Cryptic local adaptation is pervasive across environmental and evolutionary gradients

Morgan M. Sparks[1,*], Joshua C. Kraft[2], Kliffi M. Subida Blackstone[1], Gordon G. McNickle[2,3], Mark R. Christie[1,4]

1. Department of Biological Sciences, Purdue University; 915 W. State St., West Lafayette, Indiana 47907-2054 USA

2. Department of Botany and Plant Pathology, Purdue University; 915 W. State St., West Lafayette, Indiana 47907-2054 USA

3. Purdue Center for Plant Biology, Purdue University; 915 W. State St., West Lafayette, Indiana 47907-2054 USA

4. Department of Forestry and Natural Resources, Purdue University; 715 W. State St., West Lafayette, Indiana 47907-2054 USA

*corresponding author: Morgan M. Sparks; Department of Biological Sciences, Purdue University; 915 W. State St., West Lafayette, Indiana 47907-2054 USA; msparks1309@gmail.com; 360-791-1369

Keywords: countergradient variation, cogradient variation, genetic assimilation, maladaptive plasticity, perfect compensation, phenotypic plasticity

Abstract

Cryptic local adaptation—where an environmental effect masks the magnitude of the genetic contribution to a phenotype—has long been a topic of interest in ecology and evolution. Here, we systematically document the magnitude and drivers of two common forms of cryptic local adaptation—counter- and cogradient variation. Using a hierarchical Bayesian meta-analysis, we calculated the overall effect size as 1.03 for countergradient variation and 0.90 for cogradient variation. This result indicates that the genetic effect of cryptic local adaptation is approximately equal to a one standard deviation change in trait value between the most disparate populations. We also found that the abiotic and biotic covariates with the largest mean effects were temperature (2.50) and gamete size (2.78), although there was substantial variance. Our results demonstrate the pervasiveness and large effect of cryptic local adaptation in wild populations and underscores the importance of accounting for these effects in future studies.

Introduction

The persistence of populations via adaptation to local environmental conditions remains a central issue in ecology and evolutionary biology (Blanquart *et al.* 2013). Environmental gradients, where abiotic factors such as temperature, salinity, or nutrient availability vary in a predictable manner, provide a framework to formulate and test questions about the pattern and magnitude of local adaptation. Pairing environmental gradients with common garden or reciprocal transplant studies, especially when there are many populations distributed along these gradients, can lead to insights into how local adaptation covaries with specific environmental factors (Levins 1968; Hargreaves *et al.* 2020).

Shared patterns of phenotypic variation along ecological gradients spurred notable research into the environmental and genetic determinants of body size and morphology and led to the establishment of eco-geographical "rules" such as Bergmann's, Allen's, and Hesse's rules and the subsequent deviations therein (Bergmann 1848; Allen 1877; Atkinson & Sibly 1997; Blanckenhorn 2004; Gaston *et al*. 2008; Tian & Benton 2020). Foundational work by Levins (1968, 1969) revealed a pattern whereby fruit flies along an elevation gradient expressed a similar body size in their wild home environments, leading to the conclusion that the trait was canalized. However, a subsequent study revealed the higher elevation populations were much smaller than lowland populations in common environments—a response believed to be related to desiccation—revealing these populations were locally adapted to different elevations and their environmental conditions and expressed disparate phenotypes across all common environments. In concert, these findings provided the first evidence of cryptic local adaptation.

Local adaptation is commonly expressed with reaction norm plots, where the trait value of a genotype is plotted as a function of multiple environmental values (West-Eberhard 2003). In

a reaction norm, the slope of a genotype's response is representative of phenotypic plasticity (the effect of the environment) and the difference between genotype-specific reaction norms at each environmental variable is representative of the genotypic difference (the genetic effect). When different genotypes express the same reaction norm (i.e., overlapping lines, Fig. 1a, plasticity alone) this is representative of a purely environmentally driven response where distinct genotypes would express the same trait values (or phenotype) along an environmental gradient. Alternatively, populations may be locally adapted, such that the reaction norms cross when comparing different common environments, which is indicative of a genotype by environment interaction (Fig. 1b, Genetic x Environment). Finally, populations may exhibit cryptic local adaptation, where the genetic and environmental effects covary with one another, either negatively or positively (Conover *et al*. 2009). In this scenario, the reaction norms for populations may have the same or similar slopes thus maintaining consistent rank order in their trait values (Fig.1, co- and countergradient variation). For example, a high latitude population may have consistently larger body sizes across environments than a low latitude population, despite both expressing plasticity in body size. When the genetic portion of phenotypic variance opposes with the environmental portion, it is called countergradient variation (Fig. 1c), and when those two factors act in the same direction it is termed cogradient variation (Fig. 1d) (Levins 1968; Conover & Schultz 1995). That is, when comparing among wild populations, phenotypic traits appear more similar than they would if populations were compared in a common environment for countergradient variation, and vice versa for cogradient variation.

Co- and countergradient variation constitute forms of cryptic local adaptation because their true genetic effect on phenotypic variation is masked by the environment (Conover & Schultz 1995; Schlichting 2008; Conover *et al*. 2009; Kinnison *et al*. 2015) (Fig.1).

Countergradient variation can be particularly difficult to identify in wild populations because the environmental influence on a phenotype can conceal the genetic influence, such that the signal of local adaptation only appears when individuals are reared in common environments. This response has been termed perfect compensation (or exact compensation), when wild-type phenotypes are roughly equal across a gradient in their home environments (Fig. 1c) (sensu Conover *et al.* 2009). Compensation can be thought of as a measure of how similar phenotypic values are across populations in their home environments—and as an indicator of whether selection may favor similar or disparate phenotypes along an environmental gradient.

While Levins (1969) was the first to formally describe countergradient variation—which was followed by early research by Berven and others (Berven *et al.* 1979; Berven 1982) in *Ranid* frogs—it wasn't until a two-and-a-half decades later than interest in cryptic local adaptation grew substantially. This research was led by a group investigating a coastal, eastern North American fish, Atlantic silverside (*Menidia menidia*), after they discovered that populations had similar body sizes along a latitudinal gradient that spanned from the Gulf of Maine to the Floridian Atlantic Coast. Using a common garden approach, they found that the northern populations were consistently larger than southern populations, regardless of the temperature in which they were raised (*e.g.*, Fig. 1c). This finding led to substantial follow up research investigating the proximate and ultimate factors driving the trend (Conover & Present 1990; Conover *et al.* 2005), namely faster growth and more efficient metabolism in response to size-dependent winter mortality. After a theoretical review outlining the evolutionary significance of countergradient variation (Conover & Schultz 1995), subsequent researchers added to the evidence for co- and countergradient variation across diverse taxa including plants, bivalves, fishes, and insects (*e.g.*, *plants*, Chapin & Oechel 1983; *bivalves*, Dittman 1997; *fishes*,

DiMichele & Westerman 1997; Imsland *et al.* 2000; Jean M. Billerbeck *et al.* 2000; Jonassen *et al.* 2000; *insects*, Blanckenhorn 1991; Armbruster & Conn 2006; Pöykkö & Tammaru 2010), and across ecological gradients such as temperature, salinity, carotenoid availability, and urbanization (*e.g., temperature*, Yamahira & Conover 2002; Kokita 2003; *salinity*, Secor *et al.* 2000; *carotenoid availability*, Craig *et al.* 2005; Grether 2005; and *urbanisation*, Tüzün *et al.* 2017).

Despite the attention paid to counter- and cogradient variation, there have been few studies that describe the magnitude and extent of these types of local adaptation, nor has there been much investigation into the broad abiotic and biotic factors that might contribute to their occurrence (Conover *et al.* 2009; Hendry 2017; Stamp & Hadfield 2020). Without understanding the extent of cryptic local adaptation, it is difficult to make predictions about how populations expressing cryptic local adaptation may interact with a changing environment or how local adaptation may aid or hinder conservation goals. Here we use a Bayesian meta-analytical framework with hierarchical models to describe the general effect size of both counter- and cogradient variation as they appear in the literature and to investigate what biotic and abiotic factors are associated with their occurrence. The goals of these analyses are to *i)* quantify the effect size of countergradient and cogradient variation, *ii)* describe these effect sizes within and across taxa, phenotypic traits, and different environmental gradients, and *iii)* quantify the effect size of compensation for the home environment phenotypes to assess if there are clear trends for phenotypic similarity or dissimilarity along gradients.

