



CHICAGO JOURNALS



Phenotypic Selection in Natural Populations: What Limits Directional Selection?

Author(s): Joel G. Kingsolver and Sarah E. Diamond

Source: *The American Naturalist*, Vol. 177, No. 3 (March 2011), pp. 346-357

Published by: [The University of Chicago Press](#) for [The American Society of Naturalists](#)

Stable URL: <http://www.jstor.org/stable/10.1086/658341>

Accessed: 16/02/2011 12:47

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press and The American Society of Naturalists are collaborating with JSTOR to digitize, preserve and extend access to *The American Naturalist*.

<http://www.jstor.org>

Phenotypic Selection in Natural Populations: What Limits Directional Selection?

Joel G. Kingsolver* and Sarah E. Diamond†

Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599

Submitted June 18, 2010; Accepted November 24, 2010; Electronically published January 31, 2011

Dryad data: <http://datadryad.org/handle/10255/dryad.7996>.

ABSTRACT: Studies of phenotypic selection document directional selection in many natural populations. What factors reduce total directional selection and the cumulative evolutionary responses to selection? We combine two data sets for phenotypic selection, representing more than 4,600 distinct estimates of selection from 143 studies, to evaluate the potential roles of fitness trade-offs, indirect (correlated) selection, temporally varying selection, and stabilizing selection for reducing net directional selection and cumulative responses to selection. We detected little evidence that trade-offs among different fitness components reduced total directional selection in most study systems. Comparisons of selection gradients and selection differentials suggest that correlated selection frequently reduced total selection on size but not on other types of traits. The direction of selection on a trait often changes over time in many temporally replicated studies, but these fluctuations have limited impact in reducing cumulative directional selection in most study systems. Analyses of quadratic selection gradients indicated stabilizing selection on body size in at least some studies but provided little evidence that stabilizing selection is more common than disruptive selection for most traits or study systems. Our analyses provide little evidence that fitness trade-offs, correlated selection, or stabilizing selection strongly constrains the directional selection reported for most quantitative traits.

Keywords: adaptation, fitness trade-offs, microevolution, natural selection, phenotypic selection.

Introduction

Natural and sexual selection are the primary mechanisms of adaptive evolution within populations (Darwin 1859). Over the past half century, phenotypic selection on quantitative traits has been detected in scores of field studies of plants and animals. Reviews of this burgeoning literature have documented several major patterns. First, there

is abundant evidence for directional selection on morphology and life history in many study systems (Endler 1986; Kingsolver et al. 2001; Hereford et al. 2004). Second, temporal fluctuations in the magnitude and sometimes the direction of directional selection are common (Siepielski et al. 2009). Third, there is little evidence for significant stabilizing selection in most study systems. In particular, current estimates of linear and quadratic selection suggest that stabilizing selection is no more common than disruptive selection (Kingsolver et al. 2001). The magnitude of directional selection is sufficient to produce rapid microevolutionary change in many populations (Grant and Grant 1989; Hendry and Kinnison 1999; Reznick and Ghalambor 2001).

These results present an important and unresolved challenge: if most populations are well adapted to current environmental conditions, then why is directional selection on quantitative traits apparently so common and stabilizing selection apparently uncommon (Estes and Arnold 2007)? What factors reduce total selection on phenotypic traits or reduce the cumulative evolutionary responses to such selection? There are several possible answers to this challenge. First, lack of genetic variation might prevent evolutionary responses to selection. Quantitative genetic studies document significant heritable variation for phenotypic traits in most natural populations (Mousseau and Roff 1987; Roff 1997), but there is low heritability for some traits and study systems (Kellermann et al. 2009). In addition, multivariate genetic constraints may reduce evolutionary responses to selection, even with abundant genetic variation of individual component traits (Walsh and Blows 2009). Second, there may be trade-offs among different components of fitness, such that, for example, a trait value increasing survival may also decrease mating success or fecundity. As a result, total selection on the trait may be less than directional selection via each fitness component (Roff 2002). Third, phenotypic and genetic correlations between traits may cause indirect, correlated selection (Lande and Arnold 1983). As a result, direct

* Corresponding author; e-mail: jgking@bio.unc.edu.

† Present address: Department of Biology, North Carolina State University, Raleigh, North Carolina 27695.

Am. Nat. 2011. Vol. 177, pp. 346–357. © 2011 by The University of Chicago. 0003-0147/2011/17703-5225\$15.00. All rights reserved.

DOI: 10.1086/658341

selection on a trait may be balanced by opposing indirect selection on a correlated trait. Fourth, directional selection on a trait may alternate in direction in time or space, reducing the cumulative effects of selection (Siepielski et al. 2009). Finally, stabilizing selection may indeed be stronger or more common than previously reported, such that natural populations may in fact be near local fitness peaks (Estes and Arnold 2007).

As reported in a recent review and synthesis by Siepielski et al. (2009), our estimates of phenotypic selection in natural populations have increased more than threefold in the past decade. Here we combine these estimates with an earlier data set (Kingsolver et al. 2001) to explore patterns and constraints on directional selection for quantitative traits in natural populations. We develop predictions for how fitness trade-offs, indirect selection, alternating selection, and stabilizing selection may reduce total or cumulative selection or responses to selection, and we evaluate these predictions by using our combined data sets. We also consider the limitations of our current data, studies, and statistical methods for detecting nonlinear components of selection. Our analyses provide little evidence that fitness trade-offs, correlated selection, or stabilizing selection strongly constrains directional selection on most quantitative traits.

Material and Methods

The Data Set

The current data set was generated by merging two previously published selection data sets from Kingsolver et al. (2001) and Siepielski et al. (2009); redundant records shared across both data sets were eliminated. We used the redundant records to confirm the consistency of coding and estimates between the two data sets (127 estimates from six studies). The data sets consider selection studies of natural variation in quantitative phenotypes within populations under natural field conditions. Details of the inclusion criteria for particular studies and contents of the finalized selection data sets can be found elsewhere (Kingsolver et al. 2001; Siepielski et al. 2009); briefly, the combined data set contains information on taxonomic groups, the number of temporal and spatial replicates, the traits measured, the fitness components measured, and selection estimates, including linear and quadratic gradients and differentials with their associated standard errors, P values, and sample sizes. Note that the combined data set does not include all relevant selection estimates during the time period considered (1984–2009)—for example, nonreplicated selection studies published between 2002 and 2009 are excluded—but this should not bias our analyses or results.

To facilitate comparisons among different traits and study systems, we used variance-standardized differentials or gradients (Lande and Arnold 1983), which quantify selection on traits in terms of the relationship between relative fitness and variation in a quantitative trait measured in units of standard deviation. Although the potential advantages of mean-standardized coefficients have been considered (Hereford et al. 2004), we use variance-standardized coefficients for two main reasons. First, standardization by the mean is appropriate only for traits that have a natural value rather than an arbitrary zero value (Hereford et al. 2004); this would exclude important types of traits such as phenological timing. Second, most of the studies included in the data set do not report mean-standardized coefficients and do not contain the necessary data to convert variance-standardized coefficients to mean-standardized coefficients.

