VOL. 190, NO. 3 THE AMERICAN NATURALIST SEPTEMBER 2017

What Are the Environmental Determinants of Phenotypic Selection? A Meta-analysis of Experimental Studies

Christina M. Caruso,[1,*,†] Ryan A. Martin,[2,†] Nina Sletvold,[3] Michael B. Morrissey,[4] Michael J. Wade,[5] Kate E. Augustine,[6] Stephanie M. Carlson,[7] Andrew D. C. MacColl,[8] Adam M. Siepielski,[9] and Joel G. Kingsolver[6]

1. Department of Integrative Biology, University of Guelph, Guelph, Ontario N1H 2W3, Canada; 2. Department of Biology, Case Western Reserve University, Cleveland, Ohio 44106; 3. Department of Ecology and Genetics, Uppsala University, SE-752 36 Uppsala, Sweden; 4. School of Biology, University of St. Andrews, St. Andrews, Fife KY16 9TH, United Kingdom; 5. Department of Biology, Indiana University, Bloomington, Indiana 47405; 6. Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599; 7. Department of Environmental Science, Policy and Management, University of California, Berkeley, California 94720; 8. School of Life Sciences, University of Nottingham, Nottingham NG7 2RD, United Kingdom; 9. Department of Biological Sciences, University of Arkansas, Fayetteville, Arkansas 72701

Submitted May 24, 2016; Accepted April 7, 2017; Electronically published August 2, 2017

Online enhancements: appendix in four parts. *Dryad data:* http://dx.doi.org/10.5061/dryad.r19f8.

ABSTRACT: Although many selection estimates have been published, the environmental factors that cause selection to vary in space and time have rarely been identified. One way to identify these factors is by experimentally manipulating the environment and measuring selection in each treatment. We compiled and analyzed selection estimates from experimental studies. First, we tested whether the effect of manipulating the environment on selection gradients depends on taxon, trait type, or fitness component. We found that the effect of manipulating the environment was larger when selection was measured on life-history traits or via survival. Second, we tested two predictions about the environmental factors that cause variation in selection. We found support for the prediction that variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness but not for the prediction that variation is more likely to be caused by biotic factors. Third, we compared selection gradients from experimental and observational studies. We found that selection varied more among treatments in experimental studies than among spatial and temporal replicates in observational studies, suggesting that experimental studies can detect relationships between environmental factors and selection that would not be apparent in observational studies.

Keywords: causes of selection, directional selection, fitness, natural selection, selection gradient.

Introduction

There are more than 8,000 published estimates of phenotypic selection, measured as the relationship between quantitative traits and relative fitness (i.e., selection gradients and differentials; Lande and Arnold 1983). These estimates demonstrate that selection varies in space and time (Siepielski et al. 2009, 2011, 2013), presumably because the environmental factors that affect traits, fitness, or the relationship between traits and fitness also vary in space and time (e.g., Weis et al. 1992; Steele et al. 2011). However, the specific environmental factors that cause variation in selection are difficult to identify without a thorough understanding of how organisms interact with their biotic and abiotic environments (Endler 1986, pp. 164–165; Wade and Kalisz 1990). Consequently, the environmental factors that cause variation in selection have been identified in few study systems (MacColl 2011).

One commonly invoked prediction about the environmental factors that cause selection to vary in space and time is that variation in selection is more likely to be caused by biotic factors than by abiotic factors. This prediction arises from multiple sources, including Darwin (1859, chap. 3), who argued that population size is limited more by interactions with competitors and predators than by climate; Fisher (1930, pp. 41–42), who argued that the rate of adaptation is primarily limited by a species' interactions with its competitors and natural enemies because they are also evolving; Thompson (2005, pp. 3–4), who argued that most evolution is coevolution because all species depend on at least one other species to survive and reproduce; and Calsbeek et al. (2012), who argued that biotic factors are more important causes of variation in selection because they fluctuate more

---

[*] Corresponding author; e-mail: carusoc@uoguelph.ca.

[†] These authors contributed equally to this work.

ORCIDs: Caruso, http://orcid.org/0000-0001-7069-9572; Morrissey, http://orcid.org/0000-0001-6209-0177.

Am. Nat. 2017. Vol. 190, pp. 363–376. © 2017 by The University of Chicago. 0003-0147/2017/19003-57008$15.00. All rights reserved.
DOI: 10.1086/692760

rapidly and on a smaller spatial scale than abiotic factors. Consequently, although there are hypotheses that predict that abiotic factors should be important causes of evolutionary change (i.e., Court Jester hypotheses; reviewed in Barnosky 2001) and evidence that abiotic factors cause variation in natural selection at a global scale (Siepielski et al. 2017), the idea that biotic factors are important agents of selection is common in evolutionary biology. For example, selection by biotic factors is central to the theory of indirect genetic effects (Wolf et al. 1998) and the Red Queen hypothesis (Van Valen 1973).

A second commonly invoked prediction about the environmental factors that cause selection to vary in space and time is that variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness. This prediction is based on the opportunity for selection (I), which places an upper limit on the strength of selection (Crow 1958). Because the opportunity for selection is estimated as the variance in absolute fitness (V_W) divided by mean absolute fitness squared (W^2; Crow 1958; Shuster and Wade 2003, chap. 2), it will be larger in environments where mean fitness is relatively low than in environments where mean fitness is relatively high, even if the variance in absolute fitness does not differ between environments. Consequently, in environments where mean fitness is relatively low, selection should be stronger because the opportunity for selection is larger (Arnold and Wade 1984; Rundle and Vamosi 1996; Shuster and Wade 2003, chap. 2). The idea that environmental factors that have a large effect on mean fitness are important agents of selection is common in evolutionary biology. For example, pollinators are predicted to exert stronger selection on floral traits in populations where infrequent pollination causes mean fitness to be relatively low (e.g., Sletvold and Ågren 2014).

Both the prediction that variation in selection is more likely to be caused by biotic factors and the prediction that variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness have been supported in some study systems (e.g., Lau and Lennon 2011; Pelletier and Coulson 2012) but not in others (e.g., Caruso et al. 2005; Sletvold and Agren 2014). To test whether these two predictions are supported across study systems, we compiled a database of selection estimates from experimental studies. These experimental studies have not been included in previous meta-analyses of selection estimates (Kingsolver et al. 2001; Siepielski et al. 2009, 2011, 2013; Kingsolver and Diamond 2011), in part because experiments are often conducted in controlled environments (e.g., growth chambers, experimental ponds) and may produce selection estimates that are not representative of estimates from observational studies in unmanipulated populations (e.g., Irschick 2003; Zajitschek and Bonduriansky 2014). In the first type of experiment included in our database, one or more environmental factors was manipulated, and if selection differed between treatments, then the manipulated environmental factor was inferred to cause variation in selection (Wade and Kalisz 1990). In the second type of experiment, replicate populations were transplanted into two or more sites, and if selection differed between sites, then one or more of the environmental factors that differed between sites was inferred to cause variation in selection (e.g., Etterson 2004).

The two types of experimental studies included in our database can expose populations to a wider range of environments than observational studies, including environments that are infrequently found in the wild (Wootton and Pfister 1998). Consequently, experimental studies may be able to estimate relationships between environmental factors and selection estimates that would not be apparent in observational studies. If experiments expose populations to a wider range of environments, then there should be more variation in selection estimates among treatments within experimental studies than among spatial and temporal replicates within observational studies. If experiments displace populations from their adaptive peaks by exposing them to environments that are infrequently found in the wild (Estes and Arnold 2007), then selection estimates from experimental studies should be larger than estimates from observational studies.

