

Evolutionary Change in Continuous Reaction Norms

Courtney J. Murren,^{1,*} Heidi J. Maclean,² Sarah E. Diamond,³ Ulrich K. Steiner,⁴
Mary A. Heskell,⁵ Corey A. Handelsman,⁶ Cameron K. Ghalambor,⁶ Josh R. Auld,⁷
Hilary S. Callahan,⁸ David W. Pfennig,² Rick A. Relyea,⁹ Carl D. Schlichting,¹⁰ and Joel Kingsolver²

1. Department of Biology, College of Charleston, Charleston, South Carolina 29424; 2. Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599; 3. Department of Biology, North Carolina State University, Raleigh, North Carolina 27695; 4. Department of Biology and Max-Planck Odense Center on the Biodemography of Aging, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark; 5. Research School of Biology, Australian National University, Canberra, Australian Capital Territory 0200, Australia; 6. Department of Biology and Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado 80523; 7. Department of Biology, West Chester University, West Chester, Pennsylvania 19383; 8. Department of Biological Sciences, Barnard College, Columbia University, New York, New York 10027; 9. Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260; 10. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut 06269

Submitted May 28, 2013; Accepted October 18, 2013; Electronically published February 20, 2014

Online enhancement: appendix, supplementary PDF. Dryad data: <http://dx.doi.org/10.5061/dryad.4s286>.

ABSTRACT: Understanding the evolution of reaction norms remains a major challenge in ecology and evolution. Investigating evolutionary divergence in reaction norm shapes between populations and closely related species is one approach to providing insights. Here we use a meta-analytic approach to compare divergence in reaction norms of closely related species or populations of animals and plants across types of traits and environments. We quantified mean-standardized differences in overall trait means (*Offset*) and reaction norm shape (including both *Slope* and *Curvature*). These analyses revealed that differences in shape (*Slope* and *Curvature* together) were generally greater than differences in *Offset*. Additionally, differences in *Curvature* were generally greater than differences in *Slope*. The type of taxon contrast (species vs. population), trait, organism, and the type and novelty of environments all contributed to the best-fitting models, especially for *Offset*, *Curvature*, and the total differences (*Total*) between reaction norms. Congeneric species had greater differences in reaction norms than populations, and novel environmental conditions increased the differences in reaction norms between populations or species. These results show that evolutionary divergence of curvature is common and should be considered an important aspect of plasticity, together with slope. Biological details about traits and environments, including cryptic variation expressed in novel environmental conditions, may be critical to understanding how reaction norms evolve in novel and rapidly changing environments.

Keywords: reaction norm, phenotypic plasticity, meta-analysis, evolution, environmental gradient.

Introduction

Phenotypic plasticity refers to the ability of a single genotype to produce different phenotypes in different en-

vironments (Schlichting 1986; Scheiner 1993; Via et al. 1995; Pigliucci 2001). This variation in phenotype in response to the environment is graphically represented as a reaction norm, a plot of phenotype versus environmental values (Bradshaw 1965; Sultan 1987; Schlichting and Pigliucci 1998). Plasticity is ubiquitous in nature, and the magnitudes and patterns of plasticity can vary considerably among different types of organisms, traits, and environments. Reaction norms also reveal whether genotypes differ in trait means with or without changes in plasticity. A long-standing challenge in evolutionary biology is to understand how reaction norms evolve and how similarities and differences in reaction norms across closely related species reflect underlying evolutionary forces.

Plastic responses to the environment are evidenced in reaction norms through changes in slopes and curvature and can be adaptive or nonadaptive (Ghalambor et al. 2007). Research into the genetic architecture of plasticity (Via and Lande 1985; Kirkpatrick and Heckman 1989; Gomulkiewicz and Kirkpatrick 1992; Maughan et al. 2007) and studies of the effects of relaxed selection on plastic responses (the removal or reduction of a selection pressure important for maintaining plasticity; e.g., Snell-Rood et al. 2010) suggest that cross-environment genetic correlations and drift may limit the evolution and maintenance of adaptive reaction norms. Variation in responses among taxa indicate that attributes of reaction norms have evolved, and such variation represents raw material for further evolution of traits and their responses across environments (Pfennig et al. 2010). Here we document patterns of evolutionary differences in reaction norms among closely related taxa to address the question of which aspects of reaction norms are most evolutionarily variable. These

* Corresponding author; e-mail: murrenc@cofc.edu.

comparisons enable us to begin to examine the results of processes of evolution to shape the reaction norm, including a history of environmental heterogeneity, response to novel environments, chance events, and thus potential to diverge. We compare the differences between ecotypes within species to the differences between congeneric species. More specifically, we ask, When reaction norms evolve, are the differences among taxa primarily in their trait means, slopes of response, or curvature of response to an environmental gradient? Thereby, we are asking whether closely related species differ in trait responses to different environments and, if so, whether these responses are highly variable among species and environments.

A critical aspect of whether plasticity evolves and is maintained is the extent to which environmental variation is heterogeneous, including the extent to which aspects of the environment are rare or novel. One body of theory argues that genotypes exposed to novel environments may express cryptic genetic variation and, in some cases, hidden reaction norms (Schlichting and Pigliucci 1998; Dudash et al. 2005; Schlichting 2008). "Hidden" refers to that portion of the reaction norm expressed outside the normal range of internal or external environmental conditions experienced by the population (Schlichting and Murren 2004; Schlichting 2008). Novel environments may uncover variation and ultimately generate new patterns of selection on reaction norms, influencing both the slope and the curvature of the reaction norm. The evolutionary responses to such selection will depend on the availability of genetic variation for reaction norms and the subsequent effects of genetic constraints on further potentially adaptive responses. In addition to existing novel or rare aspects of environment, anthropogenic environmental change is happening at a fast pace, thus increasing the likelihood that contemporary organisms may experience novel abiotic or biotic environmental conditions.

Our ability to understand how plasticity evolves may depend on the range of environments and types of traits that are examined. Most empirical and theoretical studies of reaction norms consider trait values across two different environments, and the reaction norm is characterized by a straight-line segment and most commonly investigated as a slope (e.g., Kawecki and Ebert 2004). The inferences that can be drawn regarding the evolution of plasticity from such two-point reaction norms are limited because changes of the slope of the reaction norm cannot be separated from changes of the mean trait values in each environment. By addition of a third environment, the slope and mean differences can be independently estimated biologically and statistically (Via 1993; Schlichting and Pigliucci 1998). Continuous reaction norms across more than two environments along a gradient frequently reveal more complex shapes that vary in both slope and curvature (e.g.,

Scheiner and Lyman 1989; Gavrillets and Scheiner 1993*a*, 1993*b*; Pigliucci 2001; Beldade et al. 2011). For example, thermal reaction norms are typically convex in shape, with a central maximum and lower trait values at both lower and higher temperatures (Knies et al. 2006). In contrast, reaction norms for predator cues and other species interactions are generally monotonic in shape, or logistic, reflecting a threshold response (e.g., Relyea and Auld 2005). For morphological traits, reaction norms along other environmental axes display a wide variety of shapes (Schlichting and Pigliucci 1998; Pigliucci 2001). Given this diversity, it seems likely that evolutionary changes in curvature of reaction norms may differ among physiological rates, body size, behavior, fitness, and other types of traits. Similarly, genetic constraints on changes in reaction norm shape may depend on the type of environmental factor, such as temperature (Izem and Kingsolver 2005), environmental quality, and chemical cues associated with predators (e.g., Relyea and Auld 2005; Torres-Dowdall et al. 2012).

Comparisons among closely related species, or between populations that are locally differentiated (e.g., ecotypes or locally adapted populations), provide a powerful context for examining how evolution may shape divergence in patterns of reaction norms. Differences in reaction norms between pairs of closely related species or ecotypes can be characterized by (1) variation in trait means across environments (offset) and (2) variation in shape, including changes defined to include changes in slope and curvature. We investigate the influence of time since divergence by contrasting patterns of reaction norm divergence between congeneric and geographic populations. To the degree that variation in plasticity reflects underlying genetic variation across taxa or populations within taxa, such comparisons can also inform us about which aspects of the three main attributes of reaction norm (offset, slope, and curvature) undergo evolutionary change. Studies on continuous reaction norms permit the detection of nonlinearity in response to environmental gradients. By examining the patterns of reaction norm divergence, we can begin to investigate how evolutionary forces have shaped differences in continuous reaction norms through time.