Defining Traits and Fitness Components

Traits were assigned to major trait classes including size, other morphology, phenology, other life history, behavior, principal component (PC), interaction (e.g., behavior \times behavior or life history \times behavior; nonlinear selection gradients and differentials only), and other (a limited subset that could not readily be placed into the other classes). The trait class size included traits that indicated overall aspects of body size (e.g., body size, body length, body mass, and PCs for body size); all remaining morphology traits were classified as other morphology. The trait class phenology included traits that indicated the timing of life-history events (e.g., initial date, peak date, or duration of events such as germination, flowering, laying date, and hatching date); all remaining life-history traits were classified as other life history. PCs were placed into their own category, except for PCs reflecting aspects of overall body size (as defined by authors of the original study), which were included in the size category.

Most studies reported selection via individual components of fitness, such as survival, mating success, and fecundity. A few studies reported more comprehensive metrics of fitness, which were considered the total fitness class. These studies used measures such as net reproductive rate (R_0), intrinsic rate of increase (r), or similar metrics that combine survival and successful reproduction.

Analyses

We did not perform a formal meta-analysis because of (1) potential autocorrelation among estimates of selection based on multiple traits and/or fitness metrics from the same study, (2) limited availability of the phenotypic variance-covariance matrix needed for meta-analyses of

selection gradients, and (3) limited availability of standard errors needed for meta-analyses of selection differentials (Kingsolver et al. 2001; Hereford et al. 2004; Siepielski et al. 2009). In addition, our main goal here is not to estimate the average magnitude of selection but rather to evaluate the potential roles of fitness trade-offs, correlated selection, and other factors that may constrain directional selection and microevolution. As a result, we use graphical visualization techniques supplemented with traditional parametric and nonparametric statistical analyses and bootstrap resampling.

The vast majority of studies did not report standard errors of estimates, making it difficult to account for the effects of sampling error in these analyses. For example, of the 4,651 unique estimates of selection, 62% did not report standard errors for any estimates. Even in cases where studies reported standard errors, they often did not report these for all estimates (of the 2,538 estimates of the linear selection differential, s , 81% do not report standard errors). Sampling error may decrease the power to reject null hypotheses in some cases but should not otherwise bias our analyses or results. All graphical and statistical analyses were performed using R (ver. 2.9.1). To explore patterns of directional selection on different traits and fitness components, we graphically examined probability density distributions of linear selection gradients (β) for different trait classes and fitness components; densities were estimated using an automatic bandwidth (Silverman 1986). Sign tests were used to determine whether the distributions of β were centered at 0 or shifted toward either positive or negative values. Kruskal-Wallis rank-sum tests were used to determine whether the magnitude of β ($|\beta|$) was significantly different among trait classes or fitness components. We examined potential autocorrelation among estimates from the same replicate (a single study or a single temporal or spatial replicate within a study, where appropriate) via bootstrap resampling (500 iterations, with replacement) in which one estimate of β was randomly drawn from each replicate. We then performed visual and statistical tests on the resampled estimates to assess concordance with the analyses of the original data. Comparable resampling procedures were conducted for analyses of fitness trade-offs, the potential for indirect selection to reduce total selection, temporally alternating selection, and quadratic selection gradients (γ).

To explore the potential for fitness trade-offs to constrain directional selection, we considered the subset of linear selection gradient estimates for which two (rarely three) fitness components were measured on the same trait within the same study, species, and temporal/spatial replicate. With these data, we constructed and plotted three independent contrasts of the fitness components (estimates of total fitness were eliminated in this analysis):

fecundity versus survival, mating success versus survival, and fecundity versus mating success. If fitness components tend to be in the same direction (same sign), most estimates should fall into the lower-left and upper-right quadrants (quadrants 1 and 3) of a standard Cartesian plane; conversely, if components tend to be in opposite directions, estimates should fall into the upper-left and lower-right quadrants (quadrants 2 and 4). To test whether fitness components tend to be the same or opposite directions, we performed a χ^2 test to compare the number of estimates in quadrants 2 plus 4 and the number of estimates in quadrants 1 plus 3, relative to a null hypothesis of 50 : 50.

We assessed the potential for indirect selection to reduce total selection on a trait by examining the relationship between β , which estimates the strength of selection directly on the trait of interest, adjusted for effects of potential selection on phenotypically correlated traits, and s , which estimates the total strength of selection due to both direct and indirect (correlated) selection. (This interpretation of β assumes that all relevant correlated traits have been identified and measured in the study; see "Discussion.") We subsetted the data set to include only linear selection gradients and differentials measured on the same trait within the same study, species, and temporal/spatial replicate (excluding all PCs) for which two or more traits were considered. We explored indirect selection by using regression analysis. The slope of the regression of β on s (specifically, slopes < 1) indicated the potential for indirect selection to reduce total selection on a trait. We used reduced major axis (RMA) regression (also standardized major axis; sensu Legendre and Legendre 1998; Warton et al. 2006) to estimate the slope. In this case, RMA was more appropriate than ordinary least squares (OLS) regression for two reasons (Legendre and Legendre 1998; Warton et al. 2006): (1) our primary objective was to estimate the slope best describing the bivariate scatter of s and β rather than a slope that predicts s from β and (2) our subsequent objective was to test whether this slope deviated from the null hypothesis of slope = 1. RMA regression analyses were based on residual s and β (i.e., subtracting the sample mean from each estimate); this constrained the intercepts to pass through the origin (i.e., fixed at 0). We tested whether the slope of s as a function of β was significantly different from 1 (where slope = 1 is the null hypothesis that indirect selection does not reduce total selection on a trait), using the methods outlined by Warton et al. (2006).

To assess the potential for temporally alternating selection to constrain directional selection, we examined the number of sign changes in β for the same study, species, trait, and spatial replicate. Our method differed from that reported by Siepielski et al. (2009), as we counted the

number of differences in the sign of β in temporal sequence rather than examining the number of estimates of β with opposing signs. Thus, we estimated the probability that the direction of selection reverses from one temporal replicate to the next rather than the proportion of temporal replicates for which directional selection is positive. However, the key question was whether these temporal fluctuations were sufficient to reduce the cumulative effects of directional selection over time (Bell 2010). To address this, we examined the average of s over time for each trait for the same study, species, and spatial replicate. If directional selection on a trait was consistently in the same direction over time, $\text{mean}(|s|) = |\text{mean}(s)|$. Alternatively, if the direction of selection varied randomly over time, then $|\text{mean}(s)| \sim 0$. As a result, the slope of $\text{mean}(|s|)$ versus $|\text{mean}(s)|$ reflected the extent to which changes in the magnitude and direction of selection reduced the cumulative selection differential (Falconer and MacKay 1996).

