consideration of whether each individual trait has a relevant link to fitness during divergence and should be included in estimates of the extent of parallelism. Second, environment/ecotype categories are often very coarse (e.g., lake vs. stream) and likely subsume considerable variation among sites of a given type, in either the strength or even the presence of selective forces. Indeed, a number of studies have explained deviations from parallelism or quantitative variation in ecotype divergence as a function of quantitative variation in relevant ecological factors (e.g., Robinson et al. 2000; Landry et al. 2007; Matthews et al. 2010; Fitzpatrick et al. 2014, Moore et al. 2016; Riesch et al. 2016). For instance, the magnitude of divergence in trophic traits of lake versus stream stickleback is closely predicted by the magnitude of divergence in their diets (Berner et al. 2008, 2009; Kaeuffer et al. 2012). Third, different ecotype pairs might experience different levels of gene flow that differentially constrain the magnitude of divergence. This association also has been reported for lake-stream stickleback (Hendry and Taylor 2004) and for a number of other organisms (Hendry 2017). Fourth, evolutionary histories and/or constraints could drive deviations from strict parallelism across ecotype pairs (Travisano et al. 1995; Price et al. 2000; Langerhans and DeWitt 2004) even in the face of parallel selection. For example, allometric constraint has played an important role in marine-freshwater divergence in threespine stickleback in Norway (Voje et al. 2013), but even in such a well-studied species the generality of these results remains to be seen. Fifth, some of the traits analyzed (e.g., male guppy color) are also subject to sexual selection, which could confound or complicate the predictability of natural selection. Sixth, many of the traits—especially trophic traits (Alfaro et al. 2004; Wainwright et al. 2005)—can influence performance in multiple ways, and so traits and trait combinations might differ even when performance is similar. These examples suggest the value of explicitly recognizing and considering (non)parallelism as a profitable avenue to evolutionary insight.

Of course, recognizing the drivers but also the consequences of deviations from strict parallelism is likely to be fruitful. For example, a recent study of cases of speciation observed that nonparallel responses had been detected in 30 of the 43 study systems (Langerhans and Riesch 2013). Not only parallelism but also deviations from it might promote reproductive isolation and eventual speciation (Langerhans and Riesch 2013).

In our study, the influence of trait type on the extent of parallelism differed across analyses: R^2 results were influenced by trait type and a strong interaction between trait type and study design (fig. 3), while PCVA results were not. The interaction between trait type and study design has no obvious biological explanation, and in combination with the conflicting results from PCVA may indicate that our sample size is too small for generalizations about whether certain traits are more likely to be parallel.

Comparisons of the extent of parallelism for wild-caught fish versus common-garden fish can yield insights into the role played by plasticity in phenotypic parallelism. If plasticity increases parallelism, we might expect greater parallelism in wild-caught fish. If plasticity decreases parallelism, we might expect greater parallelism in common-garden fish. In our analysis, the extent of parallelism was similar in the two study types: mean $R^2 = 0.456 \pm 0.311$ in wild-caught fish and $R^2 = 0.480 \pm 0.351$ in common-garden fish. To a first approximation, this similarity might suggest that genetic divergence frequently underlies phenotypic parallelism. However, strong inferences require formal analyses of parallelism in wild-caught and common-garden fish from the same populations, with recent examples including work on high-predation–low-predation guppies (Torres-Dowdall et al. 2012) and lake-stream stickleback (Oke et al. 2016).

Conclusions

The extent and nature of phenotypic parallelism is highly variable among studies that are seeking to infer parallel evolution. Indeed, substantial deviations from strict parallelism were often present even in studies inferring parallelism on the basis of a significant environment or ecotype term in a statistical model. If we were to literally follow the inferential goal attending studies of parallel evolution, we might thereby conclude that selection is not very deterministic, repeatable, or predictable. However, we instead suggest that selection might act on a more local scale than that normally considered in studies of parallel evolution, such that selection differs considerably between environments that seem similar. The coarse categorization of populations into habitat types often might be oversimplified, missing subtle or less noticeable environmental variability that drives variation in selection across ecotype pairs. Alternatively, natural selection itself might be deterministic, but phenotypes might have nondeterministic components that result from plasticity, different genetic backgrounds, local variation in sexual selection, and many-to-one phenotype-to-performance mapping. The quantification and analysis of the extent of parallelism shows considerable promise for disentangling the contributions of various forces to diversification.

Acknowledgments

We thank D. Adams, M. Collyer, C. Cunningham, and A. Kaliontzopoulou, whose advice on analyses greatly improved this article. The manuscript was also much improved by feedback from four anonymous reviewers, Editor B. Langerhans at Axios Review, and Editors Y. Michalakis and R. K. Butlin at *The American Naturalist* as well as by thoughtful discussions with students of McGill's BIOL594 course and the Graphos peer writing group. K.B.O. was supported by a Natural Sciences and Engineering Research Council of Canada Canada Graduate Scholarships–Master's (NSERC-CGS M) award and an NSERC Postgraduate Scholarships–Doctoral Program (NSERC-PGS D) award.