Here we take advantage of a body of empirical literature documenting continuous reaction norms for related species and populations within species. We examine these data to answer the following four questions about the evolution of reaction norms: (1) How do differences in offset (environmentally independent differences in trait means between taxa), slope (differences in plasticity that are directional along an environmental gradient), and curvature (differences in trait responses that are not scaled directionally with the environment) of reaction norms contribute to the overall differences in reaction norms among

species and populations? (2) Do any of these components appear to be constrained (e.g., no difference between closely related species pairs or population pairs in this particular attribute of reaction norm variation; consistent with mechanisms such as lack of available genetic variation)? (3) Do patterns of reaction norm differences depend on the type of organism, trait, or environment considered? (4) Is reaction norm divergence greater among species than among populations within species?

Methods

The Literature Database

To assemble a database of phenotypic responses across environmental gradients, keyword searches (table A1; tables A1, A2 available online) were conducted in Web of Science (ISI) and Google Scholar to identify articles that reported reaction norms across three or more ordered environments for two or more congeners or populations within species. We specifically chose to examine studies for which the authors of the original work considered the populations independent evolutionary units with hypotheses for the distinct ecologies or phenotypes (e.g., ecotypes that are defined as populations). We did not include studies where the aim was to uncover genetic variation by sampling many genotypes within a population or multiple populations but which lacked a concomitant ecological hypothesis for divergence. These formal database searches were supplemented by searches through the reference lists of relevant articles, synthetic papers uncovered during our searches, papers already known to us, or our own published or unpublished data, with the aim to balance trait and environment types across animal and plant studies. We also systematically examined the past 10 years of individual journals (*Ecology*, *Evolution*, *The American Naturalist*) for publication of reaction norm studies; these detailed investigations resulted in a few additional studies that fit our criteria.

Criteria and Overview of Studies Examined

Studies were retained if phenotypic data were available in tabular or graphical format for two or more congeneric species or populations of individual species for each of three or more environments. Studies were excluded if authors did not identify how three or more discrete environments were ordered but were retained if authors identified an ordering or a gradient such that slope between environments was ecologically interpretable. We employed ImageJ (Rasband 1997–2011; Abramoff et al. 2004) to determine the trait values from graphics. We excluded studies that examined a random sample of genotypes or genetic

lines within a single population, unless the study proposed specific hypotheses about differences in reaction norms among the genotypes. If multiple traits were examined in the same study, all available phenotypic data were collected. We recorded study species, species type (vertebrate, invertebrate, plant, or microbe), number of environments investigated, location of the study (field vs. controlled environments), whether one of the environmental attributes was outside of the native range (as defined by the original study authors), phenotypic trait type, and environment type (also as defined by the authors). Phenotypic trait types and environment types were each further classified into a small number of broad categories (see below; table A2). Two types of categories received the designation “subpopulation”: (1) multiple populations within a geographic region and (2) pairs of populations grown at two levels of one environment and three or more levels of another environment. Both of these categories increase ecological variance. The assembled data set included 79 studies from 48 journals from the period 1971–2011, with more than 1,220 data records. Data are deposited in the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.4s286> (Murren et al. 2014). Our intent was not to be exhaustive but rather to develop a robust sample that captured variation in trait, organism, and environment types. An individual data record included all the phenotypes for a particular species (or ecotype) for a single trait across all environments along the same gradient where the phenotype of the species was measured.

Metrics

We aimed to quantify the differences between the mean reaction norms for pairs of species, populations, or subpopulations evaluated in the same study and to partition these differences into several distinct components that characterize the mean position (i.e., offset), and here we define shape to include both slope and curvature of the reaction norms. Suppose a trait is measured for an ordered set of environmental levels $i = 1 \dots n$. Let x_i and y_i be the mean values of the trait for two related species, x and y , measured in the environment, and X be the mean trait across species x across all environmental levels and Y be the trait mean for species y across all environmental levels. Then the set of values $\{x_1, x_2 \dots x_n\}$ represents the reaction norm for species x , and $\{y_1, y_2 \dots y_n\}$ represents the reaction norm for species y (fig. 1A). For the analyses, we excluded species that had mean trait values less than 0 in any environmental level (see below).

We can partition the differences between two reaction norms into four components: (1) *Offset*, O , which reflects the overall difference in mean trait value across all environments; (2) *Slope*, S , which characterizes the difference

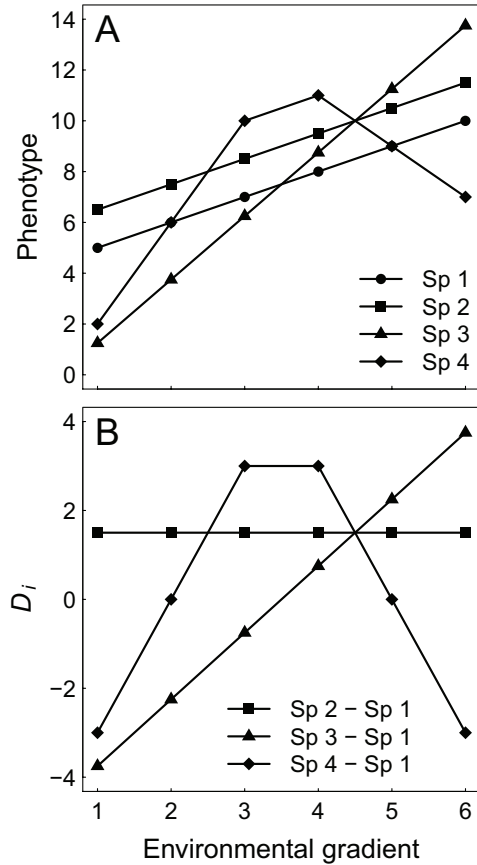


Figure 1: Quantifying differences in reaction norms. *A*, Species (or populations) may diverge from one another in reaction norms by changes in overall mean and shape. Consider the reaction norms measured at six environmental levels i for four species in the same genus: species 1 (Sp 1, black circles and solid line), species 2 (Sp 2, black squares and solid line), species 3 (Sp 3, black triangles and solid line), and species 4 (Sp 4, black diamonds and solid line). In this example, the total difference T in reaction norms between the focal and each relative's reaction norm are identical (see eqq. [1]–[4] and text). Relative to species 1, species 2 differs only in *Offset*, species 3 only in *Slope*, and species 4 only in *Curvature*. *B*, As an example, we compare the difference at each environmental level (D_i) between species 1 and each of the other three species depicted in *A* (additional graphs could be created for comparisons of species 2–4). Note the simple form of these differences (D_i) for each species. For species 2, the flat line at a positive value (zero slope and curvature) indicates a difference in *Offset*: $O = 1.5$, $S = 0$, $C = 0$ (eq. [1]). For species 3, the positive slope (zero offset and curvature) indicates a difference in *Slope*: $O = 0$, $S = 1.5$, $C = 0$ (eq. [2]). For species 4, the negative curvature (zero offset and slope) indicates a difference in *Curvature*: $O = 0$, $S = 0$, $C = 1.5$ (eq. [3]). Differences between empirical reaction norms will typically involve changes in all three of these attributes in varying proportions (see text for further description).

in overall slope; (3) *Curvature*, C , which depicts the average difference in curvature of the reaction norm, and (4) *Wiggle*, W , which describes the variability in shape that was not captured by any of the previous three measures. Using the notation from the previous paragraph and denoting D_i as the difference between two taxa for a trait measured in environment i and n as the number of environments (see also fig. 1*B*), these four components can be expressed algebraically as

$$O = \text{Offset} = \left| \frac{\sum_1^n D_i}{n} \right|, \quad (1)$$

$$S = \text{Slope} = \left| \frac{\sum_1^{n-1} S_i}{n-1} \right|, \text{ where } S_i = D_{i+1} - D_i, \quad (2)$$

$$C = \text{Curvature} = \left| \frac{\sum_1^{n-2} C_i}{n-2} \right|, \text{ where } C_i = S_{i+1} - S_i, \quad (3)$$

$$W = \text{Wiggle} = \frac{\sum_1^{n-2} |C_i|}{n-2} - C. \quad (4)$$

Here O , S , and C represent the average zeroth-, first-, and second-order differences between reaction norms per environmental level (see Discussion), allowing us to partition the total differences between reaction norms. We computed values of O , S , C , and W for each pair of congeneric species, or of differentiated populations within a species, and of subpopulations with a population for each unique combination of study and trait in the data set. Figure 1*A* depicts sample reaction norms that differ between species by only *Offset* (species 1 vs. 2), *Slope* (species 1 vs. 3), or *Curvature* (species 1 vs. 4). Variability of higher order than W was considered to be negligible. Figure 1*B* depicts the resulting D_i 's comparisons of offset only, slope only, and curvature only between the reaction norms in figure 1*A*. To allow comparisons across traits and environments at various measurement scales, we use the overall trait mean (M) for the species (or population) pair, $M = (X + Y)/2$, where X is the trait mean for species x across all environmental levels and Y is the trait mean for species y across all environmental levels. Then we define $\Delta O = O/M$, $\Delta S = S/M$, $\Delta C = C/M$, and $\Delta W = W/M$ for each pair of species or ecotype, representing the mean-standardized differences in *Offset*, *Slope*, *Curvature*, and *Wiggle*, respectively. Thus ΔO , ΔS , ΔC , and ΔW quantify the divergences in offset, slope, curvature, and higher-order terms in reaction norms between two species, relative to overall mean trait value for the two species. Finally, we define *Total* $\Delta T = T/M$, representing the total differences between the reaction norms relative to the overall trait mean, where $T = O + S + C + W$.