Here, we analyze directional selection gradients from experimental studies using meta-analytic models (Morrissey and Hadfield 2012) to address three objectives. The first objective was to explore whether the effect of manipulating the environment on selection estimates depends on taxon, trait type, or fitness component. Unlike previous meta-analyses (e.g., Kingsolver et al. 2001), which explored whether directional selection gradients (β) vary among taxa, trait types, and fitness components, our analysis explored whether the magnitude of the difference in directional selection gradients between treatments (i.e., $|\beta_i - \beta_j|$, where i and j denote different treatments) varies among taxa, trait types, and fitness components. The second objective was to test two predictions about the environmental factors that cause selection to vary in space and time. To test the prediction that variation in selection is more likely to be caused by biotic factors, we compared the mean $|\beta_i - \beta_j|$ from studies that manipulated biotic factors and studies that manipulated abiotic factors. To test the prediction that variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness, we estimated the relationship between directional selection gradients (β) and mean fitness across treatments. The third objective was to determine how selection estimates vary depending on the context in which they are measured. Specifically, we tested whether the effect of manipulating the environment on selection ($|\beta_i - \beta_j|$) differs between experiments conducted in controlled versus field environments and whether directional selection gradients (β) differ between experimental and observational studies.

Methods

Search Strategy

We searched the literature for experimental studies of phenotypic selection published between 1990 (when Wade and Kalisz [1990] described how to infer the causes of selection by experimentally manipulating the environment) and 2013. First, we searched the Web of Knowledge (ver. 5.10; Thomson Reuters, Philadelphia) for papers that cited Wade and Kalisz (1990). Second, we searched the Web of Knowledge for papers that included the keywords "experiment*" or "manipulat*" and cited Lande and Arnold (1983), Mitchell-Olds and Shaw (1987), Schluter (1988), Schluter and Nychka (1994), or Brodie et al. (1995). Third, we searched the Web of Knowledge for papers that both cited Lande and Arnold (1983) and included the keyword "adaptive plasticity"; this search strategy identified experiments that manipulated the environment and measured selection within each treatment as described in Wade and Kalisz (1990) but with the goal of determining whether phenotypic plasticity is adaptive. Fourth, we searched the papers included in a review of selection on functional traits of plants (Geber and Griffen 2003). Most papers in the database were identified using one of the first three search strategies described above.

We included a paper in our database if it met four criteria. First, selection differentials or gradients were measured for quantitative traits. Differentials estimate both direct and indirect selection on a trait, whereas gradients estimate only direct selection (Lande and Arnold 1983). Second, selection differentials or gradients were standardized by the standard deviation of the trait (Lande and Arnold 1983); although selection estimates can also be standardized by the mean of the trait (Hereford et al. 2004), few studies do so. Third, selection differentials or gradients were measured in either ≥2 experimentally manipulated environments or ≥2 transplant sites. Fourth, selection was measured on phenotypic values rather than on genotypic or breeding values (i.e., genotypic selection; Rausher 1992).

Database Construction

For each paper that we included in our database, we recorded four types of data. First, we recorded the name of the study species and whether it was a vertebrate, invertebrate, or plant. Second, we recorded a description of the experimental manipulation and treatments. Third, we recorded the mean and standard error of traits and fitness components for each treatment. Fourth, we recorded directional and quadratic selection differentials and gradients for each treatment, including their associated standard errors, *P* values, and sample sizes. Each paper contributed a minimum of two records to the database, where a record included estimates of selection on a single trait, via a single fitness component, in a single treatment level. The database is deposited in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.r19f8 (Caruso et al. 2017).

For each paper, we coded experiments into one of three experimental context categories: laboratory, mesocosm, and field (table 1). The laboratory category included experiments conducted in indoor controlled environments such as greenhouses and growth chambers. The mesocosm category included experiments conducted in outdoor controlled environments such as experimental ponds, arrays, and gardens. Both the laboratory and mesocosm categories included some experiments conducted on phenotypically or genetically manipulated populations (e.g., hybrids, recombinant inbred lines). In contrast, the field category included experiments conducted on in situ natural populations in outdoor uncontrolled environments.

Each experiment was coded into one of four environmental factor categories: transplant, biotic, abiotic, and biotic+abiotic (table 1). The transplant category included experiments that transplanted populations into different field sites. The biotic category included experiments that manipulated one or more biotic factors such as predators, mutualists, and competitors. The abiotic category included experiments that manipulated one or more abiotic factors such as tempera-

Table 1: Number of estimates of directional selection gradients in the database for each category of five moderator variables

Moderator, variable	No. records reporting SE	Total no. records (inc. reports without SE)
Taxon:		
Invertebrate	73	89
Plant	604	1,262
Vertebrate	202	356
Trait type:		
Life history	151	272
Morphology	411	829
Physiology	77	143
Size	240	463
Fitness component:		
Fecundity	635	1,077
Mating success	120	152
Performance	50	269
Survival	58	187
Total fitness	16	22
Environmental factor:		
Abiotic	195	355
Biotic	461	824
Both biotic and abiotic	107	175
Transplant	116	353
Experimental context:		
Field	262	350
Laboratory	317	641
Mesocosm	300	716

ture, soil moisture, and $[CO_2]$. The biotic+abiotic category included experiments that manipulated both a biotic and an abiotic factor in the same experiment.

Within each experiment, we recorded two types of data for each record of selection. First, we recorded the name of the trait on which selection was measured and whether it was a life-history, morphology, physiology, or size trait (table 1). A few records ($N = 21$) estimated selection on traits such as herbivore resistance that could not be coded into any of our categories; these records were excluded from all analyses and from table 1. Second, we recorded the name of the fitness component used to measure selection and whether it was an estimate of fecundity, mating, performance, survival, or total fitness (table 1). These trait and fitness component categories have been used in previous meta-analyses of phenotypic selection (Kingsolver et al. 2001; Kingsolver and Diamond 2011) with two exceptions: the physiology category, which includes records of selection on traits such as photosynthetic rate and plasma testosterone concentration; and the performance (sensu Arnold 1983) category, which includes records of selection estimated via components such as aboveground biomass that mediate the relationship between traits and fitness.

Statistical Analyses

Although we compiled estimates of directional and quadratic selection differentials and gradients, we analyzed directional selection gradients. We analyzed directional rather than quadratic selection estimates because the magnitude of quadratic selection gradients and differentials are often underestimated by 50% (Stinchcombe et al. 2008). We analyzed directional selection gradients rather than differentials because there were more than twice as many records of directional selection gradients ($N = 879$) than of directional selection differentials ($N = 402$). However, we included a supplemental analysis of directional selection differentials in the appendix, part A (appendix pts. A–D available online).

Objective 1: Conduct Exploratory Meta-analysis

For our exploratory meta-analysis, we tested whether the effect of experimentally manipulating the environment on selection varied among taxa, trait types, and fitness components. To do this, we estimated the magnitude of the difference in directional selection gradients between treatments (i.e., $|\beta_i - \beta_j|$, where i and j represent different treatments) and compared the mean $|\beta_i - \beta_j|$ between categories of taxa, trait types, and fitness components. Our focus on $|\beta_i - \beta_j|$ differs from the focus of previous exploratory meta-analyses (e.g., Kingsolver et al. 2001), which compared directional selection gradients (β) between categories of taxa, trait types, and fitness components.