Finally, we used methods comparable to those for linear selection gradients to examine patterns of nonlinear selection (γ) among different traits and fitness components. The quadratic selection gradient γ reflects the curvature of the fitness surface near the population mean phenotype. Negative values of γ are necessary (but not sufficient) for stabilizing selection, whereas positive values of γ are necessary (but not sufficient) for disruptive selection. If mean trait values in most populations are near local fitness peaks, we would predict estimates of γ to be strongly shifted toward negative values (Estes and Arnold 2007).

Results

Summary of Studies and Estimates

The combined data set included 143 studies with 2,819 estimates of linear selection gradients (β), 2,538 estimates of linear selection differentials (s), 1,250 estimates of quadratic selection gradients (γ), and 740 estimates of quadratic selection differentials (g). Collectively, these represent 4,651 distinct estimates of selection. Studies involving terrestrial plants (mainly angiosperms), invertebrates (mainly insects), and vertebrates (mainly birds and lizards) were well represented (table 1). Most estimates involved major components of fitness (survival, fecundity, and mating success), but ~2% of estimates used more integrated measures of total fitness (table 1). Aspects of size, phenology, and other morphological and life-history traits are well represented in the data set; only 1% of the estimates involved behavioral or physiological traits.

More than 98% of estimates of β and γ ranged between -2 and 2 and were roughly symmetric around 0 (fig. 1). We use $\text{median}(|\beta|)$ as a measure of the magnitude of

directional selection and $\text{median}(\beta)$ as an indicator of the tendency or shift toward positive or negative values (see “Material and Methods”). Probability densities of β for different trait types (fig. 2B) showed that directional selection on size was significantly shifted toward positive values ($\text{median}(\beta) = 0.09$, $70\% > 0$, $P < .0001$, $N = 363$) and that selection on phenology was significantly skewed toward negative values ($\text{median}(\beta) = -0.08$, $74\% < 0$, $P < .0001$, $N = 633$). These qualitative patterns held for selection via different fitness components, for different taxa, and for resampled estimates (accounting for potential autocorrelation among estimates from the same study or replicate). These results support previous studies suggesting that there is positive directional selection on aspects of size and negative selection on phenological timing in many study systems (Kingsolver and Pfennig 2004; Kingsolver 2009).

As in previous analyses, the magnitude of directional selection ($\text{median}(|\beta|)$) via survival ($\text{median}(|\beta|) = 0.08$) was significantly smaller than that for fecundity ($\text{median}(|\beta|) = 0.19$) or that for mating success ($\text{median}(|\beta|) = 0.17$; fig. 2A; Hoekstra et al. 2001; Kingsolver et al. 2001; Siepielski et al. 2011). This qualitative pattern held for selection on different traits, taxa, and resampled estimates. Importantly, the magnitude of selection via total fitness ($\text{median}(|\beta|) = 0.26$) was significantly stronger than selection via survival ($\text{median}(|\beta|) = 0.08$; Mann-Whitney test: $W = 54,343.5$, $P < .0001$), fecundity ($\text{median}(|\beta|) = 0.19$; $W = 24,113$, $P = .062$ [marginally significant]), or mating success ($\text{median}(|\beta|) = 0.17$; $W = 27,978$, $P = .026$) alone. While selection via total fitness represented only 2% of the estimates, this result does not support the hypothesis of trade-offs among fitness components.

Constraints on Directional Selection

To identify possible fitness trade-offs, we examined the subset of studies in which directional selection on a trait was estimated for multiple (usually two) fitness components. We focused on whether and how the direction of selection on a trait was reversed for different components, reflecting trade-offs among these different components (see “Material and Methods”). In 57% of cases, the direction of selection on a trait was the same among different fitness components (fig. 3). The number of cases in which the direction of selection on a trait was the same for pairs of fitness components was significantly greater than the number of cases in which the direction of selection was different ($\chi^2 = 4.84$, $df = 1$, $P = .028$, $N = 304$), increasing the total magnitude of selection on the trait. In the 43% of cases where the direction of selection differed, the magnitude of these differences was often relatively

Table 1: Summary statistics for estimates of linear and quadratic selection as functions of taxon, trait class, and fitness component

Taxonomic group	Median(β)	Median(β)	N	Trait class	Median(β)	Median(β)	N	Fitness component	Median(β)	Median(β)	N
Estimates of linear selection gradients (β):											
Invertebrates	.01	.16	547	Size	.09	.14	363	Mating success	.02	.17	796
Plants	.04	.14	656	Other morphology	.03	.14	1,395	Survival	<.01	.08	1,250
Vertebrates	<.01	.12	1,616	Phenology	-.08	.11	633	Fecundity	.04	.19	702
				Other life history	.03	.1	370	Total fitness	-.04	.26	57
				Principal component	-.01	.06	26	Net reproductive rate	.29	.29	3
				Behavior	.18	.26	29	Other	.02	.08	11
				Other	-.14	.14	3				
Estimates of quadratic selection gradients (γ):											
Invertebrates	-.03	.13	251	Size	-.04	.1	186	Mating success	-.02	.16	256
Plants	<.01	.05	326	Other morphology	<.01	.1	520	Survival	-.01	.06	593
Vertebrates	-.01	.1	673	Phenology	<.01	.07	330	Fecundity	<.01	.11	362
				Other life history	-.02	.11	68	Total fitness	.01	.12	18
				Principal component	.01	.09	25	Net reproductive rate	NA	NA	NA
				Behavior	-.09	.09	6	Other	<.01	.02	21
				Interaction	<.01	.13	109				
				Other	-.15	.15	6				

Note: The number of estimates, median, and median absolute value of the estimates are provided. NA = not available.