These values represent the contrasts derived from comparison of reaction norms for each pair of closely related taxa in a study, and thus it is the number of contrast lines (described below; table 1) rather than number of studies (described in the Literature Database section) that is critical to our analyses. Note that when reaction norms with only three environments are compared (i.e., $n = 3$), ΔO , ΔS , and ΔC account for all of the differences between reaction norms. When there are more than three environments, higher-order differences (i.e., ΔW) may also contribute. Our analyses focus on the mean-standardized values of *Offset*, *Slope*, *Curvature*, and *Total* (i.e., ΔO , ΔS , ΔC , and ΔT).

Some studies included multiple populations and multiple species, which were included in analyses. Cross-level contrasts (e.g., species \times population contrasts) were excluded from analysis owing to difficulty with their biological interpretation.

Influence of Moderators on Differences in Offset, Slope, Curvature, and Total

Single-Response Linear Mixed-Effects Models. Because values of ΔO , ΔS , and ΔC (eqq. [1]–[3]) are ratios, we natural log-transformed all values prior to statistical analysis to meet the normality assumptions of the models. We then performed separate linear mixed-effects models (nlme package in R, ver. 2.15.0; Pinheiro et al. 2012; R Development Core Team 2012) for each of the three (log-transformed) reaction norm attribute responses: ΔO , ΔS , and ΔC . Residuals were weighted by sample size (based on the mean of square-root transformed sample sizes for each of the two contrast components). Study identity was included as a random effect (random intercepts) to account for potential study-level autocorrelation when reaction norms for multiple traits were analyzed in the same study.

To examine which of the biological effects that we cataloged in our database contributed to the pairwise differences in reaction norms, we evaluated models with different combinations of moderators (i.e., fixed effects). These moderators in a meta-analysis are analogous to explanatory variables in a basic linear modeling framework, where an explanatory variable affects the response. However, moderators are distinguished by the fact that they influence the relationship between an explanatory variable and response rather than the response directly (Nakagawa and Santos 2012). We considered the following six moderators (levels listed alphabetically): (1) taxon contrast type (fixed factor; three levels: species-level contrast, population-level contrast, subpopulation-level contrast), (2) trait type (fixed factor; six levels: anatomy/morphology, behavior, development, physiology, reproductive output, size; see table A2 for complete list of categories), (3) organism type

Table 1: Number of contrast lines (pairwise comparisons of reaction norms of species or populations within studies) for each level of the six moderator variables considered in the meta-analyses of reaction norm attributes

Moderator variable	No. contrast lines
Taxon contrast type:	
Species	934
Population	1,980
Subpopulation	165
Organism type:	
Plant	1,914
Invertebrate	889
Vertebrate	276
Trait type:	
Anatomy/morphology	236
Behavior	302
Development	893
Physiology	303
Reproductive output	284
Size	1,061
Environment type:	
Elevation	61
Environmental quality	572
Latitude	911
Species interaction	123
Temperature	1,412
No. environments:	
3	2,146
4–6	246
7	346
8–12	341
Environmental range:	
Within	918
Outside	2,146

Notes: For summary purposes, the number of environments is treated here as a factor, but it is treated as a continuous covariate in the meta-analytical models. A total of 3,079 contrast lines were examined, except for environmental range, which comprises 3,064 contrast lines (reduced by those cases where environmental novelty could not be determined).

(fixed factor; three levels: plant, invertebrate, vertebrate), (4) environment type (fixed factor; five levels: elevation, environmental quality, latitude, species interaction, temperature), (5) number of environments (continuous covariate), and (6) environmental range (fixed factor; yes/no binary variable describing whether any of the environmental treatments extended outside an organism's typical environmental range). Additional moderators (e.g., whether the study was performed in the wild or in a controlled laboratory environment) and moderator levels (e.g., single-celled organisms) were initially considered but ultimately excluded from analysis owing to an insufficient number of studies in our database or collinearity with other moderators. As a diagnostic of the degree of collin-

Table 2: Summary of the model selection results for moderators of *Offset* (ΔO)

Moderator	Level	Coefficient	SE _{adj}	<i>z</i>	<i>P</i>
Intercept	...	−3.56347	.53706	6.635	<.0001
Taxon contrast type	Species (sp)	.39694	.07477	5.309	<.0001
	Subpopulation (sub)	.29109	.17141	1.698	.0895
Trait type	Behavior (beh)	.47073	.57131	.824	.4010
	Development (dev)	.36162	.10038	3.603	.0003
	Physiology (phys)	−.05146	.18042	.285	.7755
	Reproduction (rep)	.54571	.10526	5.185	<.0001
	Size	.06293	.09426	.668	.5044
Organism type	Plant	.99919	.31211	3.201	.0014
	Vertebrate (vert)	.32041	.27421	1.168	.2426
Environment type	Environmental quality (qual)	.38521	.46795	.823	.4104
	Latitude (lat)	.19126	.51466	.372	.7102
	Species interaction (inter)	1.38905	.53236	2.609	.0091
	Temperature (temp)	.84293	.46277	1.821	.0685
Environment range	Outside native range (out)	.48398	.22455	2.155	.03114
Environment number	...	−.03045	.03949	.771	.440730

Moderator name	Taxon contrast type (Contrast)	Trait type (Trait)	Organism type (Organism)	Environment type (Env.)	Environment range	Environment number
Relative importance	1	1	.97	1	.80	.32

Note: Models include all those with $\Delta AICc$ (corrected Akaike Information Criterion) values <7 , comprising the set of best-fitting models. Relative importance values for a given moderator are the sum of the model weights for each model in which that particular moderator occurs. Relative importance values range from 0 to 1, with values closer to 1 being more important moderators of *Offset* (ΔO). Moderator notation used in figure 2 is given in parentheses. Coefficients are presented as treatment contrasts, with the baseline level set as the alphabetically first level as entered in the data set.

earity present among the moderators, we computed the generalized variance-inflation factors for linear models of the reaction norm components as functions of the six moderators (models were weighted by sample size) and adjusted for degrees of freedom of the moderators (Fox and Monette 1992). All generalized variance-inflation factors (adjusted for degrees of freedom) for the six moderators we examined were less than five, which is the recommended cutoff for collinearity (Fox and Monette 1992; see table S3; tables S1–S3 available online).

We examined two sets of models. We first examined full models containing all six moderators (the biological effects described above). Moderators in meta-analyses are variables that may differ across studies and may systematically influence our effects (in our case, the moderators are the six described above, and the effects are reaction norm divergence: ΔO , ΔS , and ΔC). Second, we performed model selection based on corrected Akaike Information Criterion ($AICc$) to identify the most important moderators of reaction norm divergence (MuMIn package in R; Barton 2011) by selecting all models with $\Delta AICc$ values of less than 7 (Burnham et al. 2011). Akaike weights and model averaging guided identification of the most influential combination of moderators. We used relative importance (RI) values (the sum of Akaike weights for each model that contains the moderator of interest) to identify the most influential individual moderators for each of the

reaction norm components: ΔO , ΔS , and ΔC (Burnham and Anderson 2002). RI values range from 0, where the moderator appears in none of the best models, to 1, where the moderator is in all of the best models. We used RI to rank moderators in importance and considered RI values over 0.5 to be of biological interest. We evaluated how many of the best models included a particular moderator and the relative importance values of individual moderators to infer which of the biological factors were most important in the offset, slope, and curvature differences observed between populations, subpopulations within populations, or taxa (Belmaker and Jetz 2012).