Literature Cited

Adams, C. E., A. J. Wilson, and M. M. Ferguson. 2008. Parallel divergence of sympatric genetic and body size forms of Arctic charr, *Salvelinus alpinus*, from two Scottish lakes. Biological Journal of the Linnean Society 95:748–757.

Adams, D. C. 2010. Parallel evolution of character displacement driven by competitive selection in terrestrial salamanders. BMC Evolutionary Biology 10:72.

Adams, D. C., M. L. Collyer, and C. Goodnight. 2009. A general framework for the analysis of phenotypic trajectories in evolutionary studies. Evolution 63:1143–1154.

Adams, D. C., and E. Otarola-Castillo. 2013. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution 4:393–399.

Akaike, H. 1987. Factor analysis and AIC. Psychometrika 52:317–332.

Alfaro, M. E., D. I. Bolnick, and P. C. Wainwright. 2004. Evolutionary dynamics of complex biomechanical systems: an example using the four-bar mechanism. Evolution 58:495–503.

Arendt, J., and D. Reznick. 2008. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends in Ecology and Evolution 23:26–32.

Arnold, S. J. 1983. Morphology, performance and fitness. American Zoologist 23:347–361.

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

Berner, D., D. C. Adams, A.-C. Grandchamp, and A. P. Hendry. 2008. Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. Journal of Evolutionary Biology 21:1653–1665.

Berner, D., A. C. Grandchamp, and A. P. Hendry. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. Evolution 63:1740–1753.

Bolnick, D. I., and P. Nosil. 2007. Natural selection in populations subject to a migration load. Evolution 61:2229–2243.

Bonduriansky, R. 2011. Sexual selection and conflict as engines of ecological diversification. American Naturalist 178:729–745.

Brinsmead, J., and M. G. Fox. 2002. Morphological variation between lake- and stream-dwelling rock bass and pumpkinseed populations. Journal of Fish Biology 61:1619–1638.

Clarke, B. 1975. Contribution of ecological genetics to evolutionary theory: detecting direct effects of natural-selection on particular polymorphic loci. Genetics 79:101–113.

Collyer, M. L., and D. C. Adams. 2007. Analysis of two-state multivariate phenotypic change in ecological studies. Ecology 88:683–692.

Delignette-Muller, M. L., and C. Dutang. 2015. fitdistrplus: an R package for fitting distributions. Journal of Statistical Software 64:1–34.

Evans, M. L., L. J. Chapman, I. Mitrofanov, and L. Bernatchez. 2013. Variable extent of parallelism in respiratory, circulatory, and neurological traits across lake whitefish species pairs. Ecology and Evolution 3:546–557.

Fisher, R. A. 1922. On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London A 222:309–368.

Fitzpatrick, S. W., J. Torres-Dowdall, D. N. Reznick, C. K. Ghalambor, and W. C. Funk. 2014. Parallelism isn't perfect: could disease and flooding drive a life-history anomaly in Trinidadian guppies? American Naturalist 183:290–300.

Fox, J., and S. Weisberg. 2011. An R companion to applied regression. 2nd ed. Sage, Thousand Oaks, CA.

Franssen, N. R., J. Harris, S. R. Clark, J. F. Schaefer, and L. K. Stewart. 2013. Shared and unique morphological responses of stream fishes to anthropogenic habitat alteration. Proceedings of the Royal Society B 280:20122715.

Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critque of the adaptionist programme. Proceedings of the Royal Society B 205:581–598.

Hendry, A. P. 2017. Eco-evolutionary dynamics. Princeton University Press, Princeton, NJ.

Hendry, A. P., T. Day, and E. B. Taylor. 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. Evolution 55:459–466.

Hendry, A. P., A. S. Hendry, and C. A. Hendry. 2013*a*. Hendry Vineyard stickleback: testing for contemporary lake-stream divergence. Evolutionary Ecology Research 15:343–359.

Hendry, A. P., R. Kaeuffer, E. Crispo, C. L. Peichel, and D. I. Bolnick. 2013*b*. Evolutionary inferences from the analysis of exchangeability. Evolution 67:3429–3441.

Hendry, A. P., and E. B. Taylor. 2004. How much of the variation in adaptive divergence can be explained by gene flow? an evaluation using lake-stream stickleback pairs. Evolution 58:2319–2331.

Herczeg, G., M. Turtiainen, and J. Merilä. 2010. Morphological divergence of North-European nine-spined sticklebacks (*Pungitius pungitius*): signatures of parallel evolution. Biological Journal of the Linnean Society 101:403–416.

Hoekstra, H. E., and M. W. Nachman. 2003. Different genes underlie adaptive melanism in different populations of rock pocket mice. Molecular Ecology 12:1185–1194.

Huttegger, S. M., and P. Mitteroecker. 2011. Invariance and meaningfulness in phenotype spaces. Evolutionary Biology 38:335–351.