Methods

Literature Review and Selection

Overall, we collected 858 studies that likely investigated counter- and/or cogradient variation. To generate a full list of studies, we first analyzed studies indicated as showing counter- or cogradient variation in the qualitative analysis from Conover *et al.* (2009). We then searched Web of Science topic field using the terms "counter*gradient variation" in May of 2018 and "co*gradient variation" (while excluding the previous search, as the wildcard renders them redundant) in June of 2019, resulting in 384 and 34 results, respectively. In August 2019, we also included studies citing the Conover *et al.* (2009) review, as well as the earlier review article by Conover & Schultz (1995)—682 studies total (some of which were redundant with studies in the prior search). These methods are discussed in more depth in the Supplementary Methods.

To be included in our analysis, a study had to meet a set of qualifying criteria. Specifically, a study had to be designed to include some form of a common environment—this may have been in a common garden study or a reciprocal transplant. Additionally, studies must have included two or more common environments to determine the signal of local adaptation (countergradient, cogradient, or others such as GxE that were ultimately not included in this study), as well as two or more populations to compare the genotypic difference of the phenotypic response between populations. Occasionally, studies investigated these distinctions across multiple species in the place of populations and such studies were discarded. Furthermore, each study must have included some mean estimate for a population phenotypic value (or data that could be used to calculate a mean), a measure of variation that could be transformed into units of standard deviation, and sample sizes for these statistics. When studies reported ranges an estimate for SD was made by subtracting the maximum value from the minimum value and then dividing by four as in Davidson *et al.* (2011).

Statistical Methods

To calculate effect sizes, we took data from both tables and figures in manuscripts. For data in the form of figures, we used the software WebPlotDigitizer (Rohatgi 2020) to extract the relevant data. Briefly, the program works by importing a photo of the respective figure and the user manually sets the scales for the x and y-axes based on the figure values and then adds digital points with their cursor to the data of interest. The digital points are then translated to numeric data via the imaging algorithm.

The statistic used to calculate effect size was Hedge's *d*, which is a standardized mean difference with a correction for small sample sizes (Hedges & Olkin 1985; Rosenberg *et al.* 2013). We calculated standardized mean difference (synonymous with effect size in this paper) by comparing the most extreme populations at each treatment level. For example, if we had three populations collected from an elevational gradient at 500m, 1000m, and 1500m in common gardens with three temperature treatments (e.g., 10, 15, and 20 °C) (as was often the case where temperature was a proxy for latitude or elevation), then we would calculate the mean and variance for every population in every treatment, but calculate the effect size (standardized mean difference) by using the values from the 500m and 1500m populations at each treatment. Occasionally, not all populations were used in all treatments, in these scenarios we used only the treatments for the most extremely distributed (with respect to the gradient—*e.g.*, very highest and lowest elevations) populations. The multiple values for a given trait in a given study (for example, the mean effect size of those that were calculated at 10, 15, and 20 °C in our prior example) were averaged as a single value (Borenstein 2009). Some studies conducted multiple experiments measuring the same trait. If multiple experiments measuring the same trait were recorded, we calculated the mean effect sizes for each experiment independently.

In all, we calculated 422 total effect sizes and 119 averaged effect sizes from 1204 individual data points extracted from 83 studies. Many studies showed counter- or cogradient patterns in multiple traits and/or trials within a given study. We collected all these data and controlled for their non-independence using a robust, Bayesian hierarchical modeling approach (discussed below). Hedge's *d* was calculated using the *metafor* package in program R (version 3.6.3) (R Core Team 2020) with the *escalcl()* function (Viechtbauer 2010). Additional metadata was collected from each study and included features such as year the study was published, the species investigated, and the distance in meters (for elevational gradients) and kilometers (for latitudinal gradients), the full list of which are available in our data file metadata tab (Supplementary Data). These data were collected to be used as covariates in the mixed-effects model (Table 1).

To estimate overall effect size and the effect of covariates, Bayesian linear mixed models fit with random intercepts were used in the R package *brms* (Bürkner 2017, 2018), which is a high-level interface in R for the Stan modeling language (Stan Development Team 2021). Linear mixed-effects models were chosen due to the hierarchical nature of this data set, where a study may have effect sizes estimated for multiple traits in multiple experiments. As such, a nested random intercepts effect was used to control for this non-independence. We originally set out to nest experiment within trait within study as the most appropriate reflection of our data's non-independence, but that structure proved to be too computationally intensive because model parameterization necessary to overcome divergent transitions (specifically, max tree depth in Stan) extended model run time beyond available compute power on our computational cluster. As such, random effects were reduced to traits nested in studies.

Similarly, phylogenetic relationships may provide another source of non-independence, and was controlled for by using a phylogenetic random intercepts effect for which we created the variance-covariance matrix (Lajeunesse *et al.* 2013). To do so, we generated a pseudo-phylogeny using the *taxize* package in R (Chamberlain & Szöcs 2013). Using the *classification()* function, we extracted the full taxonomy of a given organism in our dataset from the NCBI database. We then used the *class2tree()* function to generate a phylogenetically accurate tree in terms of taxonomic rank (*e.g.*, all species within a genus would have the same distance from one another, all genera within a family would have the same distance—a full explanation of the method is available in the Supplementary Methods). The function not only clusters taxa on their specific taxonomic rank but also within the taxonomic clade, allowing for more accurate distance measures. The tree created using this method was made into a matrix and then used as a custom variance-covariance component in our hierarchical model. The mathematical representation of our random effects in all models is shown in our Supplementary Materials (Supplementary Methods).

We used weakly to moderately informative priors for our hierarchical models (Williams *et al.* 2018)—except for the metaregression, where a stronger prior was used to reign in variance—lists of priors are available in Supplementary Materials (Tables S1-S3). For our models, we used four chains, with thousands of burn-ins, tens-of-thousands of iterations for each chain, with thinning set to 1 to 10 depending on the number of iterations in the model, resulting in >5,000 estimates for our posterior distributions (exact values are available in Table S4). Trace plots and posterior predictive checks are included in Figures S1-4.

Because we were interested in the magnitude of our effect and not the sign (*e.g.*, when calculating an effect size for developmental rate a countergradient response would be quicker

time to hatch relative to other populations and thus a negative number), we used log transformed absolute values of effect sizes as our response variable in order to use a Student's t-distribution and normally distributed priors (other candidate distribution types available in *brms*, such as lognormal, disallow weighting on variance, which is needed for a meta-analysis). As such, back-transformed values tended to stack near zero, but could not drop below zero, and distributions tended to have a long tail as numbers increased. Because we used a Bayesian framework, the following effects are presented with the means and 95 percent credible intervals (indicated as 95% CI) of the posterior distributions, unless otherwise noted.

The overall effect size for both co- and countergradient variation was estimated using random-effects only models, which is the convention when assuming study effects may vary in meta-analytical approaches (Mengersen *et al.* 2013). We included the covariates of trait, gradient, publication year, and taxonomic class in our metaregression countergradient model to assess how different biotic and abiotic factors may contribute to the estimate of countergradient variation effect size. This analysis was not performed with cogradient variation because of the lack of diversity of data to include many covariates in our modeling approach.