Although we present results for ΔO , ΔS , and ΔC separately, we emphasize that they should be interpreted jointly owing to potential correlations among these metrics. Additionally, we present *Total* (ΔT ; table 5). We describe our model results based on the coefficients for each level of each moderator and the RI values of individual moderators (sum of the model weights for each model in which that moderator occurs) and use these to guide our inference of which of the biological factors were most important in differences in the offset, slope, and curvature observed between populations or species (tables 2–5). Note that model coefficients (tables 2–5) represent treatment contrasts relative to the first level (alphabetically) for each categorical moderator, which is included in the intercept of the model. For example, the intercept term incorporates

the following moderator levels: Taxon contrast type, population; Trait type, anatomical/morphological; Organism type, invertebrate; Environment type, elevation; Environmental range, inside the native range.

Markov Chain Monte Carlo Multiresponse Generalized Linear Mixed-Effects Model. Because our metrics of reaction norm position (*Offset*) and shape (*Slope*, *Curvature*, and *Wiggle*) were not necessarily independent, we examined the influence of the correlation structure among log-transformed mean-standardized *Offset*, *Slope*, *Curvature*, and *Wiggle* (ΔO , ΔS , ΔC , ΔW) with a Bayesian Markov chain Monte Carlo (MCMC) multiresponse generalized linear mixed-effects model (MCMCglmm package in R; Hadfield 2010). One benefit of the Bayesian approach is that ΔO , ΔS , and ΔC can be entered simultaneously as responses. Similar to the model selection approach described above for the single-response linear mixed-effects model, we evaluated MCMCglmm with each possible combination of the moderator variables in addition to an intercept-only model (table S1). The advantage of these models is that they account for study-level autocorrelation (i.e., individual-level reaction norm components were grouped according to the study from which they originated, with each study receiving a different [random] intercept; Pinheiro and Bates 2009; Hadfield 2010); however, these models do not account for sample size, as residuals currently cannot be weighted in the MCMCglmm function.

Results

Table 1 summarizes the data set used in the meta-analyses and the number of contrasts (reaction norm comparisons by trait, between pairs of species or populations, and within studies as described above) for each of the levels of the six moderator variables. In total, our analyses involved 3,079 contrasts. Of these, 69% involved only three environmental levels and 62% involved plants. Of the invertebrate records, 60% involved insects and more than 99% involved arthropods. Contrasts were more frequent between populations (64%) than between species (30%) or subpopulations (6%). Of the contrasts, 69% considered environmental levels outside the normal environmental range of the population or species.

Our metrics of reaction norm differences quantify the divergence in *Offset* (ΔO ; differences in means), *Slope* (ΔS), and *Curvature* (ΔC), relative to the overall trait mean for the two species or populations (see Methods). The weighted mean-standardized values of our metrics *Offset*, *Slope*, and *Curvature* were $\Delta O = 0.139$ (95% confidence intervals, 0.111–0.173); $\Delta S = 0.059$ (0.046–0.076); and $\Delta C = 0.103$ (0.0785–0.135). Total divergence ΔT was 0.43 (0.35–0.54). Recall that by definition, higher-order differ-

ences (i.e., *Wiggle*, ΔW) are 0 when there are only three environmental levels (69% of the contrasts; table 1). For cases with more than three environmental levels, the weighted mean-standardized value of *Wiggle* was $\Delta W = 0.081$ (0.049–0.129). Collectively, differences in reaction norm shape (which we define to include *Slope*, *Curvature*, and *Wiggle* together) were greater than differences in *Offset*. However, our meta-analyses show that the relative values ΔO , ΔS , and ΔC varied substantially with the kinds of traits, environments, and organisms under consideration (tables 2–4).

Model selection results from the linear mixed-effects models for ΔO (i.e., differences in trait means) are presented in table 2, displaying the set of coefficients for each moderator and level (as well as SE_{adj}) and the RI values for each moderator contributing to these models. Taxon contrast type, environment type, trait type, organism type, and environmental range all had high RI values for ΔO (RI > 0.80). Among taxon types, contrasts between species had higher ΔO than contrasts between populations (fig. 2A; where figures present coefficients from single-moderator models). Among trait types, ΔO was relatively greater for behavioral, developmental, and reproductive traits and relatively low for morphological and physiological traits (fig. 2A). Among organism types, plants had higher ΔO than invertebrates or vertebrates. Among environmental types, ΔO was greater for contrasts involving interspecific interactions, intermediate for environmental quality and temperature, and lower for elevation and latitude (table 2; fig. 2A). It was also greater for environmental conditions outside the natural range of environmental conditions, but the magnitude of this effect is modest (fig. 2A).

Model selection results from the linear mixed-effects models for *Slope* divergence (ΔS) revealed that taxon contrast type, number of environments, and trait type all had relative importance values of 1 (table 3). Meanwhile, ΔS decreased with the number of environmental levels measured. Contrasts between species had higher ΔS values than those between populations (fig. 2B). The organism type, environment type, and environmental range all had RI < 0.4. The low importance of many moderators for ΔS is reflected in the relatively small differences for the factor levels of the moderators (fig. 2B; note the different scales for fig. 2A–2C).

Model selection results from the linear mixed-effects models for *Curvature* divergence (ΔC) showed that environmental type, number of environments, organism type, and trait type all had RI values of 1 (table 4). Taxon contrast type also had high relative importance (RI = 0.76). While ΔC was greatest for reproductive and physiological traits, intermediate for behavior and size traits, and lowest for developmental and morphology (fig. 2C),