Ingley, S. J., E. J. Billman, M. C. Belk, and J. B. Johnson. 2014*a*. Morphological divergence driven by predation environment within and between species of *Brachyrhaphis* fishes. PLoS ONE 9:e90274.

Ingley, S. J., E. J. Billman, C. Hancock, and J. B. Johnson. 2014*b*. Repeated geographic divergence in behavior: a case study employing phenotypic trajectory analyses. Behavioral Ecology and Sociobiology 68:1577–1587.

Jennions, M. D., and S. R. Telford. 2002. Life-history phenotypes in populations of *Brachyrhaphis episcopi* (Poeciliidae) with different predator communities. Oecologia 132:44–50.

Johnson, J. 2001. Adaptive life-history evolution in the livebearing fish *Brachyrhaphis rhabdophora*: genetic basis for parallel divergence in age and size at maturity and a test of predator-induced plasticity. Evolution 55:1486–1491.

Kaeuffer, R., C. L. Peichel, D. I. Bolnick, and A. P. Hendry. 2012. Parallel and nonparallel aspects of ecological phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. Evolution 66:402–418.

Kavanagh, K. D., T. O. Haugen, F. Gregersen, J. Jernvall, and L. A. Vollestad. 2010. Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. BMC Evolutionary Biology 10:1–12.

Landry, L., and L. Bernatchez. 2010. Role of epibenthic resource opportunities in the parallel evolution of lake whitefish species pairs (*Coregonus* sp.). Journal of Evolutionary Biology 23:2602–2613.

Landry, L., W. F. Vincent, and L. Bernatchez. 2007. Parallel evolution of lake whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. Journal of Evolutionary Biology 20:971–984.

Langerhans, R. B., and T. J. DeWitt. 2004. Shared and unique features of evolutionary diversification. American Naturalist 164: 335–349.

Langerhans, R. B., J. H. Knouft, and J. B. Losos. 2006. Shared and unique features of diversification in Greater Antillean *Anolis* ecomorphs. Evolution 60:362–369.

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

Langerhans, R. B., and R. Riesch. 2013. Speciation by selection: a framework for understanding ecology's role in speciation. Current Zoology 59:31–52.

Losos, J. B. 2011. Convergence, adaptation, and constraint. Evolution 65:1827–1840.

Maan, M. E., and O. Seehausen. 2011. Ecology, sexual selection and speciation. Ecology Letters 14:591–602.

Matthews, B., K. B. Marchinko, D. I. Bolnick, and A. Mazumder. 2010. Specialization of trophic position and habitat use by sticklebacks in an adaptive radiation. Ecology 91:1025–1034.

McCairns, R. J. S., and L. Bernatchez. 2012*a*. Plasticity and heritability of morphological variation within and between parapatric stickleback demes. Journal of Evolutionary Biology 25:1097–1112.

Moore, M. P., R. Riesch, and R. A. Martin. 2016. The predictability and magnitude of life-history divergence to ecological agents of selection: a meta-analysis in livebearing fishes. Ecology Letters 19:435–442.

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59:181–197.

Oke, K. B., M. Bukhari, R. Kaeuffer, G. Rolshausen, K. Räsänen, D. I. Bolnick, C. L. Peichel, and A. P. Hendry. 2016. Does plasticity enhance or dampen phenotypic parallelism? a test with three lake-stream stickleback pairs. Journal of Evolutionary Biology 29:126–143.

Oke, K. B., G. Rolshausen, C. LeBlond, and A. P. Hendry. 2017. Data from: How parallel is parallel evolution? a comparative analysis in fishes. American Naturalist, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.bb206.

Østbye, K., P. A. Amundsen, L. Bernatchez, A. Klemetsen, R. Knudsen, R. Kristoffersen, T. F. Næsje, and K. Hindar. 2006. Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times. Molecular Ecology 15:3983–4001.

Østbye, K., T. F. Næsje, L. Bernatchez, O. T. Sandlund, and K. Hindar. 2005. Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in Lake Femund, Norway. Journal of Evolutionary Biology 18:683–702.

Palkovacs, E. P., K. B. Dion, D. M. Post, and A. Caccone. 2008. Independent evolutionary origins of landlocked alewife populations and rapid parallel evolution of phenotypic traits. Molecular Ecology 17:582–597.

Pfenninger, M., H. Lerp, M. Tobler, C. Passow, J. L. Kelley, E. Funke, B. Greshake, et al. 2014. Parallel evolution of *cox* genes in H2S-tolerant fish as key adaptation to a toxic environment. Nature Communications 5:3873.

Price, T., I. J. Lovette, E. Bermingham, H. L. Gibbs, and A. D. Richman. 2000. The imprint of history on communities of North American and Asian warblers. American Naturalist 156:354–367.

Ravinet, M., P. A. Prodöhl, and C. Harrod. 2013*a*. Parallel and nonparallel ecological, morphological and genetic divergence in lake-stream stickleback from a single catchment. Journal of Evolutionary Biology 26:186–204.