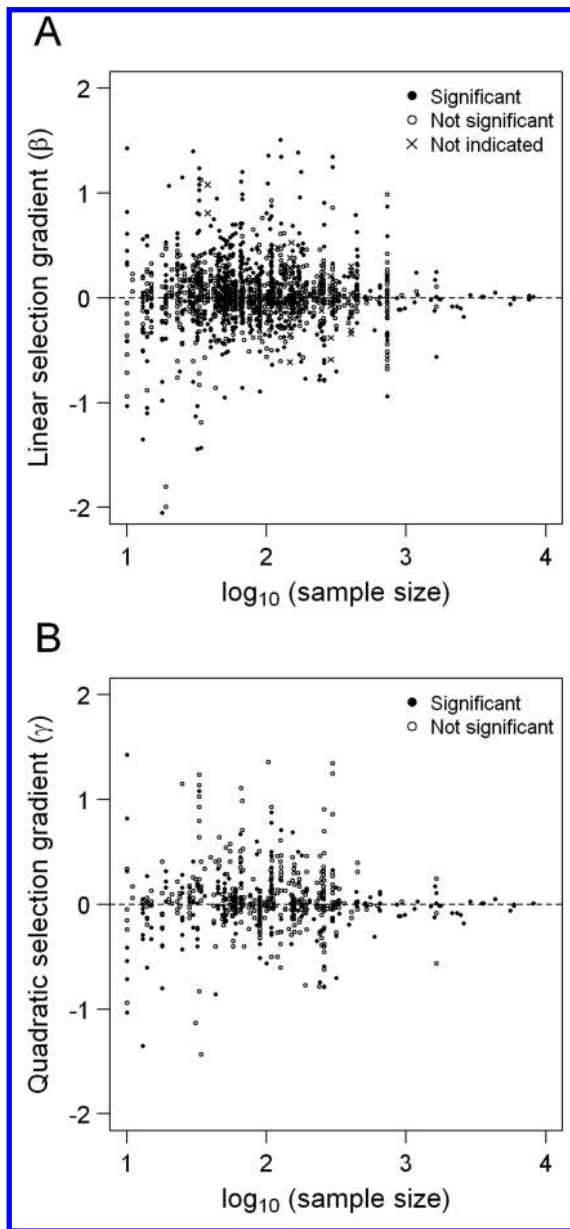


Figure 1: Summary of linear and quadratic selection gradients. Selection gradient estimates as a function of sample size (\log_{10} scale). A, Linear selection gradients (β); B, quadratic selection gradients (γ). The statistical significance (at the $P = .05$ level) of each estimate is given. Filled circles are significantly different from 0; open circles are not significant; crosses indicate that the significance of the estimate is not available.

small (fig. 3). In only 2% of cases did the magnitude and direction of selection on a trait for different fitness components balance each other to result in a net directional selection near 0 (± 0.01). In general, these qualitative patterns held for different traits, taxa, fitness component pair-

ings, and resampled pairs of estimates. These results are inconsistent with the notion that trade-offs among different fitness components generate opposing selection and reduce total fitness on most traits in natural populations (Roff 2002).

The linear selection gradient β estimates the strength of selection directly on the trait of interest, adjusted for effects of potential selection on phenotypically correlated traits, whereas the linear selection differential s estimates the total strength of selection due to both direct and indirect (correlated) selection. When both parameters are estimated for studies with multiple traits, the slope of the relationship between s and β (specifically, slopes < 1) indicates the potential for indirect selection to reduce total selection on a trait (fig. 4). RMA regression for all paired estimates (other than size traits; $N = 806$) yielded a slope of $b = 1.1$; on average, direct selection and total selection were similar, suggesting little impact of indirect selection in reducing total selection. In contrast, size showed a slope significantly < 1 ($b = 0.30$, $P < .0001$, $N = 104$). This qualitative pattern for size held for different fitness components and for resampled estimates. This suggests that indirect selection on correlated traits may frequently reduce total selection on size but not that on other types of traits.

Reversals in the direction of selection over time could also reduce the cumulative directional selection on a trait within or across generations. Using studies in which selection measurements were replicated over time (Siepielski et al. 2009) allowed us to estimate the frequency of changes in the direction (sign) of selection (fig. 5A). The direction of selection β changed in the next time interval, on average, 25% of the time; for some traits and studies, the frequency of sign changes exceeded 50% (fig. 5A). Autocorrelation of results from the same study had little effect on this result; resampling the frequency of sign change with one estimate from each study yielded a mean frequency of sign change of 26.6 (95% confidence interval [CI] = 26.3, 27.1). The frequency of sign changes decreased significantly as the magnitude of selection (median($|\beta|$)) increased ($r = -0.26$, $P < .0001$), suggesting that most of the reported sign changes occurred at low values of $|\beta|$ and may have resulted from sampling error. In addition, the RMA slope (constrained to go through the origin; see “Material and Methods”) of $\text{mean}(|s|)$ versus $|\text{mean}(s)|$ (see “Material and Methods”) is 0.95 (95% CI = 0.93, 0.98), suggesting that temporal fluctuations in directional selection do not strongly reduce cumulative selection differentials in most cases (fig. 5B). In only 21% of studies do temporal fluctuations in selection reduce cumulative selection differentials by more than 50%, that is, where $|\text{mean}(s)| < 0.5 \times \text{mean}(|s|)$. As a result, temporal changes in the direction of selection occur regularly in many study systems (Siepielski et al. 2009), but our results

do not show that such fluctuations strongly reduce the cumulative effects of directional selection over time in most systems.

Nonlinear Selection

The quadratic selection gradient γ reflects the curvature of the fitness surface near the population mean phenotype (see “Material and Methods”). Negative values of γ are necessary (but not sufficient) for stabilizing selection, whereas positive values of γ are necessary (but not sufficient) for disruptive selection. Probability densities for γ showed that estimates were distributed symmetrically around 0 for different fitness components and for total fitness (figs. 1, 6A), and median(γ) values were near 0 in each case (median(γ): survival = -0.01 ; fecundity < 0.01 ; mating success = -0.02 ; total fitness = 0.01). This result is inconsistent with the notion that stabilizing selection is much more common than disruptive selection, as would be expected if most populations are near fitness peaks (Lande 1979; Lande and Arnold 1983). Probability densities of γ for different traits (fig. 6B) showed that estimates of γ for size were significantly shifted toward negative values (median(γ) = -0.04 , 66% < 0, $P < .0001$, $N = 186$). For other types of traits, γ estimates were distributed more symmetrically around 0 (fig. 4B), and median(γ) values were near 0 in each case (median(γ): other morphology < 0.01 ; phenology < 0.01 ; other life history = -0.02). Stinchcombe et al. (2008) showed that many studies reported incorrect values of γ that were off by a factor of 2; our results were not altered by using doubled values of γ and held for different taxa (Stinchcombe et al. 2008). In agreement with those of earlier studies (Kingsolver et al. 2001), these results do not indicate negative quadratic and stabilizing selection for most traits and many study systems (Estes and Arnold 2007).