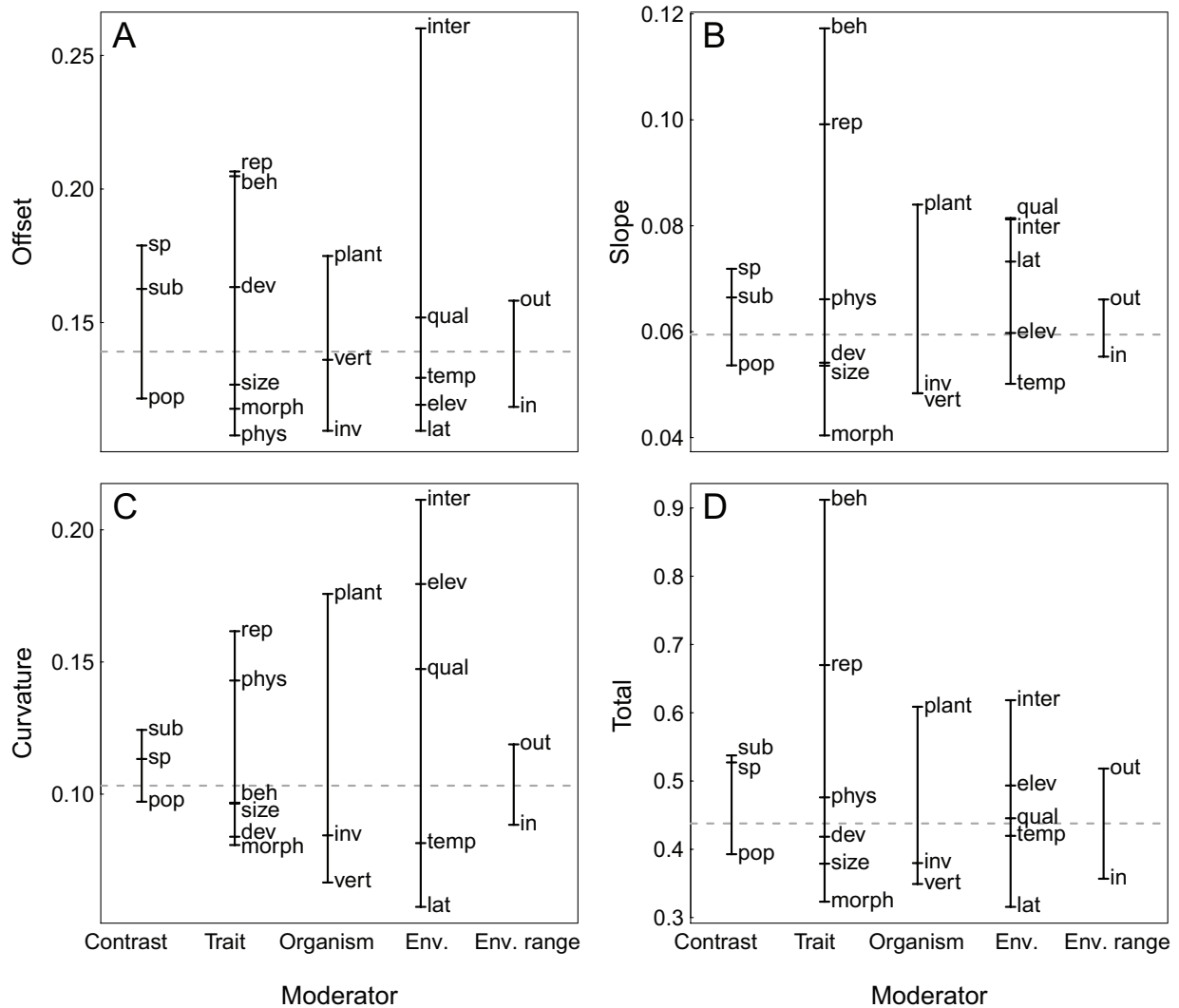


Figure 2: Mean ΔO (A), ΔS (B), ΔC (C), and ΔT (D; dashed gray lines show *Offset*, *Slope*, *Curvature*, and *Total* values, respectively) representing back-transformed (from natural log) coefficients from single-moderator linear mixed-effects models weighted by sample size and accounting for study-level autocorrelation for all levels of the five categorical moderators (solid vertical lines connect factor levels of a given moderator). The coefficient (95% confidence interval) for the continuous moderator number of environments (not shown in figure) is *Offset*: 0.968 (0.896, 1.05); *Slope*: 0.811 (0.752, 0.875), *Curvature*: 0.829 (0.766, 0.897); and *Total*: 0.938 (0.869, 1.01) from a single-moderator linear mixed-effect model weighted by sample size and accounting for study-level autocorrelation. Abbreviations for factor levels are as follows. Contrast (taxon contrast type): pop (population), sub (subpopulation), sp (species). Trait (trait type): beh (behavior), dev (development), morph (morphology), phys (physiology), rep (reproduction), size (body size). Organism (organism type): inv (invertebrate), vert (vertebrate). Env. (environment type): elev (elevation), inter (species interaction), lat (latitude), qual (environmental quality), temp (temperature). Env. range (within or outside normal environmental range): in (within), out (outside). Note the different scales for the Y-axes of A–D. *Total* values are not equivalent to the sum of A–C owing to *Wiggle* effects (not shown) and natural log-scaling effects.

it decreased with the number of environmental levels measured for reaction norms. Animals had lower ΔC than plants. Among environmental types, species interactions, elevation, and environment quality had greater ΔC than other environmental gradients. Meanwhile, ΔC was higher when reaction norms included novel environmental levels

outside the typical environmental range, but the effect was modest (fig. 2C).

We also modeled the effects of these moderators on the total differences in reaction norms (ΔT ; fig. 2D). Model selection results from the linear mixed-effects models for ΔT showed that all factors had $RI > 0.6$, with taxon type,

Table 3: Summary of the model selection results for moderators of *Slope* (ΔS)

Moderator	Level	Coefficient	SE _{adj}	z	P
Intercept	...	-2.32475	.25775	9.019	<.0001
Taxon contrast type	Species (sp)	.32185	.06971	4.617	<.0001
	Subpopulation (sub)	.18506	.16054	1.153	.2490
Trait type	Behavior (beh)	1.31032	.59107	2.217	.0266
	Development (dev)	.25604	.09340	2.745	.0061
	Physiology (phys)	.54047	.16974	3.184	.0015
	Reproduction (rep)	.78951	.09790	8.064	<.0001
	Size	.24097	.08773	2.747	.0060
Organism type	Plant	.37709	.26270	1.435	.1512
	Vertebrate (vert)	-.04909	.27497	.179	.8583
Environment type	...	0 ^a
Environment range	Outside native range (out)	.12612	.22252	.567	.5709
Environment number	...	-.20318	.03837	5.295	<.0001

Moderator name	Taxon contrast type (Contrast)	Trait type (Trait)	Organism type (Organism)	Environment type (Env.)	Environment range	Environment number
Relative importance	1	1	.39	0	.30	1

Note: Models are all those with $\Delta AICc$ (corrected Akaike Information Criterion) values <7 , comprising the set of best-fitting models. Relative importance values for a given moderator are the sum of the model weights for each model in which that particular moderator occurs. Relative importance values range from 0 to 1, with values closer to 1 being more important moderators of *Slope* (ΔS). Moderator notation used in figure 2 is given in parentheses. Coefficients are presented as treatment contrasts, with the baseline level set as the alphabetically first level as entered in data set.

^a Because environmental type does not appear in any of the best-fitting models, coefficients for the levels of moderator are not estimated.

environment type, organism type, and trait type having RI values of 1 (table 5). For contrasts between species, ΔT was greater than between populations, suggesting greater differences between reaction norms at the species level than at the population level. Behavioral and reproductive traits had greater ΔT compared to morphological and size traits, and plants had greater ΔT than animals. Species interactions had greater ΔT than other types of environmental factors. Importantly, there were greater total differences in reaction norms for environmental conditions extending outside the natural range, suggesting that novel environments generally increase the differences in reaction norms between populations and congeneric species.

We also performed Bayesian meta-analyses (see Methods for description of MCMCglmm). The benefit of this approach is that we can analyze the response variables ΔO , ΔS , and ΔC simultaneously in the same model, thus explicitly considering the correlation structure among ΔO , ΔS , and ΔC values and accounting for study-level autocorrelation. A limitation of this approach, however, is that it does not allow weighting based on sample sizes. Log-transformed ΔO , ΔS , and ΔC values were all positively correlated ($\Delta O : \Delta S$; $r = 0.52$; $\Delta O : \Delta C$; $r = 0.38$; $\Delta S : \Delta C$; $r = 0.45$). Results from the alternative Bayesian model selection for these moderators are summarized in table S1, and the results are qualitatively similar to those from the linear mixed-effects models, that is, which moderators generally contributed to the best-fitting models. Trait type, number of environments, environment type, organism

type, and taxon contrast had $RI > 0.77$, strongly supporting the importance of these factors in explaining how divergences in offset and curvature contribute to the overall differences between reaction norms.

Discussion

Evolutionary Divergence in Reaction Norms

Numerous studies and reviews have documented evolutionary divergence in trait means between species or populations (e.g., Schluter 2000). Our study is the first quantitative, comprehensive analysis of this issue for both mean and shape changes (both slope and curvature) in reaction norms. A key result of our meta-analyses is that differences in reaction norm shape—which we define to include *Slope*, *Curvature*, and *Wiggle* (ΔS , ΔC , and ΔW)—represent the majority of the overall mean standardized differences in reaction norms (fig. 2). This suggests that for closely related species and populations, evolutionary changes in reaction norm shape are generally similar to or greater in magnitude than evolutionary changes in trait means. Of course, this result is limited to study systems and traits in which at least one population or species exhibited phenotypic plasticity, given the criteria used in assembling our data set. This result highlights the importance of micro-evolutionary changes in phenotypic plasticity in many plant and animal systems (Kirkpatrick and Heckman 1989). In addition, evolutionary differences in *Curvature*

Table 4: Summary of the model selection results for moderators of *Curvature* (ΔC)

Moderator	Level	Coefficient	SE _{adj}	z	P
Intercept	...	-2.02088	.54345	3.719	.0002
Taxon contrast type	Species (sp)	.17464	.07342	2.379	.0174
	Subpopulation (sub)	.21959	.16829	1.305	.1919
Trait type	Behavior (beh)	.56130	.59530	.943	.3457
	Development (dev)	.07575	.09842	.769	.4416
	Physiology (phys)	.71992	.17869	4.029	<.0001
	Reproduction (rep)	.62463	.10307	6.060	<.0001
	Size	.13523	.09232	1.465	.1430
Organism type	Plant	.80746	.32967	2.449	.0143
	Vertebrate (vert)	-.41682	.28658	1.454	.1458
Environment type	Environmental quality (qual)	-.23289	.48453	.481	.6308
	Latitude (lat)	-.49914	.52565	.950	.3423
	Species interaction (inter)	.77758	.53689	1.448	.1475
	Temperature (temp)	.01266	.47313	.027	.9786
Environment range	Outside native range (out)	.34043	.23756	1.433	.1519
Environment number	...	-.17598	.03971	4.432	<.0001

Moderator name	Taxon contrast type (Contrast)	Trait type (Trait)	Organism type (Organism)	Environment type (Env.)	Environment range	Environment number
Relative importance	.76	1	1	1	.51	1

Note: Models are all those with $\Delta AICc$ (corrected Akaike Information Criterion) values <7 , comprising the set of best-fitting models. Relative importance values for a given moderator are the sum of the model weights for each model in which that particular moderator occurs. Relative importance values range from 0 to 1, with values closer to 1 being more important moderators of *Curvature* (ΔC). Moderator notation used in figure 2 is given in parentheses. Coefficients are presented as treatment contrasts, with the baseline level set as the alphabetically first level as entered in data set.