R Development Core Team. 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Reimchen, T. E., E. M. Stinson, and J. S. Nelson. 1985. Multivariate differentiation of parapatric and allopatric populations of threespine stickleback in the Sangan River watershed, Queen Charlotte Islands. Canadian Journal of Zoology 63:2944–2951.

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. N., F. H. Rodd, and M. Cardenas. 1996. Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. American Naturalist 147:319–338.

Riesch, R., R. A. Martin, and R. B. Langerhans. 2013. Predation's role in life-history evolution of a livebearing fish and a test of the Trexler-DeAngelis model of maternal provisioning. American Naturalist 181:78–93.

Riesch, R., M. Tobler, H. Lerp, J. Jourdan, T. Doumas, P. Nosil, R. B. Langerhans, et al. 2016. Extremophile Poeciliidae: multivariate insights into the complexity of speciation along replicated ecological gradients. BMC Evolutionary Biology 16:136.

Robinson, B. W., D. S. Wilson, and A. S. Margosian. 2000. A pluralistic analysis of character release in pumpkinseed sunfish (*Lepomis gibbosus*). Ecology 81:2799–2812.

Rosenblum, E. B., and L. J. Harmon. 2010. "Same same but different": replicated ecological speciation at white sands. Evolution 65:946–960.

Schluter, D., E. A. Clifford, M. Nemethy, and J. S. McKinnon. 2004. Parallel evolution and inheritance of quantitative traits. American Naturalist 163:809–822.

Schluter, D., and J. D. McPhail. 1992. Ecological character displacement and speciation in sticklebacks. American Naturalist 140:85–108.

Schwarz, G. 1978. Estimating the dimension of a model. Annals of Statistics 6:461–464.

Taylor, E. B., S. Harvey, S. Pollard, and J. Volpe. 1997. Postglacial genetic differentiation of reproductive ecotypes of kokanee *Oncorhynchus nerka* in Okanagan Lake, British Columbia. Molecular Ecology 6: 503–517.

Taylor, E. B., and P. Bentzen. 1993. Evidence for multiple origins and sympatric divergence of trophic ecotypes of smelt (*Osmerus*) in northeastern North America. Evolution 47:813–832.

Theis, A., F. Ronco, A. Indermaur, W. Salzburger, and B. Egger. 2014*a*. Adaptive divergence between lake and stream populations of an East African cichlid fish. Molecular Ecology 23:5304–5322.

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

Torres-Dowdall, J., C. A. Handelsman, D. N. Reznick, and C. K. Ghalambor. 2012. Local adaptation and the evolution of phenotypic plasticity in Trinidadian guppies (*Poecilia reticulata*). Evolution 66:3432–3443.

Travisano, M., J. A. Mongold, A. F. Bennett, and R. E. Lenski. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. Science 267:87–90.

Voje, K. L., A. B. Mazzarella, T. F. Hansen, K. Østbye, T. Klepaker, A. Bass, A. Herland, et al. 2013. Adaptation and constraint in a stickleback radiation. Journal of Evolutionary Biology 26:2396–2414.

Wainwright, P. C., M. E. Alfaro, D. I. Bolnick, and C. D. Hulsey. 2005. Many-to-one mapping of form to function: a general principle in organismal design? Integrative Comparative Biology 45:256–262.

Wake, D. B., M. H. Wake, and C. D. Specht. 2011. Homoplasy: from detecting pattern to determining process and mechanism of evolution. Science 331:1032–1035.

Woolf, B. 1957. The log likelihood ratio test (the *G*-test). Annals of Human Genetics 21:397–409.

References Cited Only in the Online Appendixes

Alekseyev, S. S., V. P. Samusenok, A. N. Matveev, and M. Yu. 2002. Diversification, sympatric speciation and trophic polymorphism of Arctic charr, *Salvelinus alpinus* complex, in Transbaikalia. Environmental Biology of Fishes 64:97–114.

Blair, G. R., D. E. Rogers, and T. P. Quinn. 1993. Variation in life history characteristics and morphology of sockeye salmon in the Kvichak River System, Bristol Bay, Alaska. Transactions of the American Fisheries Society 122:550–559.

Boughman, J. W., H. D. Rundle, and D. Schluter. 2005. Parallel evolution of sexual isolation in sticklebacks. Evolution 59:361–373.

Chappell, M., and J. Odell. 2004. Predation intensity does not cause microevolutionary change in maximum speed or aerobic capacity in Trinidadian guppies (*Poecilia reticulata* Peters). Physiological and Biochemical Zoology 77:27–38.

Dalziel, A. C., T. H. Vines, and P. M. Schulte. 2011. Reductions in prolonged swimming capacity following freshwater colonization in multiple threespine stickleback populations. Evolution 66:1226–1239.

Deagle, B. E., F. C. Jones, Y. F. Chan, D. M. Absher, D. M. Kingsley, and T. E. Reimchen. 2012. Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. Proceedings of the Royal Society B 279:1277–1286.