To estimate the magnitude of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) for different categories of taxa, trait types, and fitness components, we fit the following mixed model:

$$\hat{\beta}_{i,j,k} = \mu_j + b_k + m_i + e_{i,j}, \tag{1}$$

where $\hat{\beta}_{i,j,k}$ are estimated selection gradients, indexed by i. The moderator variable category (e.g., survival, when the moderator variable is fitness component; see table 1 for all moderator variable categories) is indexed by j. The study is indexed by k, where a study is defined as ≥2 records of selection gradients that are from the same experiment and share the same variables (e.g., trait type, fitness component) but were estimated in different treatments. Consequently, most experiments contained multiple studies. We include b_k in the model to distinguish between selection gradients from different studies. Here, m_i are sampling errors and $e_{i,j}$ are residuals; b_k are treated as random effects, distributed according to $b_k \sim N(0, \mathrm{SE}_i^2)$, where the variance of the b_k terms (σ_b^2) is estimated. We draw m_i from distributions with known variance according to $m_i \sim N(0, \mathrm{SE}_i^2)$, where SE_i are the reported standard errors of $\hat{\beta}_{i,j,k}$. The residuals are assumed to be drawn from separate normal distributions for each category, j, of the moderator variable $e_{i,j} \sim N(0, \sigma^2(e)_j)$, where each residual variance $\sigma^2(e)_j$ is estimated separately. Models were fit for each of the three moderator variables: taxon, trait type, and fitness component (table 1).

The model described by equation (1) estimates the variance of the distribution of selection gradients (β) for each moderator variable category. These variances were used to estimate the mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) for each moderator variable category. Specifically, because the mean absolute value of the difference between two independent random gradients drawn from the same normal distribution is $[2/(\pi)^{1/2}]\sigma(x)$ (Nair 1936), $|\beta_i - \beta_j|$ could be calculated as $[2/(\pi)^{1/2}][\sigma^2(\beta)_j]^{1/2}$, where $\sigma^2(\beta)_j$ is the estimated variance of selection gradients in any given moderator variable category.

The model described by equation (1) estimates the variance of the distribution of selection gradients (β) for each moderator variable category without controlling for correlations with other moderator variables and thus is a univariate model. However, this univariate model could be misleading because some moderator variables were correlated with each other. For example, many transplant studies estimated selection via survival or total fitness. To estimate the mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) for each moderator variable category while controlling for correlations among the five moderator variables in table 1, we fit a multivariate version of the model described by equation (1). This model took the form

$$\hat{\beta}_{i,j,k} = \mu_{j,k} + b_k + m_i + \sum_m e_{i,j,m}, \tag{2}$$

where notation is the same as for equation (1) except that effects are simultaneously included for each estimate, $\hat{\beta}_{i,j,k}$, for each associated category j of each moderator variable, indexed m, and the variance contributed by each category of each moderator variable is independently specified (i.e., such that $e_{i,j,m} \sim N(0, \sigma^2(e)_{m,j})$; see appendix, pt. B for additional information). The model described by equation (2) estimates the variance of the distribution of selection gradients (β) for each moderator variable category. These variances were used to estimate the mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) for each moderator variable category as described for equation (1), above. However, unlike $|\beta_i - \beta_j|$ estimated from the univariate models, $|\beta_i - \beta_j|$ estimated from the multivariate model was calculated relative to an arbitrary reference category of the moderator variable (e.g., morphology for the trait type variable; fig. 1*b*).

To compare the mean $|\beta_i - \beta_j|$ between moderator variable categories, we calculated Wald-type confidence intervals. These confidence intervals were calculated using the standard error of the mean absolute difference

$$\sqrt{\sigma^2[\sigma^2(\beta)]\left(\frac{1}{\sqrt{\pi}\sqrt{\sigma^2(\beta)}}\right)^2},$$

where $\sigma^2[\sigma^2(\beta)]$ is the sampling variance of $\sigma^2(\beta)$, the estimated variance of selection gradients in any given moderator variable category. If the effect of manipulating the environment on selection depends on taxon, trait type, or fitness component, then the confidence intervals for moderator variable categories (e.g., vertebrates and plants within the moderator variable taxon) should not overlap. When the confidence intervals for two moderator variable categories did not overlap in the univariate analysis but did overlap in the multivariate analysis, then we conservatively assumed that the effect of manipulating the environment on selection did not depend on taxon, trait type, or fitness component.

The models described by equations (1) and (2) and all other models described below were fitted using the MCMCglmm package (Hadfield 2010) in R (R Development Core Team 2013). Estimates of all parameters in these models were obtained using data cloning (see appendix, pt. C; Lele et al. 2007). Data cloning uses Bayesian Markov chain Monte Carlo methods to calculate maximum likelihood estimates of model parameters and their associated standard errors. These maximum likelihood estimates, unlike estimates from traditional Bayesian models, are not sensitive to the choice of priors (Lele et al. 2007).

Objective 2: Test Two Predictions about the Environmental Factors That Cause Variation in Selection

If variation in selection is more likely to be caused by biotic factors than by abiotic factors, then experimentally manipulating biotic factors should have a larger effect on selection than manipulating abiotic factors. To test whether experimentally manipulating biotic factors had a larger effect on selection, we fit the model described by equation (1) for the moderator variable environmental factor (table 1). To control for correlations between environmental factor and the other moderator variables (table 1), we also fit the multivariate model described by equation (2). Both models were used to estimate the variance of the distribution of selection gradients (β) for each environmental factor category. These variances were used to estimate the mean $|\beta_i - \beta_j|$ and associated confidence intervals for each environmental factor category as described in objective 1 above. If variation in selection is more likely to be caused by biotic factors, then the mean $|\beta_i - \beta_j|$ for studies that manipulated biotic factors should be larger than the mean $|\beta_i - \beta_j|$ for studies that manipulated abiotic factors, and their confidence intervals should not overlap.

If variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness, then selection should be stronger in treatments where mean fitness is relatively low. To test whether selection is stronger in treatments where mean fitness is relatively low, we analyzed the subset of records ($N = 493$) that reported directional selection gradients, their associated standard errors, and mean fitness for each treatment. These records were analyzed by fitting a random regression meta-analytic mixed model

$$\hat{\beta}_{i,j} = \mu + B \cdot \bar{W}_i^* + b_j + c_j \cdot \bar{W}_i^* + m_i + e_i, \tag{3}$$

where the notation is as for equations (1) and (2), except that μ and B are the overall (fixed) intercept and slopes for the regression of selection gradients on treatment-specific mean fitness $\bar{W}_i^*$, and b_j and c_j are random (among study, indexed j) slopes and intercepts of the same regression. The random slopes and intercepts are assumed to be drawn from the distribution $[{}^{b}_{c}]_j \sim N(\mathbf{O}, \sum)$, where $\sum$ is the variance-covariance matrix of random slopes and intercepts.

The covariance of random slopes and intercepts describes the dependence of the magnitude of selection on mean fitness because population mean fitness in each treatment was scaled by dividing by mean fitness across treatments within studies and subtracting 1. A negative covariance would indicate that selection gradients become less extreme as mean fitness increases (assuming that μ and B are small) and thus that selection is stronger in environments where mean fitness is relatively low.