Compensation Analysis

Finally, we analyzed whether populations appeared to be under-, over-, or perfectly compensating (*sensu* Conover *et al.* 2009) when comparing their home environment phenotypes. To do so, we used the most extreme treatments and paired them with the polarly (as in opposites, not necessarily geographically) distributed populations used in the study. For example, if we consider three populations of fish from a range of saline conditions—freshwater, intermediary, and marine—then, the most saline treatment in the experiment would be considered the home environment of the marine population and the least saline the home environment of the

freshwater fish (unless other treatments were noted to be reflective of home environments or chosen to be specifically extreme). We then compared the trait values of the two polarly distributed populations and using the previous example, if the trait value was larger for the most saline population than that of the least saline population the relationship was classified as overcompensating, close to even was classified as perfectly compensating, and negative was classified as undercompensating. To standardize these data, we again used Hedge's *d*, because the values of different traits are generally measured on different scales (Safran *et al*. 2012). Importantly, compensation was not assessed with a model, as our analysis was meant to be exploratory and because there are no quantitative definitions of compensation. The results were classified as overcompensating if values were >0.5, perfectly compensating if values were between -0.5 and 0.5, and undercompensating if values were < -0.5. Because compensation measurements were computed as Hedge's *d*, these were meant to reflect a broad distribution around 0 for perfect compensation (what amounts to a medium effect size in both positive and negative directions, Sawilowsky 2009) that encapsulates the inherent measurement error in the value.

Results

In all, our systematic review considered 858 studies. We included 83 studies that either demonstrated co-, countergradient variation, and met the requirements for inclusion in our analysis. From those, we calculated 1,204 individual trait values, and 1,024 were for traits that were exhibiting countergradient variation and 180 for cogradient variation.

Random Effects Models

The result of our random effects models that estimated overall effect size for countergradient variation was 1.03 (95% CI = 0.56-1.59) and for cogradient variation 0.90 (95% CI = 0.22-2.50)

(Fig. 2, non-truncated in Fig. S5). The heterogeneity parameter τ—the full random effect standard deviation—was modeled with its own prior distribution and resulted in 2.32 (95% CI = 1.73-3.02) for countergradient variation and 2.96 (95% CI = 1.13-6.69) for cogradient variation. We similarly included a random effect for phylogenetic relatedness which resulted in an estimate of 1.37 SD (95% CI = 1.01-2.26) for countergradient variation and 2.73 SD (95% CI = 1.03-8.36) for cogradient variation (Tables S5, S6).

Metaregression

We also implemented a mixed effects metaregression with our countergradient dataset to estimate the effect size for trait classification, gradient classification, and taxonomic class effect. The mean effect, posterior distribution and 95% credible intervals for each modeled covariate are presented in Table 1 and those same values are represented as the main effects estimated marginal means with the posterior distribution and credible intervals in Figure 3 (non-truncated in Fig. S6). For each covariate, these results illustrate similar mean values with large and overlapping credible intervals. While variability was large across all covariates, the mean coefficient estimates of the posterior distributions for temperature (mean = 2.50, 95% CI = 0.74-6.19) and gamete size (mean = 2.78, 95% CI = 0.54-8.74) were much larger than the global estimate of 1.01. By contrast, the smallest value for urbanization (mean = 0.75, 95% CI = 0.14-2.46)—which had lower credible intervals near zero and was much lower than other estimates. The between trait (nested in study) standard deviation was 2.28 (95% CI = 1.90-2.86) and the standard deviation for phylogenetic random effect was 1.51 (95% CI = 1.01-2.77). Note, these estimates are slightly different than those presented in Figure 3, which are the main effect estimated marginal means derived from the model, rather than model coefficients.

Compensation Analysis

The results of our compensation analysis were plotted to represent the conceptual relationship between home phenotypes of populations exhibiting countergradient variation (*sensu* Conover et al. 2009) (Fig. 4). Our data provide support for far more frequent over- and under compensation than perfect compensation, meaning that populations in our study generally have disparate home phenotypes as opposed to equal ones. We considered any value between -0.5 and 0.5 as perfectly compensating (or a range of 1 effect size around 0 for a conservative estimate), values less than -0.5 undercompensating, and values larger than 0.5 overcompensating. We consider this classification conservative because it allows for a medium effect size in the positive or negative direction from zero (Sawilowsky 2009), despite the theoretical description of perfect compensation being a value equal to zero (Conover et al. 2009). The distribution of these values in our data set were overcompensating = 38.6%, undercompensating = 36.0%, and perfectly compensating = 25.4%. Effect sizes ranged from -70.38 to 36.45 but 5% and 95% quantiles were -27.74 and 6.49, respectively.

Discussion

The ecological and evolutionary importance of co- and countergradient variation has been recognized for three decades (Conover & Present 1990; Conover & Schultz 1995). Here, we find that the overall effect size of co- and countergradient variation are 1.03 and 0.90, respectively. Our measure of standardized mean difference, Hedges' *d*, essentially translates the difference between populations into units of standard deviation. Thus, our results indicate that the standardized mean difference between the most disparate populations in our study—*i.e.*, the genetic effect of cryptic local adaptation—is approximately equal to a one standard deviation change in trait value. This result demonstrates that genetic differentiation among cryptic, locally

adapted populations is large but also obscured as a function of the covariance between environmental and genetic effects and their interaction with phenotype.

We used a metaregression to investigate the effect sizes of abiotic and biotic covariates of countergradient variation; covariates were statistically indistinguishable from one another due to high variance in their estimate, though many showed values that were large with distributions away from 0 (Effect sizes estimates of 1-3, Table 1, Fig. 3). Instead, we find that the inter-effect variability (τ^2, or the variability of our random effects) generally swamped the covariate-specific effects estimates in our metaregression (Table 1). While the lack of predictive results in our metaregression may be indicative that our data is not robust enough for differences among covariates to emerge, we think it is more likely that these results suggest that countergradient variation, when it does occur, is an ecologically ubiquitous phenomenon for species distributed along environmental gradients. Furthermore, we suspect the mechanism driving such adaptation is different among populations and gradients and therefore differs in its effect contributing to the variability observed in our models. In other words, while the phenotypic solution may appear similar across species and gradients, the processes driving the genetic divergence of the phenotypes of different populations is idiosyncratic. This conclusion is further supported by the lack of pattern with respect to latitudinal distance and temperature difference regressions analyses, neither of which were predictive (Fig. S7, Table S7). Thus, while populations almost always showed a large effect of cryptic local adaptation—model covariates were 2.50 and 0.99, respectively—the magnitude did not scale with temperature or physical distance.

Despite the fact that the overall effect size of cogradient variation was large, our results for cogradient variation were only based on 15 studies. This paucity reflects an overall lack of published research investigating and demonstrating cogradient variation (see also, Conover *et al.*

2009). However, in a meta-analysis of reciprocal transplant studies, Stamp & Hadfield (2020) found 60% of traits exhibited a cogradient pattern. Our results (*i.e.*, the magnitude of effect size), in conjunction with the work of Stamp & Hadfield (2020) suggest that, while rarely reported, cogradient variation constitutes an ecologically meaningful phenomenon. We further hypothesize cogradient variation may suffer from underreporting because it's natural sign mirrors environmentally induced plasticity that may not drive researchers to investigate populations using common gardens. While cogradient variation may be underreported, work such as that by Trussell & Etter (2001) show both the large effect of cogradient variation and how home phenotypic differences can disguise that effect. In a reciprocal transplant study, the researchers measured shell thickness in marine snail populations, finding that, when in their home environments, populations expressed a 3.18 standardized mean difference in shell thickness. But, when considered in shared environments—or the genetic difference due to cogradient variation—that difference reduced to 2.07. Notably, this amounts to a substantial effect size difference due to genetic variation alone, but is ~2/3 value if we observed the populations only in their home environments. More importantly, if the researchers had only investigated these populations in their home environments without common gardens, the obvious interpretation for the phenotypes would have been a strong plastic response, which, while true, did not account for even half of phenotypic variation in home environments.