Egset, C. K., G. H. Bolstad, G. Rosenqvist, J. A. Endler, and C. Pélabon. 2011. Geographical variation in allometry in the guppy (*Poecilia reticulata*). Journal of Evolutionary Biology 24:2631–2638.

Eizaguirre, C., T. L. Lenz, R. D. Sommerfeld, C. Harrod, M. Kalbe, and M. Milinski. 2011. Parasite diversity, patterns of MHC II variation and olfactory based mate choice in diverging three-spined stickleback ecotypes. Evolutionary Ecology 25:605–622.

Endler, J. A. 1982. Convergent and divergent effects of natural selection on color patterns in two fish faunas. Evolution 36:178–188.

Evans, M. L., and L. Bernatchez. 2012*a*. Oxidative phosphorylation gene transcription in whitefish species pairs reveals patterns of parallel and nonparallel physiological divergence. Journal of Evolutionary Biology 25:1823–1834.

———. 2012*b*. Data from: Oxidative phosphorylation gene transcription in whitefish species pairs reveals patterns of parallel and nonparallel physiological divergence. Journal of Evolutionary Biology 25: 1823–1834, Dryad Data Repository, http://dx.doi.org/10.5061/dryad .s4kc8.

Evans, M. L., K. Præbel, S. Peruzzi, and L. Bernatchez. 2012. Parallelism in the oxygen transport system of the lake whitefish: the role of physiological divergence in ecological speciation. Molecular Ecology 21:4038–4050.

Fenderson, O. C. 1964. Evidence of subpopulations of lake whitefish, *Coregonus clupeaformis*, involving a dwarfed form. Transactions of the American Fisheries Society 93:77–94.

Ferchaud, A., and M. M. Hansen. 2015. Data from: The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: threespine sticklebacks in divergent environments. Molecular Ecology 25:238–259, Dryad Data Repository, http://dx .doi.org/10.5061/dryad.kp11q.

———. 2016. The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: threespine sticklebacks in divergent environments. Molecular Ecology 25:238–259.

Foster, S. A., R. J. Scott, and W. A. Cresko. 1998. Nested biological variation and speciation. Philosophical Transactions of the Royal Society B 353:207–218.

Furness, A. I., M. R. Walsh, and D. N. Reznick. 2011. Convergence of life-history phenotypes in a Trinidadian killifish (*Rivulus hartii*). Evolution 66:1240–1254.

Gillespie, G. J., and M. G. Fox. 2003. Morphological and life-history differentiation between littoral and pelagic forms of pumpkinseed. Journal of Fish Biology 62:1099–1115.

Gíslason, D., M. M. Ferguson, S. Skúlason, and S. S. Snorrason. 1999. Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). Canadian Journal of Fisheries and Aquatic Sciences 56:2229–2234.

Gonda, A., G. Herczeg, and J. Merilä. 2009. Adaptive brain size divergence in nine-spined sticklebacks (*Pungitius pungitius*)? Journal of Evolutionary Biology 22:1721–1726.

Greenwood, A. K., F. C. Jones, Y. F. Chan, S. D. Brady, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, D. M. Kingsley, and C. L. Peichel. 2011. The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. Heredity 107:155–166.

Harrod, C., J. Mallela, and K. K. Kahilainen. 2010. Phenotype-environment correlations in a putative whitefish adaptive radiation. Journal of Animal Ecology 79:1057–1068.

Herczeg, G., A. Gonda, and J. Merilä. 2009*a*. Evolution of gigantism in nine-spined sticklebacks. Evolution 63:3190–3200.

———. 2009*b*. Predation mediated population divergence in complex behaviour of nine-spined stickleback (*Pungitius pungitius*). Journal of Evolutionary Biology 22:544–552.

Huss, M., J. G. Howeth, J. I. Osterman, and D. M. Post. 2014*a*. Intraspecific phenotypic variation among alewife populations drives parallel phenotypic shifts in bluegill. Proceedings of the Royal Society B 281:20140275.

———. 2014*b*. Data from: Intraspecific phenotypic variation among alewife populations drives parallel phenotypic shifts in bluegill. Proceedings of the Royal Society B 281:20140275, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.57n5j.

Jeukens, J., D. Bittner, R. Knudsen, and L. Bernatchez. 2009. Candidate genes and adaptive radiation: insights from transcriptional adaptation to the limnetic niche among coregonine fishes (*Coregonus* spp., *Salmonidae*). Molecular Biology and Evolution 26:155–166.

Johnson, S. P., S. M. Carlson, and T. P. Quinn. 2006. Tooth size and skin thickness in mature sockeye salmon: evidence for habitat constraints and variable investment between the sexes. Ecology of Freshwater Fish 15:331–338.

Jones, A. W., E. P. Palkovacs, and D. M. Post. 2013. Recent parallel divergence in body shape and diet source of alewife life history forms. Evolutionary Ecology 27:1175–1187.

Kaeuffer, R., C. Peichel, D. Bolnick, and A. Hendry. 2011. Data from: Convergence and non-convergence in ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. Evolution 66:402–418, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.k987h.