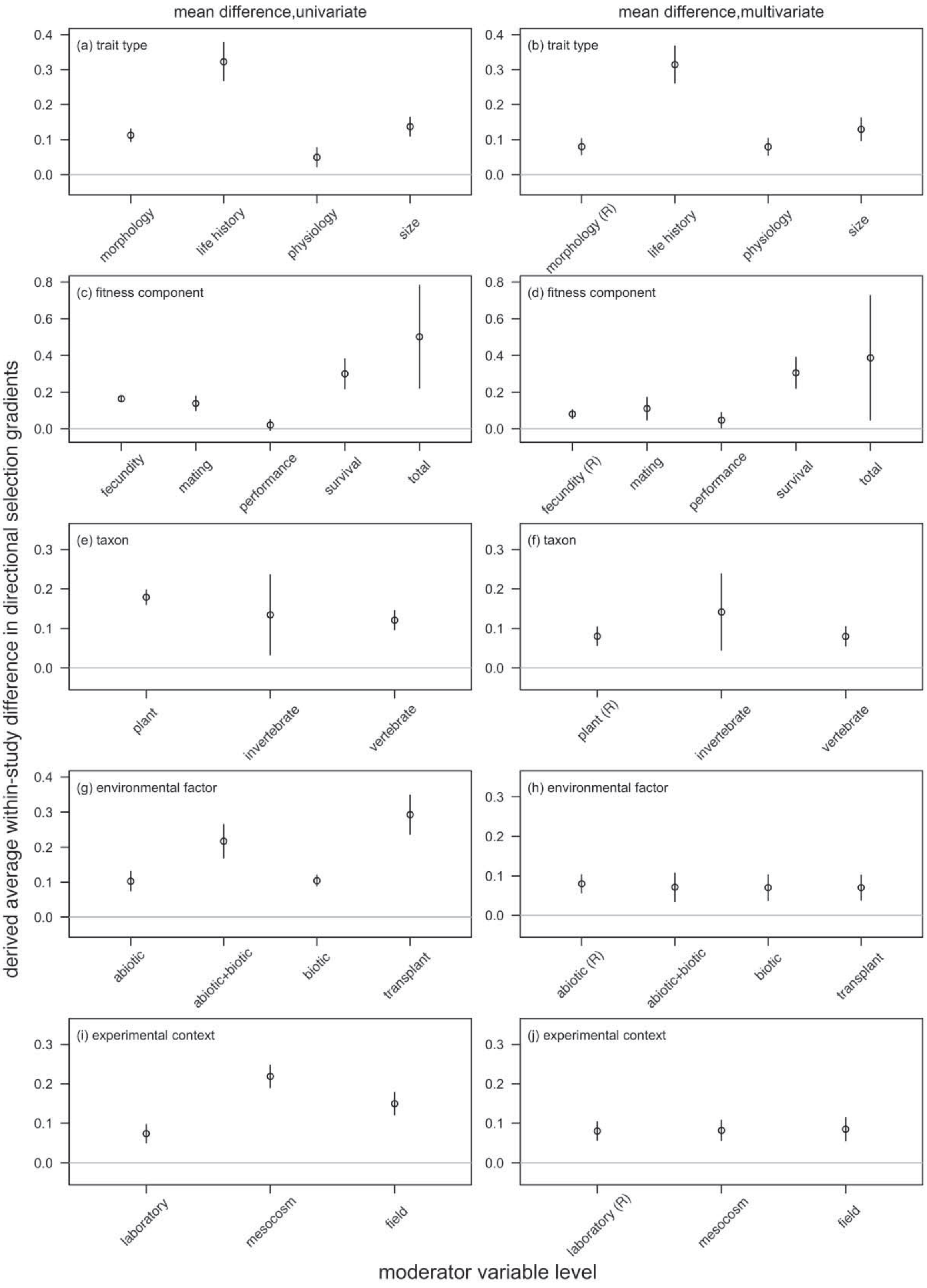

Figure 1: Mean (±95% confidence interval) absolute value of the difference in directional selection gradients between treatments (i.e., $|\beta_i - \beta_j|$, where *i* and *j* denote different treatments) for categories of five moderator variables: trait type (*a*, *b*); fitness component (*c*, *d*); taxon (*e*, *f*); environmental factor (*g*, *h*); and experimental context (*i*, *j*). The left column shows the results from univariate models that do not control for correlations with other moderator variables. The right column shows the results from a multivariate model that controls for correlations among moderator variables, where R denotes the reference category for each variable. Note that the scale of the *Y*-axis differs between moderator variables. All analyses included only the subset of records that reported directional selection gradients and their associated standard errors.

Objective 3: Test How Selection Estimates Vary Depending on the Context in Which They Are Measured

To test whether the effect of manipulating the environment on selection estimates differs between experimental studies conducted in controlled and field environments, we fit the model described by equation (1) for the moderator variable experimental context (table 1). To control for correlations between experimental context and the other moderator variables (table 1), we also fit the multivariate model described by equation (2). Both models were used to estimate the variance of the distribution of selection gradients (β) for each experimental context category. These variances were used to estimate the mean $|\beta_i - \beta_j|$ and associated confidence intervals for each experimental context category as described in objective 1 above. If the effect of manipulating the environment on selection differs between experimental studies conducted in controlled and field environments, then the confidence intervals for the moderator variable categories laboratory, mesocosm, and field should not overlap.

To test whether selection estimates differ between experimental and observational studies, we combined directional selection gradients from experimental studies ($N = 879$) with temporally and spatially replicated directional selection gradients from unmanipulated natural populations (Siepielski et al. 2009, 2013 databases; $N = 953$). The combined database included both spatially and temporally replicated selection estimates because spatial variation and temporal variation in selection are comparable (Morrissey and Hadfield 2012; Siepielski et al. 2013). Using this combined database, we fit a mixed model

$$\hat{\beta}_{i,j,k} = \mu_k + b_{j,k} + m_i + e_{i,k}, \tag{4}$$

where $\hat{\beta}_{i,j,k}$ are the estimated selection gradients, with i indexing individual estimates, j indexing studies (i.e., ≥2 records of selection gradients that share the same variables but were estimated in different treatments or replicates), and k indexing study type (experimental vs. observational). Thus, μ_k are intercepts, or mean selection gradients, for each study type. The random effects are defined according to $b_{j,k} \sim N(0, \sigma_k^2(b))$ and $e_{i,k} \sim N(0, \sigma_k^2(e))$. The model described by equation (4) was used to estimate (1) the variance (95% confidence interval [CI]) in directional selection gradients among treatments in experimental studies and among spatial and temporal replicates in observational studies and (2) the mean (95% CI) of directional selection gradients from experimental and observational studies. If experimental studies often expose populations to a wider range of environments than observational studies (Wootton and Pfister 1998), then the variance in β among treatments in experimental studies should be larger than the variance in β among spatial and temporal replicates in observational studies, and their confidence intervals should not overlap. If experimental studies often expose populations to environments that displace them from their adaptive peaks (Estes and Arnold 2007), then the mean β from experimental studies should be larger than the mean β from observational studies, and their confidence intervals should not overlap.

Effects of Subsetting the Data Set

All of our analyses included only the subset ($N = 879$ records from 51 published papers; table 1) of records that reported both directional selection gradients and their standard errors. Although there is no reason to believe that the subset of records that report standard errors is a biased sample of all records in our database, we also analyzed all of the records ($N = 1{,}707$ records from 91 published papers; table 1) that reported directional selection gradients using the models described in equations (1)–(4) modified to remove the sampling error term. However, the analyses of the full data set (see appendix, pt. D) should be interpreted with caution; above and beyond any effects of analyzing a subset of the data, not accounting for sampling error can inflate estimates of the mean and variance in selection, and either reduce or amplify any differences in selection between moderator variable categories, depending on whether standard errors vary among categories (Morrissey and Hadfield 2012; Morrissey 2016). Consequently, only the analyses of the subset of records that reported both directional selection gradients and their associated standard errors are described in the results below.

Results

Description of the Database

The database of experimental studies included 879 records that reported both directional selection gradients and their associated standard errors (table 1). These records of selection came from 51 published papers that contained a median of $N = 2$ treatments (range $= 2 - 7$). There were more records of selection from studies of plants (69%) than from studies of vertebrates (23%) or invertebrates (8%); more records of selection via fecundity (72%) than via mating (14%), performance (6%), survival (6%), or total fitness (2%); and more records of selection on morphology (47%) than on size (27%), life history (17%), or physiology (9%). Records from studies conducted in laboratory (36%), mesocosm (34%), and field (30%) environments were equally represented. There were fewer records of selection from studies that transplanted organisms into contrasting environments (13%) than from studies that manipulated the environment (87%). Among studies that manipulated the environment, there were more records of selection from studies that manipulated biotic factors (60%) than from studies that manipulated abiotic factors (26%) or both biotic and abiotic factors (14%).