(and, for studies with more than three environmental levels, in *Wiggle*) are generally greater than differences in *Slope* (fig. 2). This finding has two important implications. First, one might expect greater genetic constraints on second- and higher-order aspects of shape for continuous reaction norms (Kirkpatrick and Heckman 1989; Stinchcombe et al. 2012). However, our results suggest that microevolutionary changes in curvature and higher-order shape (*Wiggle*) may be more evolutionarily variable than simpler (first-order) aspects of reaction norm shape such as slope. Second, studies of genetic variation and evolutionary divergence in reaction norms rarely consider more than two environmental levels and quantify plasticity solely in terms of reaction norm slope (or its equivalent). Given the magnitude of *Curvature* and *Wiggle* relative to *Slope* found in our analyses, two-environment studies may underestimate the extent of evolutionary change in reaction norms and fail to capture biologically relevant patterns of plasticity.

We emphasize that divergences in both trait means (*Offset*) and reaction norm shape (*Slope*, *Curvature*, and *Wiggle* together) contribute substantially to the total evolutionary divergence between populations and congeneric species in reaction norms. Note that our analyses also show that evolutionary divergences in reaction norms are strongly moderated by aspects of the environments, traits, and organisms under study. Below we consider the implications

and contributions of these biological aspects for our understanding of reaction norm evolution.

Moderators of Reaction Norm Divergence

In general, we expect that the time since divergence between congeneric species would be longer than the divergence time between geographic populations within a species. As a result, we would predict that evolutionary divergences should be greater between species than between populations. These patterns would be consistent with expected rates of gene flow and reproductive isolation. Similarly, if genetic variation in reaction norm shape was limited or constrained, we would predict that differences in *Slope* and *Curvature* would be much greater between species than between populations. The assumption underlying these predictions is that shorter divergence times can influence these values in two ways: first, by limiting production of variation per se and, second, due to limitations in the availability of altered patterns of genetic covariance. Our results strongly support the first prediction: contrast type had high RI values for all metrics, and for each metric there was greater divergence between species than between populations. Support for the second prediction is less clear: although *Offset*, *Slope*, *Curvature*, and *Total* all had larger values for species than for pop-

Table 5: Summary of the model selection results for moderators of *Total* (ΔT)

Moderator	Level	Coefficient	SE _{adj}	z	P
Intercept	...	-1.94792	.55963	3.481	.0005
Taxon contrast type	Species (sp)	.32990	.05747	5.740	<.0001
	Subpopulation (sub)	.31751	.13202	2.405	.0162
Trait type	Behavior (beh)	1.09380	.59146	1.849	.06441
	Development (dev)	.30077	.07672	3.920	<.0001
	Physiology (phys)	.44775	.14382	3.113	.0019
	Reproduction (rep)	.73256	.07990	9.168	<.0001
	Size	.15694	.07178	2.186	.02879
Organism type	Plant	1.01982	.33751	3.022	.00251
	Vertebrate (vert)	-.0350	.29061	.120	.90415
Environment type	Environmental quality (qual)	-.05137	.50018	.103	.9182
	Latitude (lat)	-.05971	.53976	.111	.91191
	Species interaction (inter)	.81948	.54589	1.501	.1333
	Temperature (temp)	.55182	.47691	1.157	.2472
Environment range	Outside native range (out)	.41835	.24217	1.728	.0841
Environment number	...	-.06493	.03335	1.947	.0516

Moderator name	Taxon contrast type (Contrast)	Trait type (Trait)	Organism type (Organism)	Environment type (Env.)	Environment range	Environment number
Relative importance	1	1	1	1	.63	1

Note: Models are all those with $\Delta AICc$ (corrected Akaike Information Criterion) values <7 , comprising the set of best fitting models. Relative importance values for a given moderator are the sum of the model weights for each model in which that particular moderator occurs. Relative importance values range from 0 to 1, with values closer to 1 being more important moderators of *Total* (ΔT). Moderator notation used in figure 2 is given in parentheses. Coefficients are presented as treatment contrasts, with the baseline level set as the alphabetically first level as entered in data set.

ulation contrasts (fig. 2; note the different scales), statistical support for this pattern is lacking. Subpopulation contrasts also had relatively high estimated coefficients in the mean standardized models, but there is little statistical support for this effect due to the small number of subpopulation contrasts (5% of the total from 11 studies). For example, the model coefficients for subpopulation were not significantly different from population (intercept) for any of the metrics (tables 2–5). Thus, these results are consistent with the notion that evolutionary divergence time or cumulative selection contributes to divergence in reaction norms. We strongly recommend that further studies on the evolution of reaction norms (slope, curvature, and trait means) employ a phylogenetic framework with multiple species and populations within species grown across three or more environments. Such studies will help to further elucidate the importance of divergence time in contemporary differences in reaction norms.

The type of trait examined also influenced how reaction norms differ and contributed importantly to models for all of our metrics (tables 2–4, S1; fig. 2). For example, reproductive traits had relatively high divergence in all aspects of reaction norms, whereas morphological traits had relatively low divergence. This is consistent with the pattern that reproductive traits may experience stronger selection than morphological traits in many systems (Kingsolver et al. 2012), leading to greater evolutionary diver-

gence; it is not consistent with the idea that limited genetic variation in reproductive and life-history traits should limit their evolutionary responses (Mousseau and Roff 1987). Physiological traits showed a different pattern, with high divergence in curvature but low divergence in other aspects of reaction norm. Note that physiological reaction norms in response to temperature, light intensity, and other environmental factors often have negative curvature with a unimodal or plateauing shape, so evolutionary differences in curvature can indicate the evolution of environmental specialists (with steep curves) and generalists (with broader curves; e.g., Huey and Kingsolver 1989). It is noteworthy that behavioral traits exhibited the greatest divergence in *Slope* and in *Total*, suggesting that such traits—and their plasticity—are quite evolutionarily variable. This runs counter to the idea that behavioral plasticity can reduce evolutionary responses by reducing the strength of selection (Huey et al. 2003). We advocate examining whether these patterns hold within studies that consider physiological, reproductive, and behavioral traits together in the same populations or species. Additionally, further research on multivariate trait combinations across species and environments may illuminate broad differences in evolutionary variation in reaction norms or canalization of traits closely related to fitness or performance.

One motivation for our studies was to evaluate whether responses to novel environmental conditions (i.e., outside

of the normal environmental range) produced greater differences in reaction norms between populations and species than those comparisons inside the native range. Phenotypic plasticity has been implicated as a key contributor to persistence in the face of novel environmental conditions, especially in naturally weedy species, and may be of particular importance following human alterations of environments, such as via climate change, urbanization, or introduction of invasive species (e.g., Richards et al. 2006; Nicotra et al. 2010; but see Palacio-Lopez and Gianoli 2011). Environmental range had moderate to high RI values for *Offset*, *Curvature*, and *Total* divergence: in each case, divergence between taxa was greater in novel environmental conditions (fig. 2). This is consistent with the prediction that exposure to novel environmental conditions can reveal hidden aspects of reaction norm variation, possibly facilitating selection and subsequent evolution of the reaction norms (e.g., Schlichting 2008). A more refined and quantitative metric of environmental novelty, rather than the simple binary measure used here, would be valuable for further analyses of this hypothesis. More studies of species pairs along environmental gradients from inside to outside their current native ranges are needed for a better understanding of hidden reaction norms and how they may affect the potential for reaction norm evolution.

Environment type also contributed substantially to our models investigating evolutionary divergence in *Offset*, *Curvature*, and *Total* divergence. One illuminating pattern is that biotic interactions (i.e., interactions among species) produced relatively greater divergences in both trait mean and reaction norm shape (*Slope* and *Curvature*) than most abiotic aspects of environment (fig. 2). The relative importance of biotic and abiotic factors remains a recurring theme in ecology and evolution. It is especially relevant in contemporary ecology to understanding biological responses to invasive species, climate change, and other types of environmental change. Our results suggest that biotic interactions may be particularly important for generating microevolutionary changes in reaction norms. Further investigations of the clinal patterns in both trait means and other aspects of reaction norms (e.g., Bublly and Loeschcke 2005; Baird et al. 2011) and in the relative strength of selection on trait means and curvature will shed light on how multivariate environmental changes influence reaction norm evolution. One such approach to complex multivariate studies that would be particularly valuable for such investigations could include sets of populations that occur along both elevational and latitudinal gradients that vary in biotic interactions.