Kelly, C. D., J.-G. J. Godin, and J. M. Wright. 1999. Geographic variation in multiple paternity within natural populations of the guppy (*Poecilia reticulata*). Proceedings of the Royal Society B 266:2403–2408.

Kimmel, C. B., W. A. Cresko, P. C. Phillips, B. Ullmann, M. Currey, F. Von Hippel, B. K. Kristjánsson, O. Gelmond, and K. McGuigan. 2011. Independent axes of genetic variation and parallel evolutionary divergence of opercle bone shape in threespine stickleback. Evolution 66:419–434.

Kimmel, C. B., B. Ullmann, C. Walker, C. Wilson, M. Currey, P. C. Phillips, M. A. Bell, J. H. Postlethwait, and W. A. Cresko. 2005. Evolution and development of facial bone morphology in threespine sticklebacks. Proceedings of the National Academy of Sciences of the USA 102:5791–5796.

Kitano, J., S. C. Lema, J. A. Luckenbach, S. Mori, Y. Kawagishi, M. Kusakabe, P. Swanson, and C. L. Peichel. 2010. Adaptive divergence in the thyroid hormone signaling pathway in the stickleback radiation. Current Biology 20:2124–2130.

Kitano, J., S. Mori, and C. L. Peichel. 2012. Reduction of sexual dimorphism in stream-resident forms of three-spined stickleback *Gasterosteus aculeatus*. Journal of Fish Biology 80:131–146.

Kristjánsson, B. K., S. Skúlason, S. S. Snorrason, and D. L. G. Noakes. 2012. Fine-scale parallel patterns in diversity of small benthic Arctic charr (*Salvelinus alpinus*) in relation to the ecology of lava/groundwater habitats. Ecology and Evolution 2:1099–1112.

Langerhans, R. B. 2009*a*. Morphology, performance, fitness: functional insight into a post-Pleistocene radiation of mosquitofish. Biology Letters 5:488–491.

———. 2009*b*. Trade-off between steady and unsteady swimming underlies predator-driven divergence in *Gambusia affinis*. Journal of Evolutionary Biology 22:1057–1075.

Langerhans, R. B., C. A. Layman, and T. J. DeWitt. 2005. Male genital size reflects a tradeoff between attracting mates and avoiding predators in two live-bearing fish species. Proceedings of the National Academy of Sciences of the USA 102:7618–7623.

Laporte, M., S. M. Rogers, E. Normandeau, P.-A. Gagnaire, A. C. Dalziel, J. Chebib, and L. Bernatchez. 2015. RAD-QTL mapping reveals both genome-level parallelism and different genetic architecture underlying the evolution of body shape in lake whitefish (*Coregonus clupeaformis*) species pairs. G3: Genes, Genomes, Genetics 5:1481–1491.

Lavin, P. A., and J. D. McPhail. 1993. Parapatric lake and stream sticklebacks on northern Vancouver Island: disjunct distribution or parallel evolution? Canadian Journal of Zoology 71:11–17.

Leinonen, T., J. M. Cano, and J. Merilä. 2011. Genetic basis of sexual dimorphism in the threespine stickleback *Gasterosteus aculeatus*. Heredity 106:218–227.

Lu, G. Q., and L. Bernatchez. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. Evolution 53:1491–1505.

MacColl, A. D. C., A. El Nagar, and J. de Roij. 2013. The evolutionary ecology of dwarfism in three-spined sticklebacks. Journal of Animal Ecology 82:642–652.

Magurran, A. E., B. H. Seghers, G. R. Carvalho, and P. W. Shaw. 1992. Behavioural consequences of an artificial introduction of guppies (*Poecilia reticulata*) in N. Trinidad: evidence for the evolution of anti-predator behaviour in the wild. Proceedings of the Royal Society B 248:117–122.

McCairns, R. J. S., and L. Bernatchez. 2012*b*. Data from: Plasticity and heritability of morphological variation within and between parapatric stickleback demes. Journal of Evolutionary Biology 25:1097–1112, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.615sc2rv.

McGee, M. D., and P. C. Wainwright. 2012. Data from: Convergent evolution as a generator of phenotypic diversity in threespine stickleback. Evolution 67:1204–1208, Dryad Digital Repository, http://dx.doi.org/10.5061/dryad.rd4j5.

———. 2013. Convergent evolution as a generator of phenotypic diversity in threespine stickleback. Evolution 67:1204–1208.

McGurk, M. D. 2000. Comparison of fecundity-length-latitude relationships between nonanadromous (kokanee) and anadromous sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Zoology 78: 1791–1805.

Millar, N. P., and A. P. Hendry. 2011. Population divergence of private and non-private signals in wild guppies. Environmental Biology of Fishes 94:513–525.

Moser, D., M. Roesti, and D. Berner. 2012. Repeated lake-stream divergence in stickleback life history within a Central European lake basin. PLoS ONE 7:e50620.