Objective 1: Conduct Exploratory Meta-analysis

The mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) varied among trait types and fitness components (fig. 1*a*–1*d*). For trait type, $|\beta_i - \beta_j|$ was larger when selection was measured on life-history traits than on physiology, morphology, or size traits (fig. 1*a*, 1*b*). For fitness component, $|\beta_i - \beta_j|$ was larger when selection was measured through survival than through fecundity, mating, or performance (fig. 1*c*, 1*d*). These differences in $|\beta_i - \beta_j|$ between trait types and fitness components were detected in both univariate (fig. 1*a*, 1*c*) and multivariate (fig. 1*b*, 1*d*) models, suggesting that they are robust to any effects of correlations between moderator variables.

In contrast to trait type and fitness component, there was no effect of taxon type on the mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$). Although $|\beta_i - \beta_j|$ was larger for studies of plants than for studies of vertebrates in the univariate model (fig. 1*e*), $|\beta_i - \beta_j|$ did not vary with taxon type in the multivariate model that controlled for correlations with other moderator variables (fig. 1*f*).

Objective 2: Test Two Predictions about the Environmental Factors That Cause Variation in Selection

The mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) did not depend on the type of environmental factor that was manipulated. Instead, $|\beta_i - \beta_j|$ was similar for studies that manipulated biotic factors and studies that manipulated abiotic factors (fig. 1*g*, 1*h*). Relative to $|\beta_i - \beta_j|$ from studies that manipulated either biotic or abiotic factors, the mean absolute value of the difference in directional selection gradients between treatments was larger for studies that manipulated both biotic and abiotic factors and for studies that transplanted populations into contrasting environments (fig. 1*g*). However, these differences in $|\beta_i - \beta_j|$ between studies that manipulated different types of environmental factors were not found in the multivariate model that controlled for correlations with other moderator variables (fig. 1*h*).

Selection was stronger in treatments where mean fitness was relatively low. Across treatments, the covariance between directional selection gradients and mean fitness was negative (slope-intercept covariance [95% CI] $= -0.021$ [-0.023, -0.019]). This negative covariance indicates that within experimental studies, treatments with relatively low mean fitness had larger positive or negative directional selection gradients than treatments with relatively high mean fitness (fig. 2*a*). Consequently, the mean variance in directional selection gradients was larger in treatments with relatively low mean fitness than in treatments with relatively high mean fitness (fig. 2*b*).

Objective 3: Test How Selection Estimates Vary Depending on the Context in Which They Are Measured

The mean absolute value of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) did not differ between studies conducted in controlled and field environments. Although $|\beta_i - \beta_j|$ was larger for studies conducted in field or mesocosm environments than in laboratory environments in the univariate model (fig. 1*i*), this difference was not found in the multivariate model that controlled for other moderator variables (fig. 1*j*).

Directional selection gradients (β) from experimental and observational studies (fig. 3*a*) differed in two respects. First, the variance in β (95% CI) among treatments in experimental studies (0.0214 [0.0209, 0.0220]) was about three times greater than the variance in β among spatial and temporal replicates in observational studies (0.0069 [0.0067, 0.0071]). This difference between experimental and observational studies in part reflects the smaller standard errors associated with directional selection gradients from experimental studies (fig. 3*b*). Second, the mean β estimated in experimental studies (0.1115 [0.1083, 0.1149]) was about two times larger than the mean β estimated in observational studies (0.0591 [0.0549, 0.0637]).

Discussion

Our analysis of selection estimates from experimental studies was guided by three objectives. First, we conducted an exploratory meta-analysis and found that the effect of experimentally manipulating the environment on selection ($|\beta_i - \beta_j|$) depended on trait type and fitness component. Second, we tested two predictions about the environmental factors that cause selection to vary in space and time and found support for the prediction that variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness. Third, we tested how selection estimates depend on the context in which they are measured and found that there were differences between selection estimates from experimental and observational studies. These results are discussed below.

Objective 1: Conduct Exploratory Meta-analysis

The effect of experimentally manipulating the environment on selection depended on trait type; $|\beta_i - \beta_j|$ was larger when selection was measured on life-history traits than on other trait types (fig. 1*a*, 1*b*). A majority of these life-history traits were estimates of phenology, such as the timing of germination or flowering, in short-lived plants. In short-lived organisms, phenology traits are associated with stress avoidance because an individual that accelerates its phenology can develop rapidly and complete its life cycle prior to the onset of biotic or abiotic stress (e.g., Stanton et al. 2000). If the intensity