Discussion

Constraints on Directional Selection

A major goal of our analyses was to explore the potential of fitness trade-offs, indirect selection, and fluctuating selection for constraining or limiting overall directional selection. Our results on this issue are decidedly mixed. For the small number of studies in which total fitness was considered, the magnitude of directional selection via total fitness was greater than that for individual fitness components (table 1), the opposite of what would be predicted if there were trade-offs among fitness components. When we consider directional selection on a single trait via different fitness components, the direction of selection was the same for different fitness components for more than

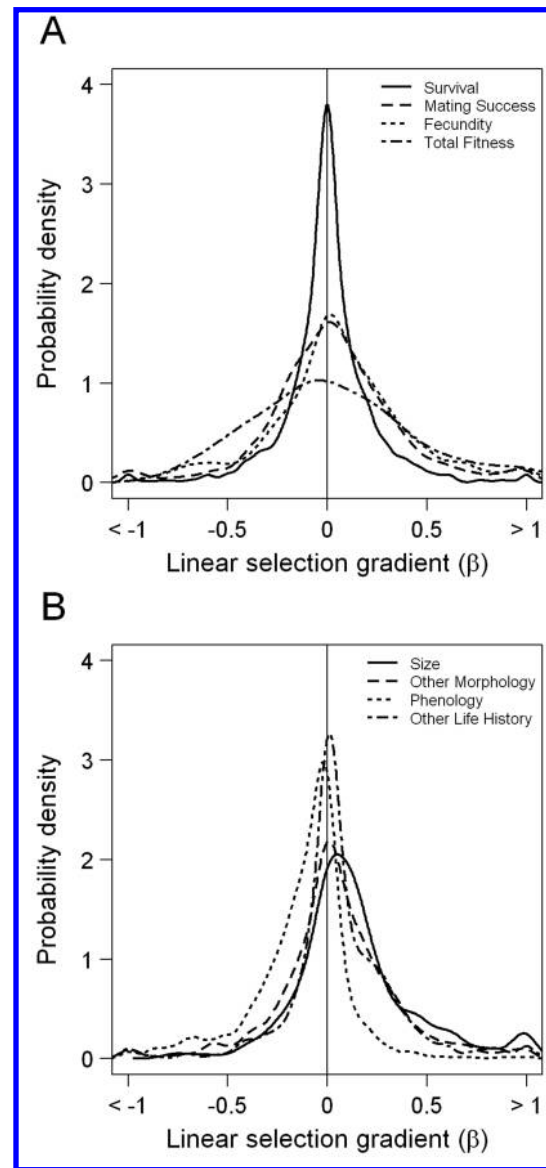


Figure 2: Probability densities for linear selection gradients (β). Here and for all subsequent probability densities of β and quadratic selection gradients (γ), estimates < -1 were assigned a value of -1 , and estimates > 1 were assigned a value of 1 . Separate probability densities are provided for (A) fitness components and total fitness (solid line, survival; dashed line, mating success; dotted line, fecundity; dash-dotted line, total fitness) and (B) major trait classes (solid line, size; dashed line, other morphology; dotted line, phenology; dash-dotted line, other life history).

57% of the estimates, magnifying rather than reducing total selection on the trait (fig. 3). Opposing selection among fitness components rarely resulted in no net selection on a trait. These results are inconsistent with the idea that trade-offs among different components of fitness

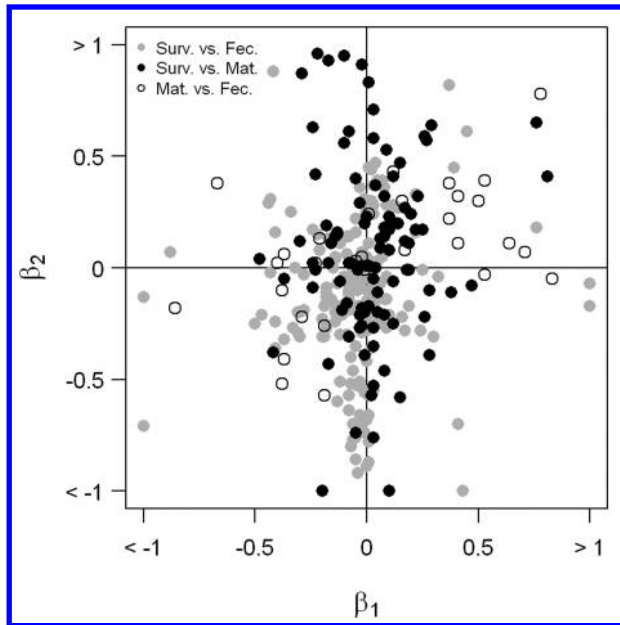


Figure 3: Potential for opposing selection on fitness components to reduce total selection on a trait. Scatterplot of linear selection gradient estimates for which two fitness components were measured on the same trait (within the same study, species, and temporal/spatial replicate). Symbols correspond to the type of fitness component pairing: survival versus fecundity (gray), survival versus mating success (black), and mating success versus fecundity (white).

limit the total directional selection on important phenotypic traits. Fitness trade-offs are widely anticipated theoretically (Roff 2002) and are documented empirically in some study systems (Zera and Harshman 2001; Labonne and Hendry 2010). However, the use of phenotypic data alone to detect fitness trade-offs can be problematic (Reznick 1985) and can yield environmental bias in estimates of phenotypic selection (Rausher 1992; Stinchcombe et al. 2002). If phenotypes are not randomly distributed across environmental conditions or if local environmental conditions affect both phenotype and fitness, this can produce correlations between phenotypic values and fitness (Rausher 1992). Genotypic selection analyses with plants suggest that such environmental covariation can have strong effects on estimates of phenotypic selection (Stinchcombe et al. 2002). If environmental variation affects multiple components of fitness, this could obscure underlying fitness trade-offs. Similarly, phenotypic selection studies cannot detect selection before trait expression, which can bias estimates of the magnitude and even the direction of selection (Mojica and Kelly 2010). Experimental studies or genotypic selection analyses of trade-offs among fitness components may help resolve this issue (Reznick 1985; Stinchcombe et al. 2002).

Indirect selection on correlated traits is another potential mechanism that can reduce or modify the overall directional selection on a trait. In this case, we would expect the slope of the relationship between direct selection (measured by β) and total selection (measured by s) to be <1 . In general, our analyses do not support this prediction for most traits, except for body size (fig. 4). As a result, the current evidence indicates that indirect selection does not generally reduce the total directional selection for most traits. Selection on correlated but unmeasured traits can alter estimates of direct selection (β), and rarely are all other traits that may influence selection on a given trait considered (Mitchell-Olds and Shaw 1987). This can make it more difficult to distinguish direct selection from indirect selection, but this should not bias the estimated slope of the relationship between β and s to be >1 or <1 (fig. 4).