O'Steen, S., A. J. Cullum, and A. F. Bennett. 2002. Rapid evolution of escape ability in Trinidadian guppies (*Poecilia reticulata*). Evolution 56:776–784.

Oufiero, C. E., M. R. Walsh, D. N. Reznick, and T. Garland. 2011. Swimming performance trade-offs across a gradient in community composition in Trinidadian killifish (*Rivulus hartii*). Ecology 92:170–179.

Palacios, M., L. Arias-Rodriguez, M. Plath, C. Eifert, H. Lerp, A. Lamboj, G. Voelker, and M. Tobler. 2013*a*. The rediscovery of a long described species reveals additional complexity in speciation patterns of poeciliid fishes in sulfide springs. PLoS ONE 8:e71069.

———. 2013*b*. Data from: The rediscovery of a long described species reveals additional complexity in speciation patterns of poeciliid fishes in sulfide springs. PLoS ONE 8:e71069, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.h8j76.

Park, P. J., and M. A. Bell. 2010. Variation of telencephalon morphology of the threespine stickleback (*Gasterosteus aculeatus*) in relation to inferred ecology. Journal of Evolutionary Biology 23: 1261–1277.

Plath, M., and I. Schlupp. 2008. Parallel evolution leads to reduced shoaling behavior in two cave dwelling populations of Atlantic mollies (*Poecilia mexicana*, Poeciliidae, Teleostei). Environmental Biology of Fishes 82:289–297.

Præbel, K., R. Knudsen, A. Siwertsson, M. Karhunen, K. K. Kahilainen, O. Ovaskainen, K. Østbye, S. Peruzzi, S.-E. Fevolden, and P.-A. Amundsen. 2013. Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. Ecology and Evolution 3:4970–4986.

Quinn, T. P., A. P. Hendry, and L. A. Wetzel. 1995. The influence of life history trade-offs and the size of incubation gravels on egg size variation in sockeye salmon (*Oncorhynchus nerka*). Oikos 74:425–438.

Ravinet, M., P. A. Prodöhl, and C. Harrod. 2013*b*. On Irish stickleback: morphological diversification in a secondary contact zone. Evolutionary Ecology Research 15:271–294.

Robinson, B. W. 2013*a*. Evolution of growth by genetic accommodation in Icelandic freshwater stickleback. Proceedings of the Royal Society B 280:20132197.

———. 2013*b*. Data from: Evolution of growth by genetic accommodation in Icelandic freshwater stickleback. Proceedings of the Royal Society B 280:20132197, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.616n0.

Siwertsson, A., R. Knudsen, K. Præbel, C. E. Adams, J. Newton, and P.-A. Amundsen. 2013. Discrete foraging niches promote ecological, phenotypic, and genetic divergence in sympatric whitefish (*Coregonus lavaretus*). Evolutionary Ecology 27:547–564.

Sullam, K. E., B. E. R. Rubin, C. M. Dalton, S. S. Kilham, A. S. Flecker, and J. A. Russell. 2015. Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. ISME Journal 9:1508–1522.

Theis, A., F. Ronco, A. Indermaur, W. Salzburger, and B. Egger. 2014*b*. Data from: Adaptive divergence between lake and stream populations of an East African cichlid fish. Molecular Ecology 23:5304–5322, Dryad Data Repository, http://dx.doi.org/10.5061/dryad.pp0q1.

Tobler, M., D. M. Alba, L. Arias-Rodriguez, and P. D. Jeyasingh. 2015*a*. Using replicated evolution in extremophile fish to understand diversification in elemental composition and nutrient excretion. Freshwater Biology 61:158–171.

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

Velotta, J. P., S. D. McCormick, R. J. O'Neill, and E. T. Schultz. 2014. Relaxed selection causes microevolution of seawater osmoregulation and gene expression in landlocked alewives. Oecologia 175:1081–1092.

Walsh, M. R., D. F. Fraser, R. D. Bassar, and D. N. Reznick. 2011. The direct and indirect effects of guppies: implications for life-history evolution in *Rivulus hartii*. Functional Ecology 25:227–237.

Walsh, M. R., and D. N. Reznick. 2009. Phenotypic diversification across an environmental gradient: a role for predators and resource availability on the evolution of life histories. Evolution 63:3201–3213.

———. 2010*a*. Experimentally induced life-history evolution in a killifish in response to the introduction of guppies. Evolution 65: 1021–1036.

———. 2010*b*. Influence of the indirect effects of guppies on life-history evolution in *Rivulus hartii*. Evolution 64:1583–1593.