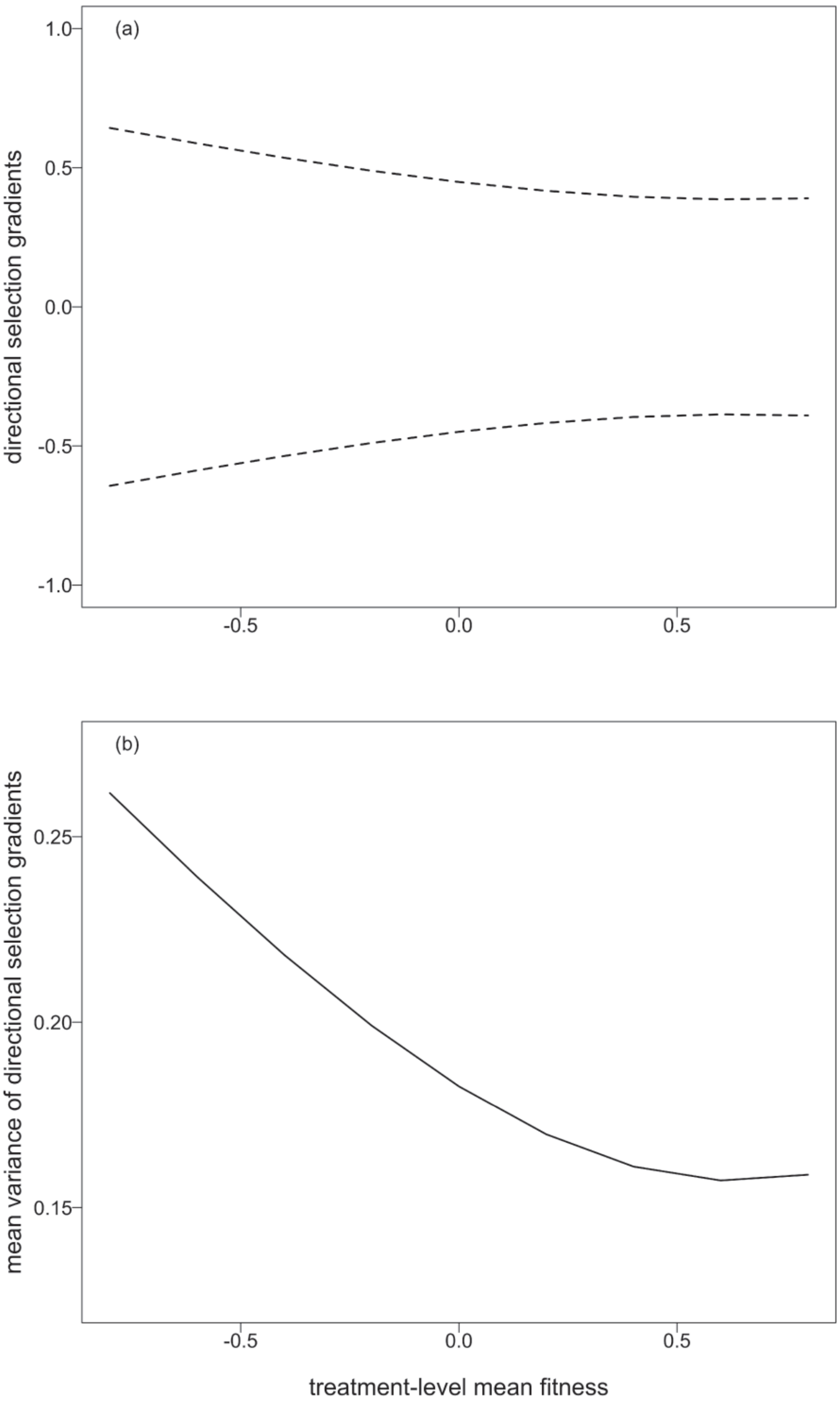

Figure 2: Relationship between directional selection gradients and treatment mean fitness, inferred from a random regression mixed model. Fitness is expressed relative to the average fitness across treatments and is centered on a mean of zero. *a*, The 95% confidence interval (dashed line) for the distribution of directional selection gradients across treatments that differ in mean fitness. A wider confidence interval indicates that selection gradients are larger; a narrower confidence interval indicates that selection gradients are smaller. *b*, The relationship between the mean variance in directional selection gradients and mean fitness across treatments. A negative slope indicates that directional selection is stronger in treatments with relatively low mean fitness than in treatments with relatively high mean fitness. All analyses included only the subset of records that reported directional selection gradients and their associated standard errors.

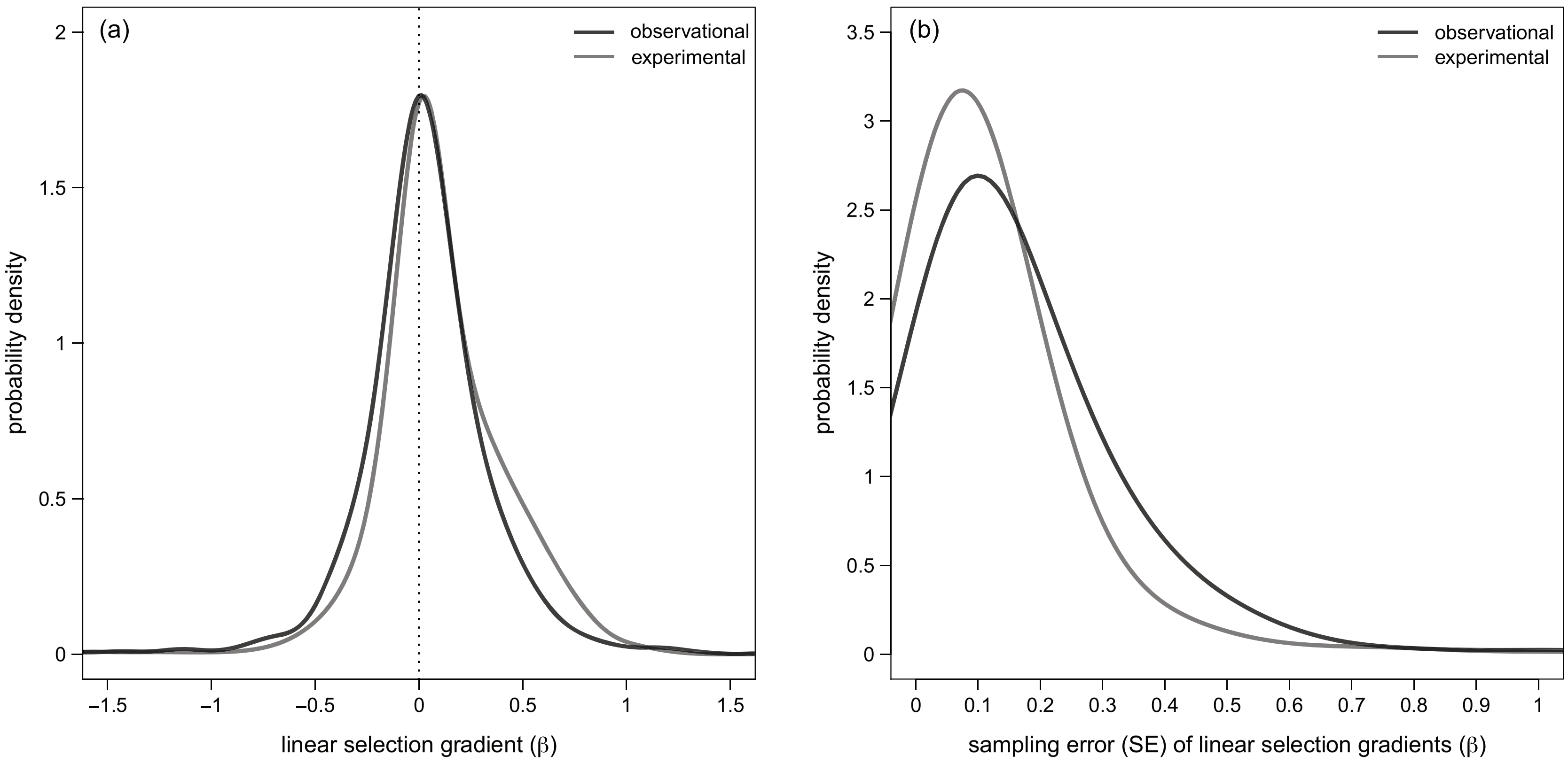

Figure 3: Distribution of estimated directional selection gradients (*a*) and their associated standard errors (*b*). Estimates from experimental studies that manipulated the environment are shown with the gray line. Estimates from observational studies in unmanipulated natural populations are shown with the black line.

of biotic or abiotic stress differs between treatments in experimental studies, then $|\beta_i - \beta_j|$ should be larger when selection is measured on phenology traits that are associated with stress avoidance.

The effect of experimentally manipulating the environment on selection also depended on fitness component; $|\beta_i - \beta_j|$ was larger when selection was estimated via survival than via other fitness components (fig. 1*c*, 1*d*). Selection estimated via survival may differ from selection estimated via other fitness components for two reasons. First, because survival is multiplicative, the most accurate estimates of selection via survival should come from studies that measure fitness over a short time period. In contrast, estimates of selection via other fitness components should not be sensitive to study duration (Hoekstra et al. 2001). If many experimental studies measure fitness over a short time period, they may more accurately estimate the effect of manipulating the environment on selection via survival than via other fitness components. Second, selection via survival, by reducing the phenotypic variance, can constrain subsequent episodes of selection via other fitness components (Wade and Kalisz 1989). This constraint could explain why $|\beta_i - \beta_j|$ was larger when selection was estimated via survival than via other fitness components.

Objective 2: Test Two Predictions about the Environmental Factors That Cause Variation in Selection

There was not support for the prediction that variation in selection is more likely to be caused by biotic factors than by abiotic factors; the magnitude of the difference in directional selection gradients between treatments ($|\beta_i - \beta_j|$) was similar for studies that manipulated biotic factors and studies that manipulated abiotic factors (fig. 1*g*, 1*h*). This comparison between studies that manipulated different types of environmental factors assumes that the effects of the biotic environment and the abiotic environment on selection are independent of each other. However, abiotic factors can indirectly affect selection by changing the biotic environment (Calsbeek et al. 2012). For example, selection for larger beak size in Darwin's finches was caused by extreme drought (an abiotic factor) that reduced seed production and thus indirectly increased the intensity of competition for food (a biotic factor; Boag and Grant 1981). Similarly, manipulating [CO_2] (an abiotic factor) indirectly decreased the importance of competition (a biotic factor) as an agent of selection on *Arabidopsis* (Lau et al. 2010). Even when the effects of the biotic and abiotic environment on selection are independent of each other, it may be difficult to determine their relative importance from short-term studies (such as the studies in our database) that manipulate the environment within a single generation. If biotic factors are more important causes of selection because Red Queen dynamics cause the biotic environment to change in a consistent direction across many generations (reviewed in Brockhurst et al. 2014), then long-term studies will be necessary to determine the relative importance of biotic and abiotic factors as causes of selection.