Our compensation analysis for countergradient variation revealed that populations in their home environments were more commonly over- or undercompensasting, as opposed to perfectly compensating (Fig. 4). In other words, the genetic effect of countergradient variation was generally not equal to that of the environmental effect. This result was clear with conservative bounds for perfect compensation (-0.5 to 0.5 standardized mean difference). Reducing those

bounds to the less conservative values of -0.2 to 0.2 (considered to be a small effect size, Sawilowsky 2009) for perfect compensation resulted in only 11.7 percent of studies demonstrating perfect compensation. While we caution that these results may not necessarily point to clear biological trends about how common and in what scenarios over- and under-compensation may occur, it does suggest that either: 1. selection for maintaining optimal phenotypes across broad environmental gradients may not be the norm (*sensu* Conover & Schultz 1995; Conover *et al.* 2009), or 2. the response to selection may not be strong enough to effectively minimize phenotypic variation across populations. If the latter, the response to selection may be minimized by insufficient adaptive genetic variation or countervailing selection on correlated traits specific to each local population (Lande & Arnold 1983).

One of the confounding aspects of investigating clinal trends in locally adapted populations is the often non-continuous nature of their phenotypic values. For example, the non-continuous nature of population responses is particularly clear in Hice *et al.* (2012), in which they investigated 37 sites across the entire range of Atlantic silversides and found an overall trend of countergradient variation in growth rate, but multiple breakpoints along the cline. The results presented in Hice *et al.* (2012), constitutes one of the few attempts to characterize cryptic local adaptation across the full environmental gradient for a species or set of populations. However, if the full distribution of populations is ignored or nor correctly accounted for, such sample designs may present a false signal of cryptic local adaptation. Few studies in this analysis included measures of genetic differentiation among the populations they were measuring. For example, hypothetical meta-populations distributed along a gradient such as latitude with little or no gene flow between them may experience divergent selection on the same trait but from different modes of selection and could appear as cryptic local adaptation.

The observation that change in treatment temperature or latitude was not predictive of effect size may be explained by the heterogeneity of environments (Fig. S7). A useful example of why this result may occur is microgeographic variation, in which large adaptive phenotypic divergence occurs over small spatial scales where environmental factors may be more disparate than initially predicted. Indeed, microgeographic variation was observed in some of the studies that were reviewed as part of this study. Researchers found that in temperature choice trials, wood frogs (*Rana sylvatica*) captured from ponds with more forest canopy, preferred higher temperatures (Freidenburg & Skelly 2004) across a single 3800-hectare forest. Companion work from a common garden with two temperature treatments, found that frogs from more shaded ponds developed faster than those from more open ponds (Skelly 2004). While broad patterns in temperature may make for useful categorizations of clines, the habitat level specifics with respect to local adaptation may be more complicated.

Clinal variation has been the subject of numerous neutral and quantitative genetic studies (Richter-Boix *et al.* 2010, 2013; Hangartner *et al.* 2012; Frei *et al.* 2014), however, the pace of genomic investigation into cryptic local adaptation lags behind other forms of local adaptation. In fact, fundamental theory about countergradient variation often assumes directional selection drives the maladapted population/s opposite of the environmental effect to achieve the negative covariance that characterizes countergradient variation (Grether 2005; Crispo 2007). But, work by Wilder *et al.* (2020), found substantial genomic divergence at specific chromosomal sites among northern and southern populations of Atlantic silverside, despite these being continuously distributed populations with substantial gene flow and low overall differentiation. These most differentiated genomic sites were largely driven by a single population, with the warmest temperatures, distinguishing it from all the other more northern populations. These results

suggest that strong selection maintains these regions, despite genomic homogenization from gene flow. This research is only beginning to scratch the surface of understanding the functional processes at play at the genomic level in counter- and cogradient variation, and what tradeoffs may exist in shaping the development of local adaptation.

Conclusion

Our findings illustrate that cryptic local adaptation is pervasive across multiple environmental gradients in a diverse set of taxa. With this meta-analysis, we quantitatively characterized the genetic effect on phenotypic divergence in cryptic local adaptation and find that the result is at or close to one full standard deviation from the mean trait value of a population. While the effect of co- and countergradient variation may be large, the variability of response observed across studies is also quite high, as indicated in the values of our random effects. Moreover, when analyzing the mean trait difference between home phenotypes of populations exhibiting countergradient variation, we find no support for a strong signal of perfect compensation, indicating that selection for optimal phenotypes across a gradient does not or cannot minimize variation in mean phenotypic trait values. We conclude that the cryptic local adaptation is both ubiquitous and has a large effect size and argue that cryptic local adaptation deserves continued attention in theoretical, experimental, and applied approaches. Undesired consequences can occur when cryptic local adaptation is ignored, (*e.g.*, Philipp & Whitt 1991) and could lead to especially problematic outcomes in conservation and/or translocation-oriented applications. Much like local adaptation more generally, cryptic local adaptation necessitates close consideration when working with populations distributed across ecological gradients.

Acknowledgements

We thank N. Emery and C. Oakley for their helpful discussion regarding cryptic local adaptation. K. Horvath offered useful suggestions about instituting a phylogenetic random effect in our models, D. Pascall provided valuable insights on implementing the pseudo-phylogeny, and R. Swihart gave useful direction on Bayesian model setup. We also thank A. Harder, A. Lee, and A. Nalesnik for their helpful comments on a past version of this manuscript. This work was supported, in part, by funding from the National Science Foundation (OCE-1924505 ; DEB-1856710) to MRC.

Data Archiving statement

Code for these analyses is posted on the GitHub account of MMS (https://github.com/morgan-sparks/CnGV-CoGV-Meta-Analysis) and cloned at the GitHub account of MRC. Data and model files will be stored permanently at DataDryad.

Author Contributions

MMS and MRC designed the research, MMS, JRK, and KMSB collected data, MMS performed the data analysis, and MMS, GGM, and MRC wrote the manuscript with input from all authors.

References

Allen, J.A. (1877). The influence of physical conditions in the genesis of species. *Radical review*, 1, 108–140.

Armbruster, P. & Conn, J.E. (2006). Geographic Variation of Larval Growth in North American *Aedes albopictus* (Diptera: Culicidae). *Annals of the Entomological Society of America*, 99, 1234–1243.

Atkinson, D. & Sibly, R.M. (1997). Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends in Ecology & Evolution*, 12, 235–239.

Bergmann, C. (1848). *Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse*.

Berven, K.A. (1982). The genetic basis of altitudinal variation in the wood frog Rana sylvatica II. An experimental analysis of larval development. *Oecologia*, 52, 360–369.

Berven, K.A., Gill, D.E. & Smith-Gill, S.J. (1979). Countergradient selection in the green frog, *Rana clamitans*. *Evolution*, 33, 609–623.

Blanckenhorn, W.U. (1991). Life-History Difference in Adjacent Water Strider Populations: Phenotypic Plasticity or Heritable Responses to Stream Temperature? *Evolution*, 45, 1520.

Blanckenhorn, W.U. (2004). Bergmann and Converse Bergmann Latitudinal Clines in Arthropods: Two Ends of a Continuum? *Integrative and Comparative Biology*, 44, 413–424.

Blanquart, F., Kaltz, O., Nuismer, S.L. & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology Letters*, 16, 1195–1205.

Borenstein, M. (Ed.). (2009). *Introduction to meta-analysis*. John Wiley & Sons, Chichester, U.K.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80, 1–28.

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10, 395–411.

Chamberlain, S.A. & Szöcs, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Res*, 2, 191.

Chapin, F.S. & Oechel, W.C. (1983). Photosynthesis, respiration, and phosphate absorption by *Carex aquatilis* ecotypes along latitudinal and local environmental gradients. *Ecology*, 64, 743–751.

Conover, D. & Present, T. (1990). Countergradient variation in growth rate: Compensation for length of the growing season among Atlantic silversides from different latitudes. *Oecologia*, 83, 316–324.

Conover, D.O., Arnott, S.A., Walsh, M.R. & Munch, S.B. (2005). Darwinian fishery science: lessons from the Atlantic silverside (*Menidia menidia*). *Canadian Journal of Fisheries and Aquatic Sciences*, 62, 730–737.

Conover, D.O., Duffy, T.A. & Hice, L.A. (2009). The covariance between genetic and environmental influences across ecological gradients: reassessing the evolutionary significance of countergradient and cogradient variation. *Ann. N. Y. Acad. Sci.*, 1168, 100–129.

Conover, D.O. & Schultz, E.T. (1995). Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends Ecol. Evol.*, 10, 248–252.