One unexpected finding to emerge from our analyses is that plants had greater evolutionary divergence than animals, with strong statistical support for this pattern for *Offset*, *Curvature*, and *Total* divergence. The importance

of phenotypic plasticity in plants has long been recognized (Bradshaw 1965), and Bradshaw (1972) proposed that plants (and other sessile organisms) should experience stronger phenotypic selection than (mobile) animals (but see Huey et al. 2002). Our analyses suggest that closely related taxa of plants have relatively greater evolutionary divergence in both trait means (*Offset*) and plasticity (*Curvature*). Differences in selection, gene flow, or genetic variation between plants and animals could generate this pattern.

Some Limitations

One limitation of our approach here, and more generally of any linear or additive analysis, is that it cannot adequately characterize changes or shifts in reaction norms along the horizontal (environmental) axis (consistent with range shifts or survivorship variation along an environmental gradient). For nonlinear reaction norms, simple horizontal shifts in position (e.g., the same response curve with a shift in the optimum environment) involve nonlinear transformations that can influence slope and curvature metrics (Izem and Kingsolver 2005). Other analytical methods are needed to further address this issue, such as curve registration or nonlinear regression. Additionally, this aspect of reaction norm shift may be important in species divergence and divergence of habitat ranges of closely related species.

Our metrics for reaction norms— ΔO , ΔS , ΔC , and ΔW —represent the average zeroth-, first-, second-, and higher-order differences between reaction norms per environmental level (see Methods) relative to the overall mean trait value of the two species or populations. This approach allows us to partition to total differences between reaction norms. However, for a given range of environments, ΔS , ΔC , and ΔW necessarily decline as the number of measured environmental levels (n) increases (eq. [2]–[4]; tables 2–4). An alternative approach is to consider contributions to reaction norm differences (D_i) at low and high environmental levels. For example, we can define the rise R as an alternative metric of total slope, where $R = S \times (n - 1) = |D_n - D_1|$ (eq. [2]). Similarly, we can define Q as an alternative metric of total curvature, where $Q = C \times (n - 2) = |(D_n + D_1) - (D_{n-1} + D_2)|$ (eq. [3]). The metrics $\Delta R = R/M$ and $\Delta Q = Q/M$ are not necessarily affected by n . Analyses of these metrics (table S3; fig. S1; fig. S1 available online) yield weighted mean-standardized values of $\Delta R = 0.154$ (95% confidence intervals, 0.116–0.203) and $\Delta Q = 0.173$ (0.135–0.221), which are larger than the values for ΔS and ΔC (see Results). However, the qualitative results for these metrics remain the same: in general differences in curvature are greater than differences in slope (i.e., $\Delta Q > \Delta R$), and

changes in reaction norm shape (i.e., $\Delta Q + \Delta R$) are greater than those for *Offset* (ΔO). In addition, the RI values of the moderators are very similar for ΔS and ΔR and for ΔC and ΔQ (tables 2, 3, S3). As expected, the one exception is for the number of environments: this has a low RI value (0.27) for ΔR but a large RI value (0.95) for ΔQ . Interestingly, increasing the number of measured environments increases ΔQ but decreases ΔC . Finally, the relative factor levels for each moderator are similar (though not identical) for ΔS and ΔR (figs. 2B, S1A) and for ΔC and ΔQ (figs. 2C, S1B). The use of other metrics and approaches to standardization may yield additional insights about divergence in reaction norms; we encourage other researchers to use our data set to explore alternative metrics.

As with all meta-analyses, the data available may be biased or censored (“the file drawer problem”). Studies with no apparent differences in any aspect of the reaction norm (trait means, slope, or curvature) may not be published. On the other hand, studies we have included were likely to be those where species or populations differed in their overall means. Our analyses, therefore, are limited to investigating the question of which aspect of the reaction norm are most often evolutionary labile, when at least one of the aspects of the reaction norm have been quantified to differ among species. Additional studies from a canalization perspective may require further knowledge of those species pairs that do not differ.

Investigators may not measure phenotypes in novel conditions, and phenotypic measurements cannot be made if experimental replicates die in particular conditions: if one species did not survive in an environment but another species did, such differences are unlikely to be captured in our data set. Differences in survivorship reaction norms are worthy of continued investigation, as they may indicate strong ecological divergence between species or ecotypes. As scholarship in statistical modeling approaches for meta-analyses continue to be developed, a reinvestigation of the patterns presented here would be warranted if models that are capable of evaluating covariances among responses, random effects, sample size weighting, and interaction of moderators in a single comprehensive analysis are developed.

Conclusion

Here we have uncovered that populations and species can diverge in trait means and reaction norms can diverge beyond simply slope to include divergence in curvature and wiggle. We recommend that future studies examining reaction norm evolution incorporate both a phylogenetic perspective and experimental designs with more than three levels to explore the importance of slope, curvature, and

higher-order components in reaction norm evolution. Our results show that patterns of reaction norm vary across types of organisms, traits, and environments. We advocate additional studies that include additional environment types not investigated here to extend our understanding of evolutionary change of reaction norms. In particular, we encourage the expansion of investigations of reaction norms in novel environments beyond the broad analysis presented here and their effects on phenotypic variation. These studies will contribute to our ability to predict future responses to environmental change. Future investigations will further refine our understanding of the divergence in reaction norms we have uncovered, and in particular we urge the examination of evolutionary mechanisms behind the patterns described here.

Acknowledgments

We acknowledge support from the U.S. National Evolutionary Synthesis Center (NESCent) and especially from members of the Costs of Plasticity working group, including J. Masel, H. Maughan, S. Seiter, and E. Snell-Rood. Additional data collection was provided by C. Acasuso and A. Demarez. We gratefully acknowledge funding from NESCent (National Science Foundation [NSF] grant EF-0905606) for the Costs of Plasticity working group (coled by C.J.M. and C.D.S.) and a short-term NESCent fellowship to C.J.M. Additionally, we acknowledge funding support from the NSF to C.J.M. (IOS-1146977), H.S.C. and C.J.M. (IOS-1052262), C.K.G. (DEB-0846175), R.E.R. (DEB-0518250), and J.K. (IOS-1120500), and from the NSF Partnerships for International Research and Education (OISE-0623341) and NSF Dimensions of Biodiversity (DEB-1046328) to C.D.S.

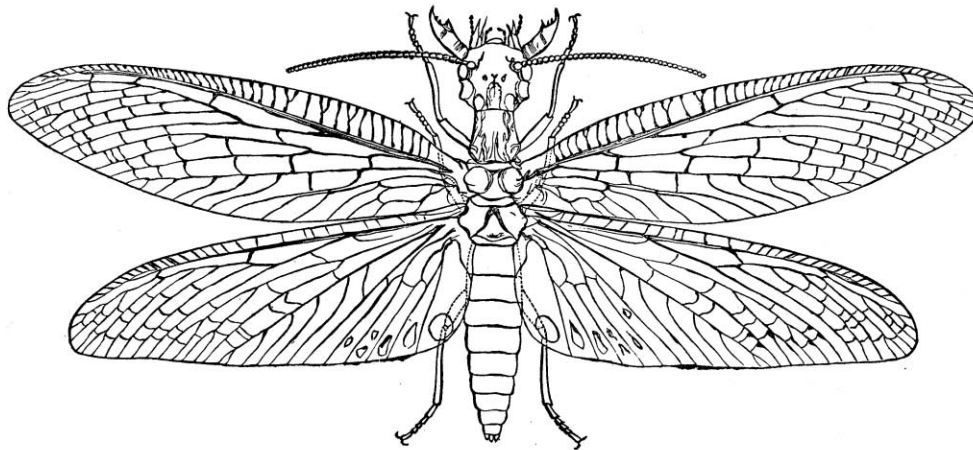
Literature Cited

- Abramoff, M. D., P. J. Magalhaes, and S. J. Ram. 2004. Image processing with ImageJ. *Biophotonics International* 11:36–42.
- Baird, T. D., C. A. Henson, N. M. Daily, F. V. Baccari, and C. J. Murren. 2011. Differential patterns of plasticity to water availability along native and naturalized latitudinal gradients. *Evolutionary Ecology Research* 13:55–73.
- Bartoń, K. 2011. MuMIn: multi-model inference. R package, version 1.2.4. <http://r-forge.r-project.org/projects/mumin/>.
- Beldade, P., A. R. A. Mateus, and R. A. Keller. 2011. Evolution and molecular mechanisms of adaptive developmental plasticity. *Molecular Ecology* 20:1347–1363.
- Belmaker, J. C., and W. Jetz. 2012. Regional pools and environmental controls of vertebrate assemblages. *American Naturalist* 179:512–523.
- Bradshaw, A. D. 1965. Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics* 13:115–155.