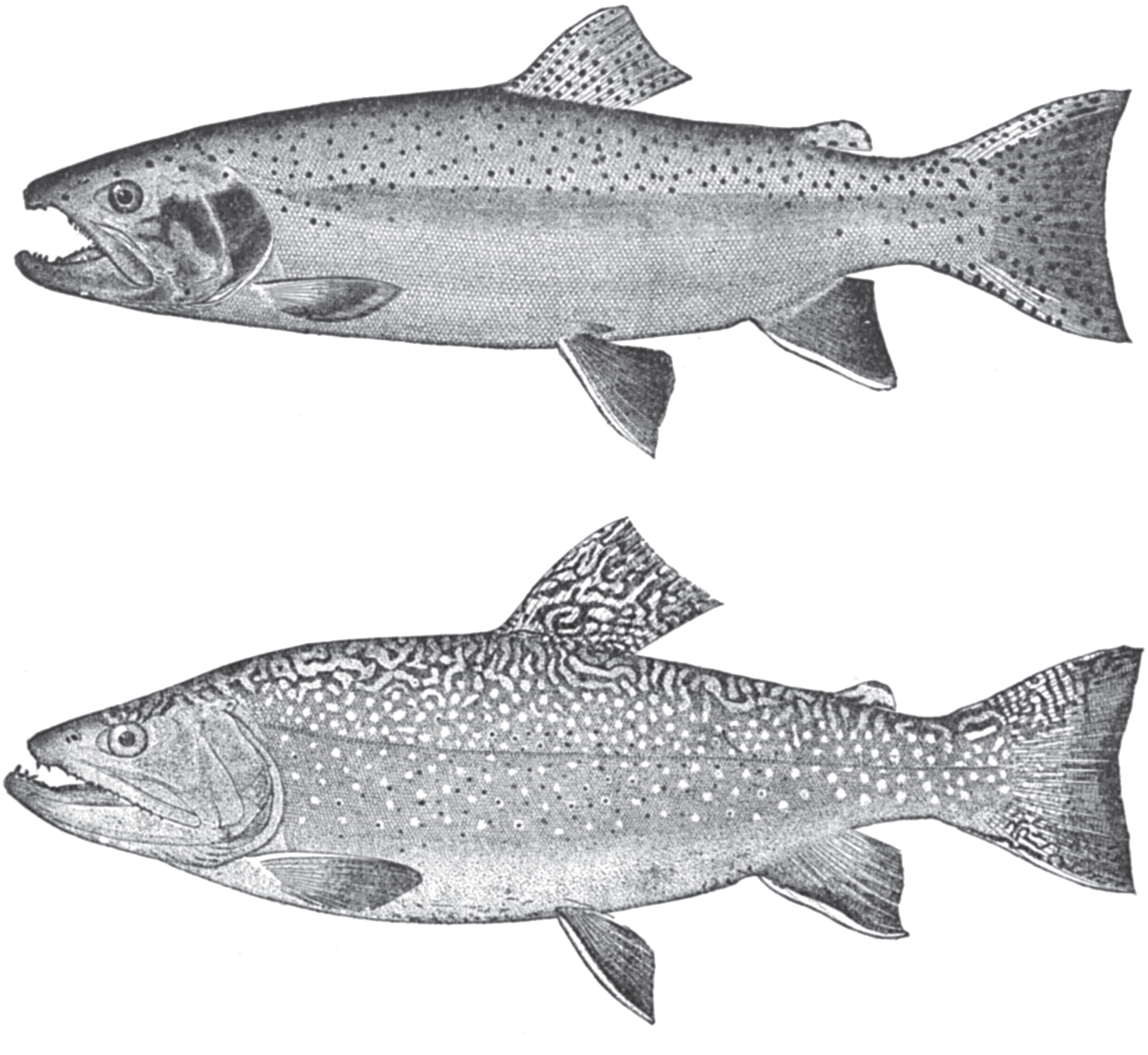
Wark, A. R., and C. L. Peichel. 2010. Lateral line diversity among ecologically divergent threespine stickleback populations. Journal of Experimental Biology 213:108–117.

Wood, C. C., and C. J. Foote. 1996. Evidence for sympatric genetic divergence of anadromous and nonanadromous morphs of sockeye salmon (*Oncorhynchus nerka*). Evolution 50:1265–1279.

Zandonà, E., S. K. Auer, S. S. Kilham, J. L. Howard, A. López-Sepulcre, M. P. O'Connor, R. D. Bassar, A. Osorio, C. M. Pringle, and D. N. Reznick. 2011. Diet quality and prey selectivity correlate with life histories and predation regime in Trinidadian guppies. Functional Ecology 25:964–973.

Zandonà, E., S. K. Auer, S. S. Kilham, and D. N. Reznick. 2015. Contrasting population and diet influences on gut length of an omnivorous tropical fish, the Trinidadian guppy (*Poecilia reticulata*). PLoS ONE 10:e0136079.

Associate Editor: Roger K. Butlin
Editor: Yannis Michalakis



"The family of Salmonidæ—embracing the white fishes, the salmons, and the trouts—is one of the most important of the temperate and arctic regions of the world." Figured: above, Rainbow Trout (*Salmo irideus*), Verona, Missouri; *below*, Brook Trout (*Salvelinus fontinalis*), New York Market. From "Distribution and Some Characters of the Salmonidæ" by Tarleton H. Bean (*The American Naturalist*, 1888, 22:306–314).