Craig, J.K., Foote, C.J. & Wood, C.C. (2005). Countergradient variation in carotenoid use between sympatric morphs of sockeye salmon (*Oncorhynchus nerka*) exposes nonanadromous hybrids in the wild by their mismatched spawning colour. *Biological Journal of the Linnean Society*, 84, 287–305.

Crispo, E. (2007). The Baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, 61, 2469–2479.

Davidson, A.M., Jennions, M. & Nicotra, A.B. (2011). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters*, 14, 419–431.

DiMichele, L. & Westerman, M.E. (1997). Geographic variation in development rate between populations of the teleost *Fundulus heteroclitus*. *Marine Biology*, 128, 1–7.

Dittman, D.E. (1997). Latitudinal compensation in oyster ciliary activity. *Functional Ecology*, 11, 573–578.

Frei, E.R., Hahn, T., Ghazoul, J. & Pluess, A.R. (2014). Divergent selection in low and high elevation populations of a perennial herb in the Swiss Alps. *Alp Botany*, 124, 131–142.

Freidenburg, L.K. & Skelly, D.K. (2004). Microgeographical variation in thermal preference by an amphibian: Microgeographical thermal preference. *Ecology Letters*, 7, 369–373.

Gaston, K.J., Chown, S.L. & Evans, K.L. (2008). Ecogeographical rules: elements of a synthesis. *Journal of Biogeography*, 35, 483–500.

Grether, G.F. (2005). Environmental change, phenotypic plasticity, and genetic compensation. *The American Naturalist*, 166, 115–123.

Hangartner, S., Laurila, A. & Räsänen, K. (2012). Adaptive divergence in Moor frog (*Rana arvalis*) populations along an acidification gradient: Inferences from Qst–Fst correlations. *Evolution*, 66, 867–881.

Hargreaves, A.L., Germain, R.M., Bontrager, M., Persi, J. & Angert, A.L. (2020). Local adaptation to biotic interactions: A meta-analysis across latitudes. *The American Naturalist*, 195, 395–411.

Hedges, L.V. & Olkin, I. (1985). Statistical methods for meta-analysis. Academic Press.

Hendry, A.P. (2017). *Eco-evolutionary Dynamics*. Princeton University Press.

Hice, L., Duffy, T., Munch, S., & David Conover. (2012). Spatial scale and divergent patterns of variation in adapted traits in the ocean. *Ecology Letters*, 15, 568–575.

Imsland, A.K., Foss, A., Nevdal, G., Cross, T., Bonga, S.W., Ham, E.A., *et al.* (2000). Countergradient variation in growth and food conversion efficiency of juvenile turbot. *Journal of Fish Biology*, 57, 1213–1226.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution*, 1, 1210-1226.

Jean M. Billerbeck, Eric T. Schultz & David O. Conover. (2000). Adaptive variation in energy acquisition and allocation among latitudinal populations of the Atlantic silverside. *Oecologia*, 122, 210–219.

Jonassen, T.M., Imsland, A.K., Fitzgerald, R., Bonga, S.W., Ham, E.V., Naevdal, G., *et al.* (2000). Geographic variation in growth and food conversion efficiency of juvenile Atlantic halibut related to latitude. *Journal of Fish Biology*, 56, 279–294.

Kinnison, M.T., Hairston Jr., N.G. & Hendry, A.P. (2015). Cryptic eco-evolutionary dynamics. *Annals of the New York Academy of Sciences*, 1360, 120–144.

Kokita, T. (2003). Potential latitudinal variation in egg size and number of a geographically widespread reef fish, revealed by common-environment experiments. *Marine Biology*, 143, 593–601.

Lajeunesse, M.J., Rosenberg, M.S. & Jennions, M.D. (2013). Phylogenetic nonindependence and meta-analysis. In: *Handbook of Meta-analysis in Ecology and Evolution* (eds. Koricheva, J., Gurevitch, J. & Mengersen, K.). Princeton University Press, pp. 284–299.

Levins, R. (1968). *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton University Press.

Levins, R. (1969). Thermal acclimation and heat Resistance in drosophila species. *The American Naturalist*, 103, 483–499.

Mengersen, K., Jennions, M.D. & Schmid, C.H. (2013). Statistical models for the meta-analysis of nonindependent Data. In: *Handbook of Meta-analysis in Ecology and Evolution* (eds. Mengersen, K., Koricheva, J. & Gurevitch, J.). Princeton University Press, pp. 255–283.

Philipp, D.P. & Whitt, G.S. (1991). Survival and growth of northern, Florida, and reciprocal Fl hybrid largemouth bass in central Illinois. *Transactions of the American Fisheries Society*, 120, 58–64.

Pöykkö, H. & Tammaru, T. (2010). Countergradient vs. cogradient variation in growth and diapause in a lichen-feeding moth, *Eilema depressum* (Lepidoptera: Arctiidae). *Journal of Evolutionary Biology*, 23, 1278–1285.

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

Richter-Boix, A., Quintela, M., Kierczak, M., Franch, M. & Laurila, A. (2013). Fine-grained adaptive divergence in an amphibian: genetic basis of phenotypic divergence and the role of nonrandom gene flow in restricting effective migration among wetlands. *Molecular Ecology*, 22, 1322–1340.

Richter-Boix, A., Teplitsky, C., Rogell, B. & Laurila, A. (2010). Local selection modifies phenotypic divergence among *Rana temporaria* populations in the presence of gene flow. *Molecular Ecology*, 19, 716–731.

Rohatgi, A. (2020). *WebPlotDigitizer*. Version 4.1. https://automeris.io/WebPlotDigitizer/

Rosenberg, M.S., Rothstein, H.R. & Gurevitch, J. (2013). Effect Sizes: Conventional choices and calculations. In: *Handbook of Meta-analysis in Ecology and Evolution* (eds. Gurevitch, J., Koricheva, J. & Mengersen, K.). Princeton University Press, pp. 61–71.

Safran, R., Flaxman, S., Kopp, M., Irwin, D.E., Briggs, D., Evans, M.R., *et al.* (2012). A robust new metric of phenotypic distance to estimate and compare multiple trait differences among populations. *Current Zoology*, 58, 426–439.

Sawilowsky, S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, 8, 597-599.

Schlichting, C.D. (2008). Hidden reaction norms, cryptic genetic variation, and evolvability. *Annals of the New York Academy of Sciences*, 1133, 187–203.

Secor, D.H., Gunderson, T.E. & Karlsson, K. (2000). Effect of temperature and salinity on growth performance in anadromous (Chesapeake Bay) and nonanadromous (Santee-Cooper) strains of striped bass *Morone saxatilis*. *Copeia*, 2000, 291–296.

Skelly, D.K. (2004). Microgeographic countergradient variation in the wood grog, *Rana sylvatica*. *Evolution*, 58, 160–165.

Stamp, M.A. & Hadfield, J.D. (2020). The relative importance of plasticity versus genetic differentiation in explaining between population differences; a meta-analysis. *Ecology Letters*, 23, 1432–1441.

Stan Development Team. (2021). *Stan Modeling Language*.

Tian, L. & Benton, M.J. (2020). Predicting biotic responses to future climate warming with classic ecogeographic rules. *Current Biology*, 30, R744–R749.

Trussell, G.C. & Etter, R.J. (2001). Integrating genetic and environmental forces that shape the evolution of geographic variation in a marine snail. In: *Microevolution Rate, Pattern,*

Process, Contemporary Issues in Genetics and Evolution (eds. Hendry, A.P. & Kinnison, M.T.). Springer Netherlands, Dordrecht, pp. 321–337.

Tüzün, N., Beeck, L.O. de, Brans, K.I., Janssens, L. & Stoks, R. (2017). Microgeographic differentiation in thermal performance curves between rural and urban populations of an aquatic insect. *Evolutionary Applications*, 10, 1067–1075.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36.

West-Eberhard, M.J. (2003). *Developmental Plasticity and Evolution*. OUP USA.