As with fitness trade-offs, environmental covariation can also obscure the effects of indirect selection (Rausher 1992; Stinchcombe et al. 2002). The situation for size is of particular interest. Because there is usually positive directional selection favoring increased size in most systems (fig. 2; Kingsolver and Pfennig 2004), our results suggest that negative indirect selection due to correlated traits often re-

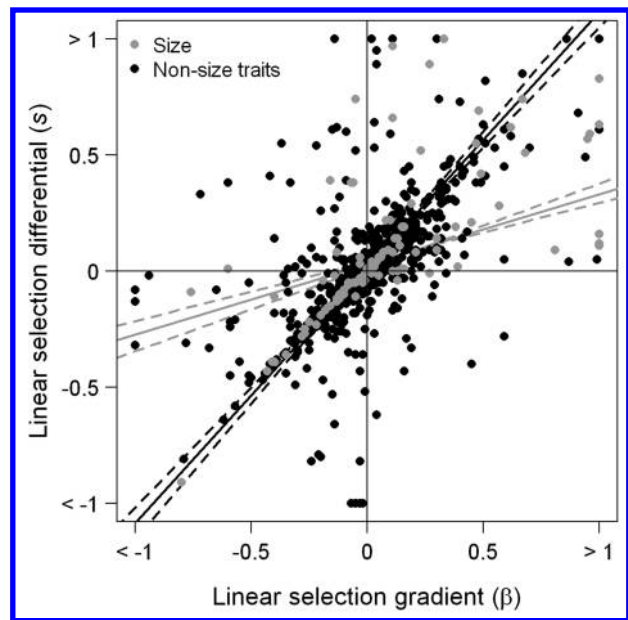


Figure 4: Potential for indirect selection to reduce total selection on a trait. Linear selection differential estimates (s) as a function of linear selection gradient estimates (β) from studies in which both estimates were reported and for which two or more traits were considered. The RMA regression (solid line) and 95% confidence interval (dashed lines) are shown for size (gray lines and symbols) and nonsize (black lines and symbols) traits.

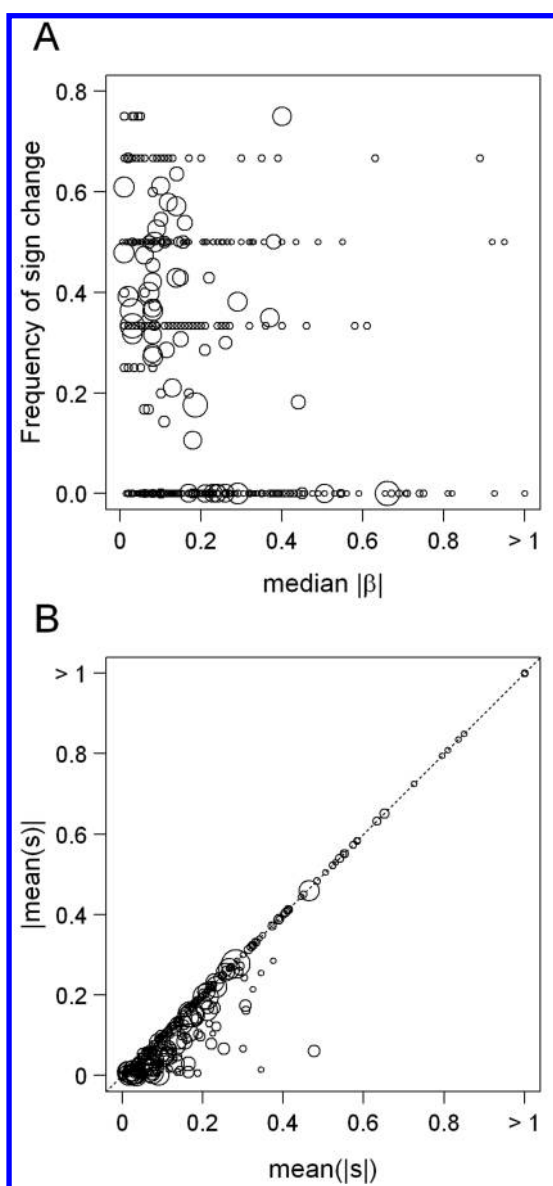


Figure 5: Potential for temporal fluctuations in selection to reduce cumulative selection on a trait. *A*, Frequency of sign changes in the linear selection gradient as a function of the median absolute value of the linear selection gradient for temporally replicated studies. *B*, Absolute value of the mean selection differential over time ($|\text{mean}(s)|$) as a function of the mean of the absolute value of the selection differential over time ($\text{mean}(|s|)$). The dashed line represents a 1:1 correspondence line. In *A* and *B*, each circle corresponds to results for a unique combination of study, species, and trait. Larger circle areas indicate greater temporal replication.

duces the total net selection on size (Lande 1979). One would anticipate, for example, that direct selection for increased size might be countered by opposing indirect selection on phenology or development time (Roff 1980);

unfortunately, there are too few simultaneous estimates of selection on size and development time currently available to evaluate this hypothesis.

Temporal fluctuations in selection could also alter the consequences of directional selection (Siepielski et al. 2009). The key issue is whether such temporal (or spatial) fluctuations are sufficient to reduce the cumulative effects of directional selection over time (Bell 2010). The frequency of reversals in the direction of selection varied widely among traits and studies (fig. 5*A*) but with an overall median of 27%. The duration or number of temporal replications in a study does not appear to alter this basic pattern, but the frequency of sign changes decreases as the average magnitude of selection increases (fig. 5*A*). Some changes in sign are expected simply due to sampling error, especially when selection is weak. Importantly, our analyses suggest that temporal fluctuations in directional selection do not strongly reduce cumulative selection differentials in most studies (fig. 5*B*). As a result, temporal changes in the direction of selection occur regularly in many study systems (Siepielski et al. 2009; van de Pol et al. 2009), but this does not substantially reduce the cumulative effects of directional selection over time (Bell 2010).

In summary, our analyses of directional selection provide little evidence that fitness trade-offs or indirect selection substantially reduces the total directional selection on most traits, though we do detect important indirect selection on size. Environmental covariation may obscure underlying fitness trade-offs and indirect selection not detected by standard phenotypic analyses of selection (Rausher 1992; Stinchcombe et al. 2002). Temporal variation in the direction of selection may occur regularly in natural populations (Siepielski et al. 2009) but appears to have limited impact on cumulative directional selection in most study systems.

Of course, sustained directional selection will lead to evolutionary changes only in populations that possess sufficient genetic variation. Many quantitative genetic studies have documented heritable variation in quantitative traits in natural populations (Mousseau and Roff 1987; Roff 1997), but low genetic variances and multivariate genetic constraints can limit evolutionary responses to selection in some systems (Kellermann et al. 2009; Walsh and Blows 2009). In addition, increased environmental variance in the field can sometimes substantially reduce the effective heritabilities of traits in natural populations (Ruiz et al. 1991; Weigensberg and Roff 1996). Unfortunately, few long-term studies have documented the temporal patterns of selection, genetic variation, and evolutionary responses of phenotypic traits in natural populations to evaluate whether low genetic variation strongly limits directional evolution (Grant 1986; Siepielski et al. 2009).