- . 1972. Some of the evolutionary consequences of being a plant. *Evolutionary Biology* 5:25–47.
- Bubli, O. A., and V. Loeschke. 2005. Variation of life-history and morphometrical traits in *Drosophila buzzatii* and *Drosophila simulans* collected along an altitudinal gradient from a Canary island. *Biological Journal of the Linnean Society* 84:119–136.
- Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, New York.
- Burnham, K. P., D. R. Anderson, and K. P. Huyvaert. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65:23–35.
- Dudash, M. R., C. J. Murren, and D. E. Carr. 2005. Using *Mimulus* as a model system to understand the role of inbreeding in conservation: genetic and ecological approaches. *Annals of the Missouri Botanical Garden* 92:36–51.
- Fox, J., and G. Monette. 1992. Generalized collinearity diagnostics. *Journal of the American Statistical Society* 3:178–183.
- Gavrilets, S., and S. M. Scheiner. 1993a. The genetics of phenotypic plasticity. 5. Evolution of reaction norm shape. *Journal of Evolutionary Biology* 6:31–48.
- . 1993b. The genetics of phenotypic plasticity. 6. Theoretical predictions for directional selection. *Journal of Evolutionary Biology* 6:49–68.
- Ghalambor, C. K., J. K. McKay, S. P. Carroll, and D. N. Reznick. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21:394–407.
- Gomulkiewicz, R., and M. Kirkpatrick. 1992. Quantitative genetics and the evolution of reaction norms. *Evolution* 46:390–411.
- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* 33:1–22. <http://www.jstatsoft.org/v33/i02/>.
- Huey, R. B., M. Carlson, L. Crozier, M. Frazier, H. Hamilton, C. Harley, A. Hoang, and J. G. Kingsolver. 2002. Plants versus animals: do they deal with stress in different ways? *Journal of Integrative and Comparative Biology* 42:232–240.
- Huey, R. B., P. E. Hertz, and B. Sinervo. 2003. Behavioral drive versus behavioral inertia in evolution: a null model approach. *American Naturalist* 161:357–366.
- Huey, R. B., and J. G. Kingsolver. 1989. Evolution of thermal sensitivity of ectotherm performance. *Trends in Ecology and Evolution* 4:131–135.
- Izem, R., and J. G. Kingsolver. 2005. Variation in continuous reaction norms: quantifying directions of biological interest. *American Naturalist* 166:277–289.
- Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7:1225–1241.
- Kingsolver, J. G., S. E. Diamond, A. M. Siepielski, and S. M. Carlson. 2012. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evolutionary Ecology* 26:1101–1118.
- Kirkpatrick, M., and N. Heckman. 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *Journal of Mathematical Biology* 27:429–450.
- Knies, J. L., R. Izem, K. L. Supler, J. G. Kingsolver, and C. L. Burch. 2006. The genetic basis of thermal reaction norm evolution in lab and natural phage populations. *PLoS Biology* 4:1257–1264.
- Maughan, H., J. Masel, C. W. Birky, and W. N. Nicholson. 2007. The roles of mutation accumulation and selection in loss of sporulation in experimental populations of *Bacillus subtilis*. *Genetics* 177:937–948.
- Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. *Heredity* 59:181–197.
- Murren, C. J., H. J. Maclean, S. E. Diamond, U. K. Steiner, M. A. Heskell, C. A. Handelsman, C. K. Ghalambor, et al. 2014. Data from: Evolutionary change in continuous reaction norms. *American Naturalist*, Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.4s286>.
- Nakagawa, S., and E. S. A. Santos. 2012. Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology* 26:1253–1274.
- Nicotra, A. B., O. K. Atkin, S. P. Bonser, A. M. Davidson, E. J. Finnegan, U. Mathesius, P. Poot, et al. 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* 15:684–692.
- Palacio-Lopez, K., and E. Gianoli. 2011. Invasive plants do not display greater phenotypic plasticity than their native or non-invasive counterparts: a meta-analysis. *Oikos* 120:1393–1401.
- Pfennig, D. W., M. A. Wund, E. C. Snell-Rood, T. Cruickshank, C. D. Schlichting, and A. P. Moczek. 2010. Phenotypic plasticity's impacts on diversification and speciation. *Trends in Ecology and Evolution* 25:459–467.
- Pigliucci, M. 2001. *Phenotypic plasticity: beyond nature and nurture*. Johns Hopkins University Press, Baltimore.
- Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and the R Development Core Team. 2012. nlme: linear and nonlinear mixed effects models. R package, version 3.1–103. <http://cran.r-project.org/web/packages/nlme/index.html>.
- Pineheiro, J. C., and D. M. Bates. 2009. *Mixed-effects models in S and S-PLUS*. Springer, New York.
- R Development Core Team. 2012. R: a language environment for statistical computing. Vienna, Austria. <http://www.R-project.org>.
- Rasband, W. S. 1997–2011. ImageJ. U.S. National Institutes of Health, Bethesda, MD. <http://imagej.nih.gov/ij/>.
- Relyea, R. A., and J. R. Auld. 2005. Predator- and competitor-induced plasticity: how changes in foraging morphology affect phenotypic trade-offs. *Ecology* 86:1723–1729.
- Richards, C. L., O. Bossdorf, N. Z. Muth, J. Gurevitch, and M. Pigliucci. 2006. Jack of all trades, master of some? on the role of phenotypic plasticity in plant invasions. *Ecology Letters* 9:981–993.
- Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics* 24:35–68.
- Scheiner, S. M., and R. F. Lyman. 1989. The genetics of phenotypic plasticity. 1. Heritability. *Journal of Evolutionary Biology* 2:95–107.
- Schlichting, C. D. 1986. The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics* 17:667–693.
- . 2008. Hidden reaction norms, cryptic genetic variation, and evolvability. *Annals of the New York Academy of Sciences* 1133:187–203.
- Schlichting, C. D., and C. J. Murren. 2004. Evolvability and the raw materials for adaptation. Pages 18–29 *in* Q. C. B. Cronk, J. Whitton, R. H. Ree, and I. E. P. Taylor, eds. *Plant adaptation: molecular genetics and ecology*. NRC Research, Ottawa.
- Schlichting, C. D., and M. Pigliucci. 1998. Phenotypic evolution: a reaction norm perspective. Sinauer, Sunderland, MA.
- Schluter, D. 2000. *The ecology of adaptive radiations*. Oxford University Press, New York.

- Snell-Rood, E. C., J. D. Van Dyken, T. Cruickshank, M. J. Wade, and A. P. Moczek. 2010. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *Bioessays* 32:71–81.
- Stinchcombe, J. R., Function-Valued Traits Working Group, and M. Kirkpatrick. 2012. Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. *Trends in Ecology and Evolution* 27:637–647.
- Sultan, S. E. 1987. Evolutionary implications of phenotypic plasticity in plants. *Evolutionary Biology* 21:127–178.
- 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.
- Via, S. 1993. Adaptive phenotypic plasticity: target or by-product of selection in a variable environment. *American Naturalist* 142:352–365.
- Via, S., R. Gomulkiewicz, G. DeJong, S. M. Scheiner, C. D. Schlichting, and P. H. van Tienderen. 1995. Adaptive phenotypic plasticity: consensus and controversy. *Trends in Ecology and Evolution* 10: 212–217.
- Via, S., and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39:505–522.

Associate Editor: Anurag Agrawal
 Editor: Troy Day



THE HORNED CORYDALUS.

“One of the largest and most formidable looking, though perfectly harmless, insects we have is the *Corydalus cornutus*. Its large size, its broad net-veined wings and slow-stupid flight, and aquatic habits, besides many other characteristics, place it very low in the scale of insect life.” From “Zoology” (*The American Naturalist*, 1867, 1:434–439).