Wilder, A.P., Palumbi, S.R., Conover, D.O. & Therkildsen, N.O. (2020). Footprints of local adaptation span hundreds of linked genes in the Atlantic silverside genome. *Evolution Letters*, 4, 430–443.

Williams, D.R., Rast, P. & Bürkner, P.-C. (2018). Bayesian meta-analysis with weakly informative prior distributions. *PsyArXiv*. doi:10.31234/osf.io/7tbrm

Yamahira, K. & Conover, D.O. (2002). Intra- vs. interspecific latitudinal variation in growth: adaptation to temperature or seasonality? *Ecology*, 83, 1252.

Table Captions

Table 1: Results of Bayesian metaregression mixed model with effect size fit as a function of trait, gradient, and organism class. Random effects included trait nested in paper and paper alone, as well as a phylogenetic species effects. Shown is the intercept and the effect of each level of a covariate along with their estimated error and 95% credible intervals.

Figure Captions

Figure 1: Conceptual illustration of four types of phenotypic responses for two populations of a marine fish distributed along the Atlantic coast of the United States. Each row represents four related, but different responses: a) environment alone (phenotypic plasticity without local adaptation), b) a genetic by environment response, c) countergradient variation, d) and cogradient variation. The first two columns illustrate how each of the four responses manifest in different environments—shown is the expected phenotypic responses of both populations in their home environments ("wild") and if they were exposed to a reciprocal transplant experiment in their respective home habitats ("reciprocal transplant"). We also include the reaction norm for each response, where the base of the arrow indicates a population in its home environment and the arrowhead its away environment. Notice that for the counter-gradient variation example, both populations are the same size in the wild (i.e., perfect compensation), illustrating one way that local adaptation can be cryptic.

Figure 2: Posterior distributions of effect size estimates from random-effects meta-analysis for counter- and cogradient variation. The circle indicates the mean and the diamond the median of the posterior distribution (overlapping in countergradient) and the solid, black line, the 95% credible intervals. The mean and credible interval values are also indicated with text on the right. A total of 140 of the largest values (of 50,000 in each distribution) were removed because they extended beyond the x-axis margin (see Fig. S5 for all data points). Purple represents the posterior distribution for the effect size of countergradient variation and green, cogradient variation. The dashed line illustrates an effect size of 1.

Figure 3: Posterior distributions of estimated marginal means for the covariates in a metaregression analysis for countergradient variation. Yellow distributions depict taxonomic class, green distributions depict environmental gradients, and purple distributions depict trait covariates. The black circle indicates the mean of the posterior distribution (medians omitted because they overlapped with means). The thick solid, horizontal black line depicts the 50% credible intervals for each covariate and the grey, thinner line the 95% credible intervals (also indicated with text to the right). A total of the 314 largest values (of 60,000 thinned to 6,000 in each distribution) were removed because they extended beyond the x-axis margin (all points can be viewed in Figure S6). The dashed line illustrates each posterior distribution's relationship to the study-wide effect size of 1.03 estimated from the random effects model and shown in Fig. 2.

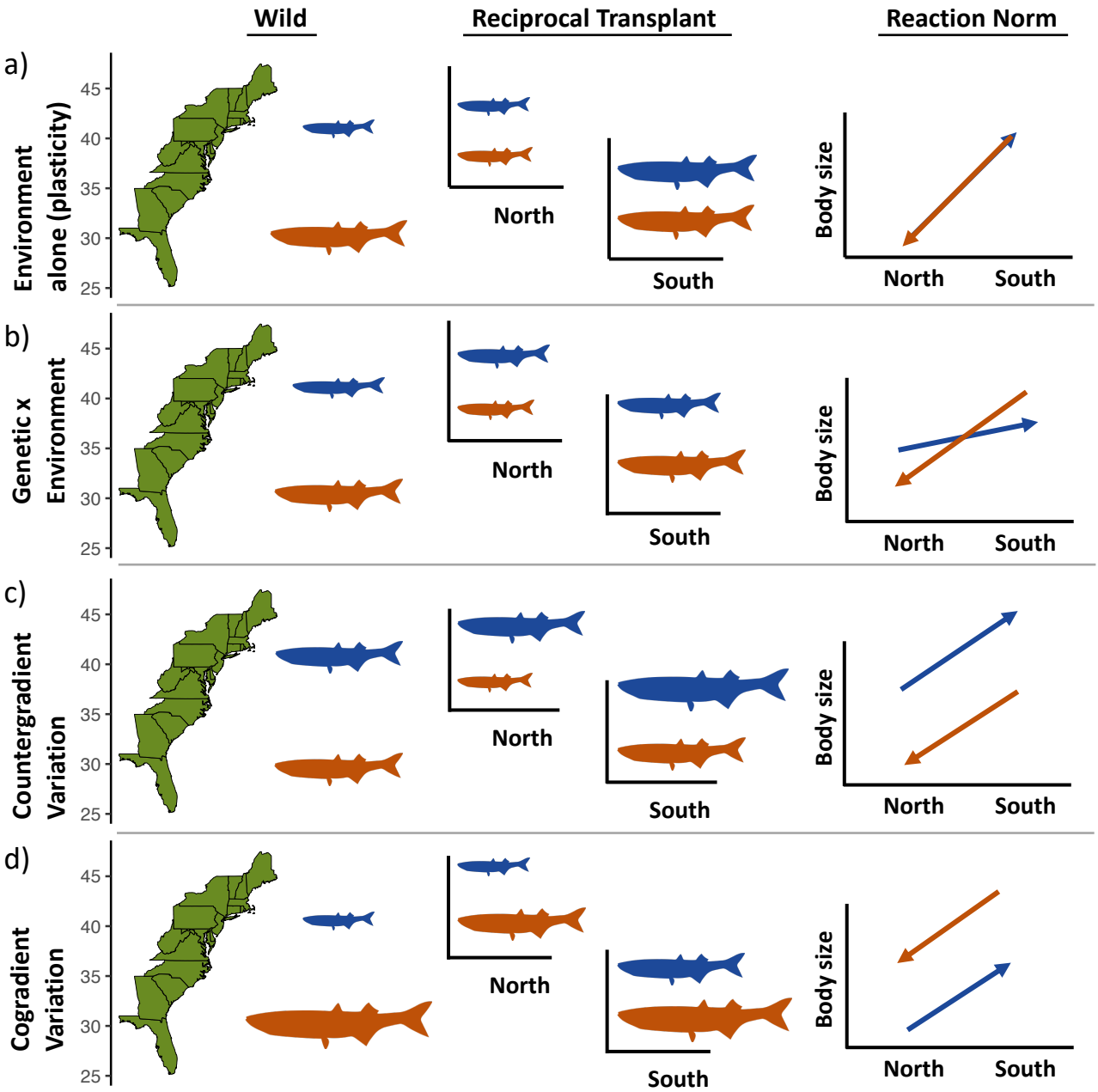
Figure 4: Conceptual representation of two countergradient populations in their home environments and the reaction norms of those populations in the home and away environments. Also shown are the results of our compensation analysis. a) Represents an overcompensating relationship where the northern population is larger than the southern population in all environments and the slope of the dashed line connecting their home phenotype (arrow bases) has a negative sign as it decreases from the larger northern population to the smaller southern population. b) Represents a perfectly compensating relationship, where phenotypes are equal in their respective home environments and thus the slope is flat. c.) Represents an undercompensating relationship, where the southern population is larger than the northern population in their respective home environments and the slope of the dashed line connecting their home phenotype (arrow bases) has a positive sign as it increases from the smaller northern population to the larger southern population. d.) represents the results of our compensation