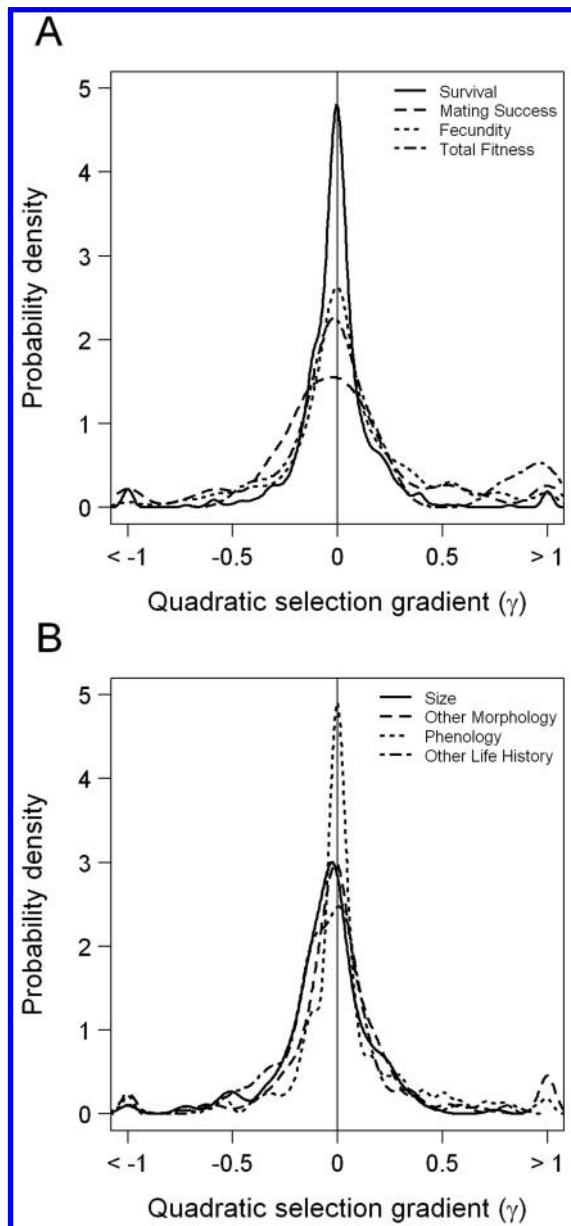


Figure 6: Probability densities for quadratic selection gradients (γ). Separate probability densities are provided for (A) fitness components and total fitness (solid line, survival; dashed line, mating success; dotted line, fecundity; dash-dotted line, total fitness) and (B) major trait classes (solid line, size; dashed line, other morphology; dotted line, phenology; dash-dotted line, other life history).

Nonlinear Selection: Patterns and Problems

If most populations and traits are near maxima in fitness surfaces, we would expect negative γ to be widespread. We see evidence partially supporting this prediction for size (fig. 6B): estimates of γ for size are significantly shifted

toward negative values (median = -0.04). Joint estimates of β and γ (see “Material and Methods”) suggest that mean sizes in some study populations are within 1–2 standard deviations of a nearby fitness peak, supporting the hypothesis that at least some populations are well adapted with respect to body size (Estes and Arnold 2007). This result and interpretation present an interesting puzzle: if populations are near fitness peaks for size, why is directional selection for increased size widespread? Models of stochastic environmental variation in the location of fitness peaks would predict random variation in the direction of selection, so that positive and negative selection on size should be equally common (Lynch and Lande 1993).

For traits other than size, the distribution of γ is symmetric around 0 (fig. 6B). Although correlational selection among traits can alter the estimated magnitude of stabilizing or disruptive selection (Blows and Brooks 2003; Blows 2007), this will not alter the sign of the quadratic selection or the qualitative results reported here. These patterns of quadratic selection do not support the widespread expectation that stabilizing selection on quantitative traits occurs in most natural populations.

There are several possible explanations for this surprising result. First, many studies may choose traits because of an expectation that directional selection may be occurring, so that stabilizing selection is less commonly observed. In that case, the available data are not representative of the actual patterns of quadratic selection in nature. For example, perhaps studies of stabilizing selection use experimental manipulations of phenotype or environment and so would not be represented in our data set or analyses. Further field studies of traits in populations for which stabilizing selection is anticipated would help address this possibility. Second, temporal and spatial variation in the location of fitness peaks may prevent mean trait values from staying near such peaks in most populations. The observed temporal variation in the direction of selection (e.g., fig. 5) is consistent with this interpretation (Siepielski et al. 2009; Bell 2010).

Third, there are important limitations of OLS regression analyses, including the Lande-Arnold approach used in most studies, for estimating nonlinear components of phenotypic selection (Arnold and Wade 1984; Mitchell-Olds and Shaw 1987; Schluter 1988; Geyer et al. 2007; Shaw et al. 2008). Even for relatively smooth fitness surfaces, quadratic selection coefficients can give very misleading information about the curvature of the fitness surface if population mean trait values are not within 1–2 standard deviations of fitness maxima; in particular, the sign of γ does not necessarily reflect the direction of curvature of the fitness surface (Mitchell-Olds and Shaw 1987; Shaw and Geyer 2010). To address this problem, more flexible methods such as cubic splines or projection pursuit re-

gression are used increasingly to fit and visualize fitness surfaces (Mitchell-Olds and Shaw 1987; Schluter 1988; Phillips and Arnold 1989; Schluter and Nychka 1994). However, these methods usually do not provide statistical support for key features of the fitness surface (e.g., curvature or location of fitness maxima or minima), and their results are difficult to compare or generalize across studies. In addition, the statistical distribution of total fitness within a population is typically very irregular and rarely normal or even monotonic (Shaw and Geyer 2010). Although general linear (e.g., OLS regression), generalized (e.g., logistic regression), and additive (e.g., splines) models can be estimated, statistical inferences about features of the fitness surface from these models are suspect if fitness distributions are nonmonotonic (Geyer et al. 2007; Shaw et al. 2008; Shaw and Geyer 2010). Geyer, Shaw, and colleagues (Geyer et al. 2007; Shaw et al. 2008; Shaw and Geyer 2010) have recently developed a likelihood approach for estimating phenotypic selection that combines different statistical distributions for different components of fitness into an integrated statistical framework. This approach holds considerable promise for a more rigorous understanding of patterns of nonlinear selection.

Studies and analyses that consider more complex and integrated aspects of fitness rather than single fitness components are sorely needed. However, progress in our understanding is not entirely limited by lack of adequate data; there are 22 studies (320 pairs of fitness components) already available in which multiple fitness components or total fitness has been reported (fig. 3). Unfortunately, only summary statistics and estimated parameter values are available for the vast majority of these studies. The underlying data on trait values and fitness for each individual in these studies remain inaccessible and, as a result, useless to the research community. Publishing these primary data in public, searchable data repositories such as Dryad (<http://datadryad.org/repo>), thus making them available for additional analyses, would be the fastest and most efficient way to advance our understanding of phenotypic selection in natural populations (Whitlock et al. 2010).