In contrast to the prediction that variation in selection is more likely to be caused by biotic factors, the prediction that variation in selection is more likely to be caused by environmental factors that have a large effect on mean fitness was supported; directional selection gradients were larger in treatments where environmental factors caused mean fitness to be relatively low than in treatments where environmental factors caused mean fitness to be relatively high (fig. 2). This relationship between the strength of selection and mean fitness is, in principle, inevitable because the opportunity for selection, which places an upper limit on the strength of selection, is larger in environments where mean fitness is relatively low (Rundle and Vamosi 1996). However, in practice, the expected relationship between the strength of selection and mean fitness may be obscured for three reasons. First, selection will always be weaker than the upper limit placed by the opportunity for selection (Crow 1958; Krakauer et al. 2011). Second, selection cannot act when mean fitness is zero because no individuals survive or reproduce. Third, environmental factors that do not have a large effect on mean fitness can still cause selection to vary in space and time by affecting the shape of the relationship between a trait and fitness (i.e., the fitness function) or by affecting the distribution of a trait (Weis et al. 1992; Steele et al. 2011). Yet we were still able to detect a relationship between the strength of selection and mean fitness, suggesting that the opportunity for selection is an important determinant of the strength of selection not just in principle but also in practice (Krakauer et al. 2011; Sletvold and Ågren 2016).

Objective 3: Test How Selection Estimates Vary Depending on the Context in Which They Are Measured

The effect of manipulating the environment on selection estimates was similar for experimental studies conducted in controlled and field environments; $|\beta_i - \beta_j|$ did not differ between studies conducted in the laboratory, in mesocosms, and in the field (fig. 1*j*). Most previous meta-analyses of selection estimates (i.e., Kingsolver et al. 2001; Siepielski et al. 2009, 2011, 2013; Kingsolver and Diamond 2011) focused on studies conducted in unmanipulated natural populations, and thus did not compare selection estimates between controlled and field environments. However, one meta-analysis that included experimental studies (Geber and Griffen 2003) found that selection on plant functional traits did not consistently differ between studies conducted in greenhouses, common gardens, and natural populations. These results suggest that selection estimates from studies conducted in controlled and field environments are similar, despite the difficulty of adequately simulating a complex field environment in the laboratory (Irschick 2003; Zajitschek and Bonduriansky 2014).

Although selection estimates did not differ between experimental studies conducted in controlled and field environments, they did differ between experimental and observational studies. Selection gradients varied more among treatments in experimental studies than among spatial and temporal replicates in observational studies, as expected if experimental studies expose populations to a wider range of environments than observational studies (Wootton and Pfister 1998). Selection gradients from experimental studies were also larger than gradients from observational studies, as expected if experiments displace populations from their adaptive peaks (Estes and Arnold 2007). Finally, selection gradients from experimental studies had smaller standard errors than gradients from observational studies (fig. 3*b*), indicating that experimental studies produce more precise selection estimates than observational studies and thus should have more power to detect the effect of an environmental factor on selection than observational studies. These results suggest that experimental studies should be the preferred approach to identifying the environmental factors that cause variation in selection not just because experiments can unambiguously determine causality (Wade and Kalisz 1990) but also because they can estimate relationships between environmental factors and selection estimates that would not be apparent in observational studies.

Limitations of the Database

Although we compiled >1,700 records of selection from the >90 published papers that have been published since Wade and Kalisz (1990), our database had four limitations. First, nearly half of the records (49%) in the database did not report both directional selection gradients and their associated standard errors. These records were excluded from our primary analyses, representing a loss of information about the environmental factors that cause variation in selection. Second, more than half of the records (71%) in the database did not report estimates of mean fitness in addition to selection gradients and their associated standard errors. This lack of fitness estimates limited our ability to test whether the relationship between the strength of selection and mean fitness varied depending on taxa, trait types, fitness components, or environmental factors. Third, the species used in the experimental studies included in our database largely did not overlap with the species used in observational studies included in other databases of selection estimates (Siepielski et al. 2009, 2011, 2013; Kingsolver and Diamond 2011). This lack of overlap is problematic because observational studies can generate hypotheses about the environmental factors that cause variation in selection that can then be tested with experimental studies (Wade and Kalisz 1990). Fourth, the database contained too few experimental studies to test more specific hypotheses about the environmental factors that cause variation in selection. For example, few experimental studies of phenotypic selection have manipulated interspecific competition, which limits our ability to test the hypothesis that competitor-mediated selection causes the evolution of biodiversity (reviewed in Pfennig and Pfennig 2012).

Conclusions

Our meta-analysis provides two guidelines for future studies of the causes of selection. First, future studies should not assume that biotic environmental factors are more important causes of variation in selection than abiotic factors. Although there are specific study systems where biotic factors are more important agents of selection (e.g., Lau and Lennon 2011), these results cannot be generalized across study systems. Second, future studies should focus on the effects of environmental factors on mean fitness (e.g., Benkman 2013; Sletvold et al. 2017). If environmental factors do not interact, then the factor that has the largest effect on mean fitness is predicted to be the most important cause of variation in selection. Although these guidelines are not a substitute for thoroughly understanding how organisms interact with their environment, they can facilitate the study of the causes of selection by making predictions about which environmental factors are most likely to cause selection to vary in space and time.

Acknowledgments

We thank the other members of the National Evolutionary Synthesis Center (NESCent, National Science Foundation [NSF], EF-0905606) working group on the Environmental Determinants of Natural Selection for their feedback, K. Eisen for assistance with data entry, and two anonymous reviewers for comments on the manuscript. We also thank the following sources of financial support: NESCent for funding a sabbatical fellowship to M.J.W.; NESCent for funding a sabbatical fellowship and two short-term fellowships to C.M.C.; the National Institute for Mathematical and Biological Synthesis (NIMBioS, NSF, DBI-1300426) for funding a postdoctoral fellowship to R.A.M.; and the NSF (DEB-1255318 and DEB-1620046) for support to A.M.S.

Literature Cited

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

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: theory. Evolution 38:709–719.

Barnosky, A. D. 2001. Distinguishing the effects of the Red Queen and the Court Jester on Miocene mammal evolution in the northern Rocky Mountains. Journal of Vertebrate Paleontology 21:172–185.

Benkman, C. W. 2013. Biotic interaction strength and the intensity of selection. Ecology Letters 16:1054–1060.

Boag, P. T., and P. R. Grant. 1981. Intense natural selection in a population of Darwin's finches (Geospizinae) in the Galapagos. Science 214:82–85.

Brockhurst, M. A., T. Chapman, K. C. King, J. E. Mank, S. Paterson, and G. D. Hurst. 2014. Running with the Red Queen: the role of biotic conflicts in evolution. Proceedings of the Royal Society B 281: 20141382.

Brodie, E. D., A. J. Moore, and F. J. Janzen. 1995. Visualizing and quantifying natural selection. Trends in Ecology and Evolution 10:313–318.

Calsbeek, R., T. P. Gosden, S. R. Kuchta, and E. I. Svensson. 2012. Fluctuating selection and dynamic adaptive landscapes. Pages 89–109 *in* E. I. Svensson and R. Calsbeek, eds. The adaptive landscape in evolutionary biology. Oxford University Press, New York.