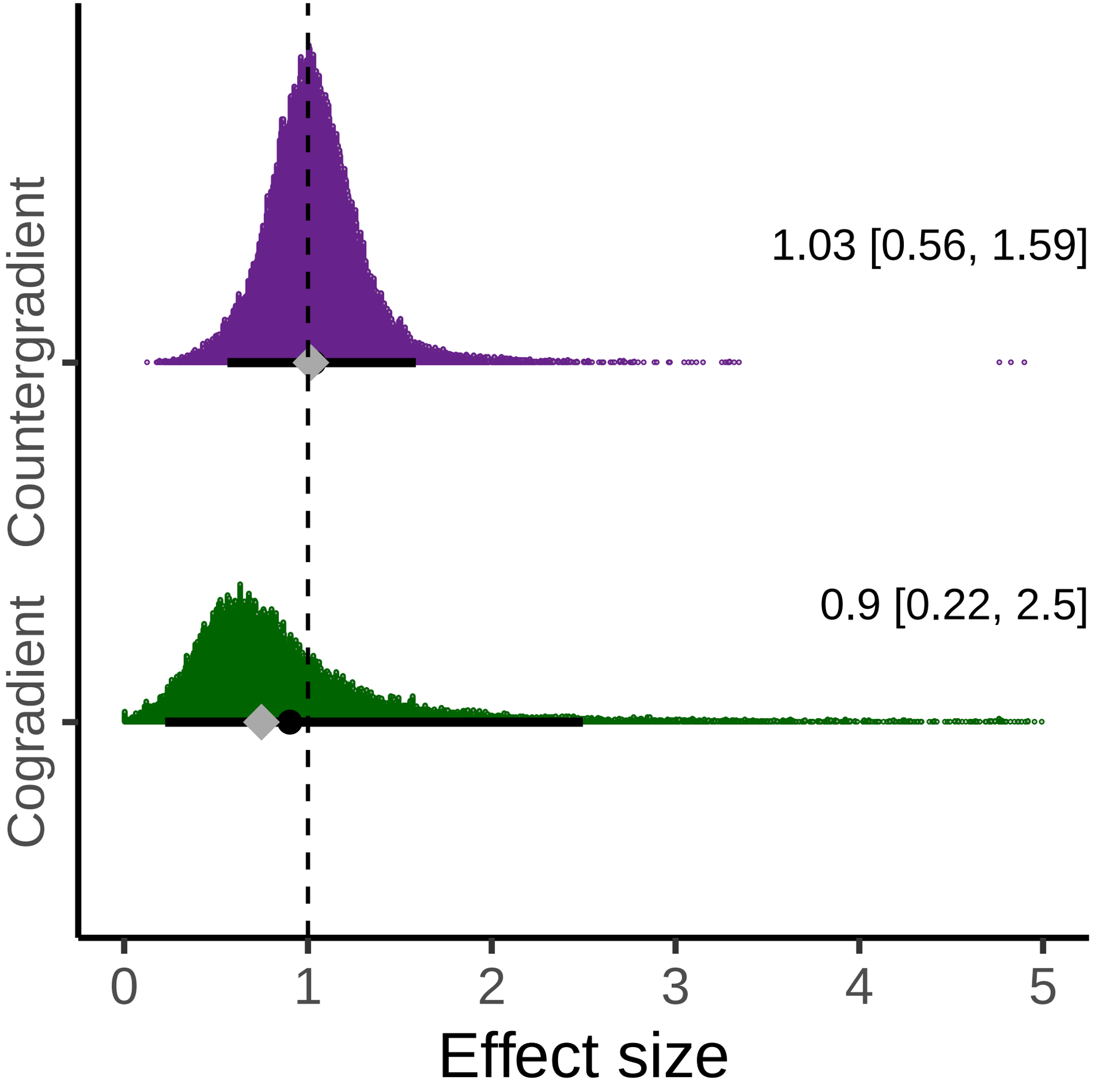
analysis where overcompensating was defined as an effect size > 0.5, perfect as 0.5 ≤ and ≥ -0.5, and undercompensating < -0.5. Of these three responses, our data was distributed as 38.6% overcompensating, 25.4% perfectly compensating, and 36.0% undercompensating. Points on the left are the calculated effect sizes and are connected to a point at 0 to show a slope to match the conceptual panels. The larger figure is bounded at -15 and 7 (removing 7 values) to better show the results near 0 and perfect compensation, whereas the inset figure includes all data.