Acknowledgments

We thank S. Carlson, J. DiBattista, and A. Siepielski for making their selection data set publicly available for further analysis. R. Martin, R. Shaw, A. Siepielski, and two anonymous reviewers provided useful comments on previous versions of the manuscript. The Whiteley Center at Friday Harbor and the National Evolutionary Synthesis Center (NESCent) provided ideal venues for planning and initiating the selection analyses and writing the article. Funding was provided by National Science Foundation

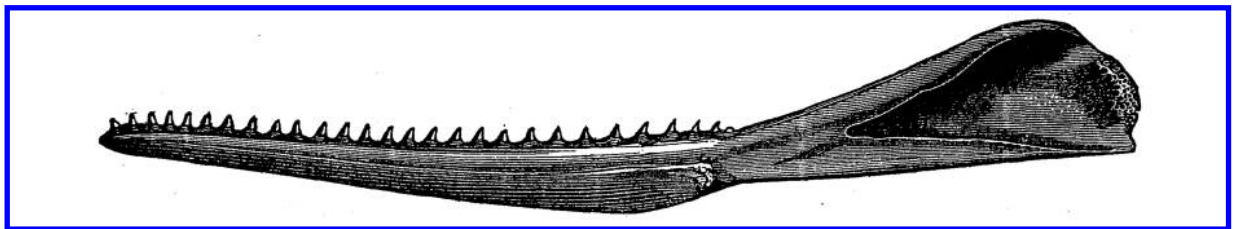
grants IOS-0641179 to J.G.K. and EF-0905606 to NESCent.

Literature Cited

- Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: theory. *Evolution* 38:709–719.
- Bell, G. 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:87–97.
- Blows, M. W. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* 20:1–8.
- Blows, M. W., and R. Brooks. 2003. Measuring nonlinear selection. *American Naturalist* 162:815–820.
- 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.
- Falconer, D. S., and T. F. C. MacKay. 1996. *Introduction to quantitative genetics*. Longman, Essex.
- Geyer, C. J., S. Wagenius, and R. G. Shaw. 2007. Aster models for life history analysis. *Biometrika* 94:415–426.
- Grant, B. R., and P. R. Grant. 1989. Natural selection in a population of Darwin's finches. *American Naturalist* 133:377–393.
- Grant, P. R. 1986. *Ecology and evolution of Darwin's finches*. Princeton University Press, Princeton, NJ.
- Hendry, A. P., and M. T. Kinnison. 1999. The pace of modern life: measuring rates of contemporary microevolution. *Evolution* 53:1637–1653.
- 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.
- Kellermann, V., B. van Heerwaarden, C. M. Sgro, and A. A. Hoffmann. 2009. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science* 325:1244–1247.
- Kingsolver, J. G. 2009. The well-temperated biologist. *American Naturalist* 174:755–768.
- Kingsolver, J. G., and D. W. Pfennig. 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution* 58:1608–1612.
- 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.
- Labonne, J., and A. P. Hendry. 2010. Natural and sexual selection give and take away reproductive barriers: models of population divergence in guppies. *American Naturalist* 176:26–39.
- Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. *Evolution* 33:402–416.
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.

- Legendre, P., and L. Legendre. 1998. Numerical ecology. 2nd ed. Elsevier Science, Amsterdam.
- Lynch, M., and R. Lande. 1993. Evolution and extinction in response to environmental change. Pages 234–250 in P. M. Kareiva, J. G. Kingsolver, and R. B. Huey, eds. Biotic interactions and global change. Sinauer, Sunderland, MA.
- Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* 41:1149–1151.
- Mojica, J. P., and J. K. Kelly. 2010. Viability selection prior to trait expression is an essential component of natural selection. *Proceedings of the Royal Society B: Biological Sciences* 277:2945–2950.
- Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. *Heredity* 59:181–197.
- Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. *Evolution* 43:1209–1222.
- Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to the environmental covariances between traits and fitness. *Evolution* 46:616–625.
- Reznick, D. 1985. Costs of reproduction: an evaluation of the empirical evidence. *Oikos* 44:257–267.
- Reznick, D. N., and C. K. Ghalambor. 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112–113:183–198.
- Roff, D. A. 1980. Optimizing development time in a seasonal environment: the “ups and downs” of clinal variation. *Oecologia (Berlin)* 45:202–208.
- . 1997. *Evolutionary quantitative genetics*. Chapman & Hall, New York.
- . 2002. *Life-history evolution*. Sinauer, Sunderland, MA.
- Ruiz, A., M. Santos, A. Barbadilla, J. E. Quezada-Díaz, E. Hasson, and A. Fontdevila. 1991. Genetic variance for body size in a natural population of *Drosophila buzzatii*. *Genetics* 128:739–750.
- 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.
- Shaw, R. G., and C. J. Geyer. 2010. Inferring fitness landscapes. *Evolution* 64:2510–2520.
- Shaw, R. G., C. J. Geyer, S. Wagenius, H. H. Hangelbroek, and J. R. Etterson. 2008. Unifying life-history analyses for inference of fitness and population growth. *American Naturalist* 172:E35–E47.
- 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* 12:1–16.
- Siepielski, A. M., J. D. DiBattista, J. A. 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: Biological Sciences* (forthcoming).
- Silverman, B. W. 1986. *Density estimation*. Chapman & Hall, London.
- Stinchcombe, J. R., M. T. Rutter, D. S. Burdick, M. D. Rausher, and R. Mauricio. 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: theory and practice. *American Naturalist* 160:511–523.
- 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.
- van de Pol, M., L. Brouwer, B. J. Ens, K. Oosterbeek, and J. M. Tinbergen. 2009. Fluctuating selection and the maintenance of individual and sex-specific diet specialization in free-living oystercatchers. *Evolution* 64:836–851.
- Walsh, B., and M. W. Blows. 2009. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annual Review of Ecology, Evolution, and Systematics* 40:41–59.
- Warton, D. I., I. J. Wright, D. S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. *Biological Reviews* 81:259–291.
- Weigensberg, L., and D. A. Roff. 1996. Natural heritabilities: can they be reliably estimated in the laboratory? *Evolution* 50:2149–2157.
- Whitlock, M. C., M. A. McPeck, M. D. Rausher, L. Rieseberg, and A. J. Moore. 2010. Data archiving. *American Naturalist* 175:145–146.
- Zera, A. J., and L. G. Harshman. 2001. The physiology of life-history trade-offs in animals. *Annual Review of Ecology and Systematics* 32:95–126.

Associate Editor: David N. Reznick
 Editor: Mark A. McPeck



Lower jaw of *Physeter macrocephalus*. From “The Sperm Whales, Giant and Pygmy” by Theodore Gill, MD, PhD (*American Naturalist*, 1871, 4:725–743).