Caruso, C. M., R. A. Martin, N. Sletvold, M. B. Morrissey, M. J. Wade, K. E. Augustine, S. M. Carlson, A. D. C. MacColl, A. M. Siepielski, and J. G. Kingsolver. 2017. Data from: What are the environmental determinants of phenotypic selection? a meta-analysis of experimental studies. American Naturalist, Dryad Digital Repository, http://dx.doi.org/10.5061/dryad.r19f8.

Caruso, C. M., D. L. Remington, and K. E. Ostergren. 2005. Variation in resource limitation of plant reproduction influences natural selection on floral traits of *Asclepias syriaca*. Oecologia 146:68–76.

Crow, J. F. 1958. Some possibilities for measuring selection intensities in man. Human Biology 30:1–13.

Darwin, C. 1859. On the origin of species by means of natural selection. J. Murray, London.

Endler, J. A. 1986. Natural selection in the wild. Princeton University Press, Princeton, NJ.

Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. American Naturalist 169:227–244.

Etterson, J. R. 2004. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the Great Plains. Evolution 58:1446–1458.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford University Press, Oxford.

Geber, M. A., and L. R. Griffen. 2003. Inheritance and natural selection on functional traits. International Journal of Plant Sciences 164:S21–S42.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software 33:1–22.

Hereford, J., T. F. Hansen, and D. Houle. 2004. Comparing strengths of directional selection: how strong is strong? Evolution 58:2133–2143.

Hoekstra, H. E., J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver. 2001. Strength and tempo of directional selection in the wild. Proceedings of the National Academy of Sciences of the USA 98:9157–9160.

Irschick, D. J. 2003. Measuring performance in nature: implications for studies of fitness within populations. Integrative and Comparative Biology 43:396–407.

Kingsolver, J. G., and S. E. Diamond. 2011. Phenotypic selection in natural populations: what limits directional selection? American Naturalist 177:346–357.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. American Naturalist 157:245–261.

Krakauer, A. H., M. S. Webster, E. H. Duval, A. G. Jones, and S. M. Shuster. 2011. The opportunity for sexual selection: not mismeasured, just misunderstood. Journal of Evolutionary Biology 24:2064–2071.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lau, J. A., and J. T. Lennon. 2011. Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. New Phytologist 192:215–224.

Lau, J. A., R. G. Shaw, P. B. Reich, and P. Tiffin. 2010. Species interactions in a changing environment: elevated CO_2 alters the ecological and potential evolutionary consequences of competition. Evolutionary Ecology Research 12:435–455.

Lele, S. R., B. Dennis, and F. Lutscher. 2007. Data cloning: easy maximum-likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. Ecology Letters 10:551–563.

MacColl, A. D. C. 2011. The ecological causes of evolution. Trends in Ecology and Evolution 26:514–522.

Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. Evolution 41:1149–1161.

Morrissey, M. B. 2016. Meta-analysis of magnitudes, differences and variation in evolutionary parameters. Journal of Evolutionary Biology 29:1882–1904.

Morrissey, M. B., and J. D. Hadfield. 2012. Directional selection in temporally replicated studies is remarkably consistent. Evolution 66: 435–442.

Nair, U. S. 1936. The standard error of Gini's mean difference. Biometrika 28:428–436.

Pelletier, F., and T. Coulson. 2012. A new metric to calculate the opportunity for selection on quantitative characters. Evolutionary Ecology Research 14:729–742.

Pfennig, D. W., and K. S. Pfennig. 2012. Evolution's wedge: competition and the origins of diversity. University of California Press, Berkeley.

Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. Evolution 46:616–626.

R Development Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rundle, H. D., and S. M. Vamosi. 1996. Selection may be strongest when resources are scarce: a comment on Wilson. Evolutionary Ecology 10:559–563.

Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait. Evolution 42:849–861.

Schluter, D., and D. Nychka. 1994. Exploring fitness surfaces. American Naturalist 143:597–616.

Shuster, S. M., and M. J. Wade. 2003. Mating systems and mating strategies. Princeton University Press, Princeton, NJ.

Siepielski, A. M., J. D. DiBattista, and S. M. Carlson. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. Ecology Letters 12:1261–1276.

Siepielski, A. M., J. D. DiBattista, J. Evans, and S. M. Carlson. 2011. Differences in the temporal dynamics of phenotypic selection among fitness components in the wild. Proceedings of the Royal Society B 278:1572–1580.

Siepielski, A. M., K. M. Gotanda, M. B. Morrissey, S. E. Diamond, J. D. DiBattista, and S. M. Carlson. 2013. The spatial patterns of directional phenotypic selection. Ecology Letters 16:1382–1392.

Siepielski, A., M. B. Morrissey, M. Buoro, S. M. Carlson, C. M. Caruso, S. M. Clegg, T. Coulson, et al. 2017. Precipitation drives global variation in natural selection. Science 355:959–962.

Sletvold, N., and J. Ågren. 2014. There is more to pollinator-mediated selection than pollen limitation. Evolution 68:1907–1918.

———. 2016. Experimental reduction in interaction intensity strongly affects biotic selection. Ecology 97:3091–3098.

Sletvold, N., M. Tye, and J. Ågren. 2017. Resource- and pollinator-mediated selection on floral traits. Functional Ecology 31:135–141.

Stanton, M. L., B. A. Roy, and D. A. Thiede. 2000. Evolution in stressful environments. I. Phenotypic variability, phenotypic selection, and response to selection in five distinct environmental stresses. Evolution 54:93–111.

Steele, D. B., A. M. Siepielski, and M. A. McPeek. 2011. Sexual selection and temporal phenotypic variation in a damselfly population. Journal of Evolutionary Biology 24:1517–1532.

Stinchcombe, J. R., A. F. Agrawal, P. A. Hohenlohe, S. J. Arnold, and M. W. Blows. 2008. Estimating nonlinear selection gradients using quadratic regression coefficients: double or nothing? Evolution 62: 2435–2440.

Thompson, J. N. 2005. The geographic mosaic of coevolution. University of Chicago Press, Chicago.

Van Valen, L. 1973. A new evolutionary law. Evolutionary Theory 1: 1–30.

Wade, M. J., and S. Kalisz. 1989. The additive partitioning of selection gradients. Evolution 43:1567–1569.

———. 1990. The causes of natural selection. Evolution 44:1947–1955.

Weis, A. E., W. G. Abrahamson, and M. C. Andersen. 1992. Variable selection on *Eurosta*'s gall size, I: the extent and nature of variation in phenotypic selection. Evolution 46:1674–1697.

Wolf, J. B., E. D. Brodie III, J. M. Cheverud, A. J. Moore, and M. J. Wade. 1998. Evolutionary consequences of indirect genetic effects. Trends in Ecology and Evolution 13:64–69.

Wootton, J. T., and C. A. Pfister. 1998. The motivation of and context for experiments in ecology. Pages 350–369 *in* W. J. Resetarits and J. Bernardo, eds. Experimental ecology: issues and perspectives. Oxford University Press, New York.

Zajitschek, F., and R. Bonduriansky. 2014. Quantitative genetics of wild populations of arthropods. Pages 147–159 *in* A. Charmantier, D. Garant, and L. E. B. Kruuk, eds. Quantitative genetics in the wild. Oxford University Press, New York.

Associate Editor: Luke J. Harmon
Editor: Yannis Michalakis

Some of the species included in the database of selection estimates compiled by the authors. *Top left*, ambush bug (*Phymata americana*); *bottom left*, great blue lobelia (*Lobelia siphilitica*); *right*, Lapland march orchid (*Dactylorhiza lapponica*). Photo credits: David Punzalan (*top left*), Brian Husband (*bottom left*), and Nina Sletvold (*right*).