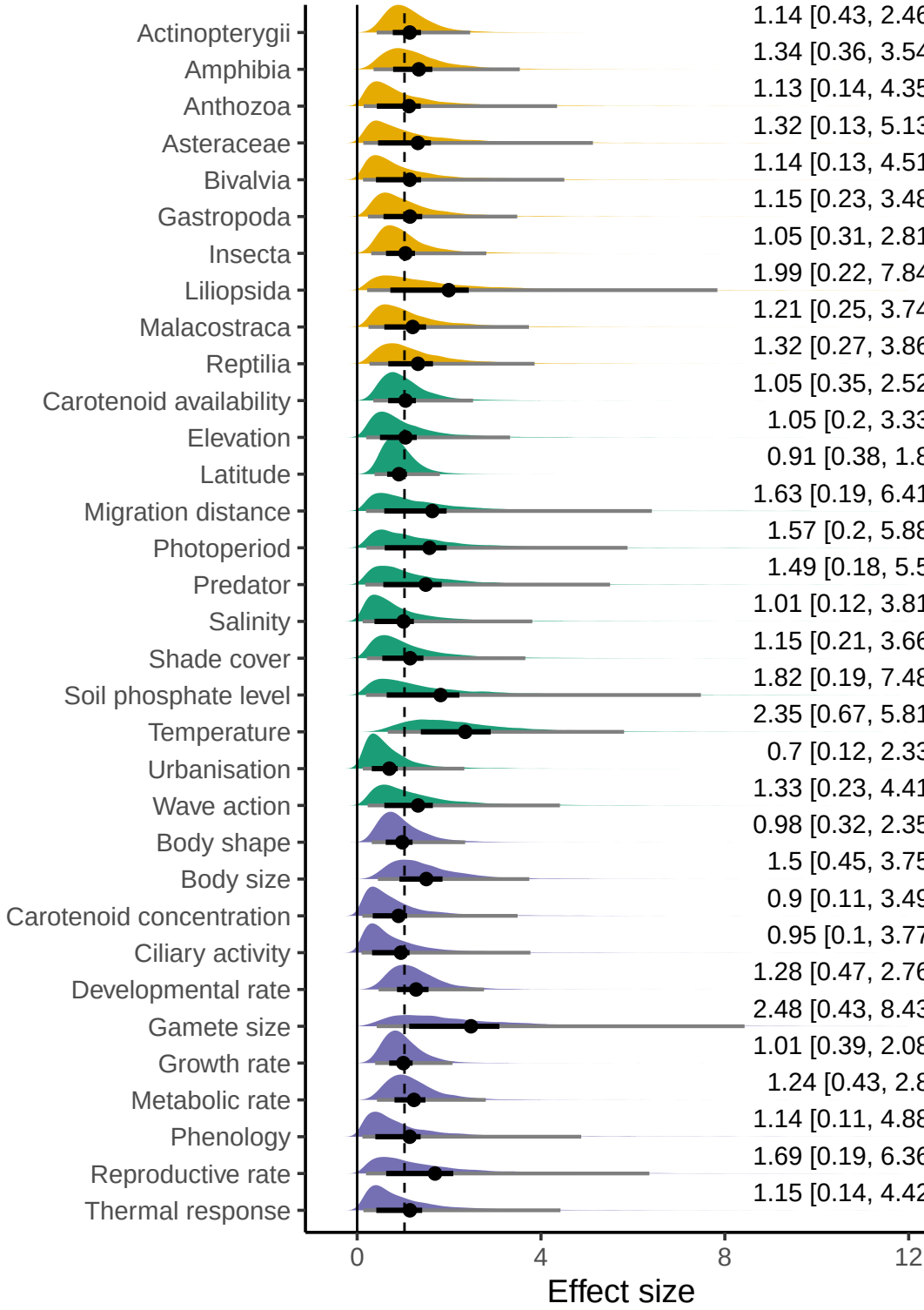
Table 1

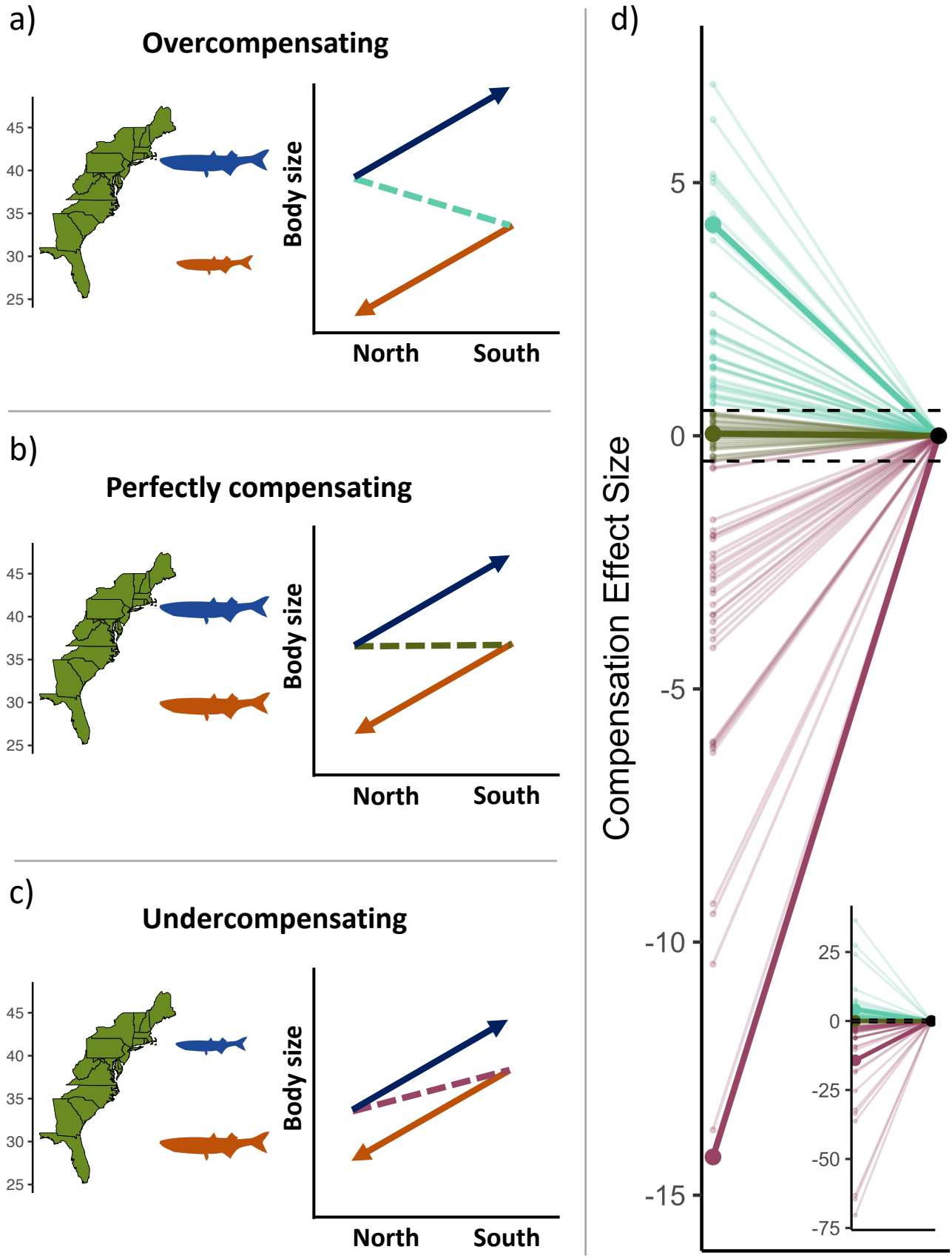
	Estimate	Est.Error	Q2.5	Q97.5
Intercept	1.09	0.74	0.27	2.98
Body size	1.72	0.91	0.56	4.03
Carotenoid concentration	0.98	0.94	0.15	3.37
Ciliary activity	1.03	1.05	0.13	3.79
Developmental rate	1.48	0.68	0.55	3.17
Gamete size	2.78	2.27	0.54	8.74
Growth rate	1.16	0.51	0.47	2.41
Metabolic rate	1.43	0.69	0.52	3.12
Phenology	1.27	1.38	0.15	5.16
Reproductive rate	1.82	1.73	0.28	5.97
Thermal response	1.25	1.22	0.18	4.52
Elevation	1.08	0.83	0.24	3.2
Latitude	0.99	0.44	0.38	2.1
Migration distance	1.65	1.72	0.23	6.11
Photoperiod	1.6	1.5	0.24	5.43
Predator presence	1.49	1.37	0.23	5.24
Salinity	1.01	0.91	0.15	3.47
Shade cover	1.19	0.89	0.26	3.59
Soil phosphate	1.86	1.92	0.23	7.09
Temperature	2.5	1.43	0.74	6.19
Urbanisation	0.75	0.67	0.14	2.46
Wave action	1.38	1.14	0.26	4.29
Amphibia	1.23	0.66	0.4	2.87
Anthozoa	1.02	0.97	0.15	3.55
Asteraceae	1.24	1.41	0.14	4.76
Bivalvia	1.04	1.12	0.14	3.8
Gastropoda	1.07	0.78	0.26	2.98
Insecta	0.97	0.53	0.36	2.35
Liliopsida	1.85	1.89	0.23	7.08
Malacostraca	1.1	0.74	0.28	3.07
Reptilia	1.21	0.78	0.32	3.21

SD Paper Number	2.31	0.37	1.65	3.11
SD Paper Number:Trait	2.28	0.25	1.9	2.86
SD Species	1.51	0.49	1.01	2.77


Wild
Reciprocal Transplant
Reaction Norm
a)
Environment alone (plasticity)
b)
Genetic x Environment
c)
Countergradient Variation
d)
Cogradient Variation
45
40
35
30
25
North
South
Body size
North
South


Countergradient
Cogradient
1.03 [0.56, 1.59]
0.9 [0.22, 2.5]
0
1
2
3
4
5
Effect size


Actinopterygii
Amphibia
Anthozoa
Asteraceae
Bivalvia
Gastropoda
Insecta
Liliopsida
Malacostraca
Reptilia
Carotenoid availability
Elevation
Latitude
Migration distance
Photoperiod
Predator
Salinity
Shade cover
Soil phosphate level
Temperature
Urbanisation
Wave action
Body shape
Body size
Carotenoid concentration
Ciliary activity
Developmental rate
Gamete size
Growth rate
Metabolic rate
Phenology
Reproductive rate
Thermal response
1.14 [0.43, 2.46]
1.34 [0.36, 3.54]
1.13 [0.14, 4.35]
1.32 [0.13, 5.13]
1.14 [0.13, 4.51]
1.15 [0.23, 3.48]
1.05 [0.31, 2.81]
1.99 [0.22, 7.84]
1.21 [0.25, 3.74]
1.32 [0.27, 3.86]
1.05 [0.35, 2.52]
1.05 [0.2, 3.33]
0.91 [0.38, 1.8]
1.63 [0.19, 6.41]
1.57 [0.2, 5.88]
1.49 [0.18, 5.5]
1.01 [0.12, 3.81]
1.15 [0.21, 3.66]
1.82 [0.19, 7.48]
2.35 [0.67, 5.81]
0.7 [0.12, 2.33]
1.33 [0.23, 4.41]
0.98 [0.32, 2.35]
1.5 [0.45, 3.75]
0.9 [0.11, 3.49]
0.95 [0.1, 3.77]
1.28 [0.47, 2.76]
2.48 [0.43, 8.43]
1.01 [0.39, 2.08]
1.24 [0.43, 2.8]
1.14 [0.11, 4.88]
1.69 [0.19, 6.36]
1.15 [0.14, 4.42]
0
4
8
12
Effect size


a)
Overcompensating
45
40
35
30
25
Body size
North
South
b)
Perfectly compensating
45
40
35
30
25
Body size
North
South
c)
Undercompensating
45
40
35
30
25
Body size
North
South
d)
Compensation Effect Size
5
0
-5
-10
-15
25
0
-25
-50
-75