

Patterns and Power of Phenotypic Selection in Nature

JOEL G. KINGSOLVER AND DAVID W. PFENNIG

Phenotypic selection occurs when individuals with certain characteristics produce more surviving offspring than individuals with other characteristics. Although selection is regarded as the chief engine of evolutionary change, scientists have only recently begun to measure its action in the wild. These studies raise numerous questions: How strong is selection, and do different types of traits experience different patterns of selection? Is selection on traits that affect mating success as strong as selection on traits that affect survival? Does selection tend to favor larger body size, and, if so, what are its consequences? We explore these questions and discuss the pitfalls and future prospects of measuring selection in natural populations.

Keywords: adaptive landscape, Cope's rule, natural selection, rapid evolution, sexual selection

Phenotypic selection occurs when individuals with different characteristics (i.e., different phenotypes) differ in their survival, fecundity, or mating success. The idea of phenotypic selection traces back to Darwin and Wallace (1858), and selection is widely accepted as the primary cause of adaptive evolution within natural populations. Yet Darwin never attempted to measure selection in nature, and in the century following the publication of *On the Origin of Species* (Darwin 1859), selection was generally regarded as too weak to be observed directly in natural populations. Several now-classic demonstrations of selection in the wild were published between 1950 and 1975, most notably the case of industrial melanism in peppered moths (Kettlewell 1973). As late as the 1970s, however, industrial melanism remained the primary example of selection in action.

The view that selection is too weak to be measured in the wild has changed dramatically. In the past 25 years, selection has been detected and quantified in hundreds of populations in nature (Endler 1986, Kingsolver et al. 2001, Hereford et al. 2004). Indeed, there are literally thousands of estimates of phenotypic selection in natural populations (Endler 1986, Kingsolver et al. 2001). These data demonstrate that selection occurs routinely in nature and that researchers can measure its action. We are therefore in a position to ask more general questions about phenotypic selection: How strong is selection? Does selection always tend to increase (or decrease) trait values, or are other patterns possible? Do different types of traits experience different patterns or levels of selection? Is

selection on traits that affect survival stronger than on those that affect only mating success? In this article, we explore these and other questions about the patterns and power of phenotypic selection in nature.

What is selection, and how does it work?

Selection is the nonrandom differential survival or reproduction of phenotypically different individuals. Thus, selection requires variation, whereby individuals differ in some of their characteristics, and differential reproduction, whereby some individuals have more surviving offspring than others because of their distinctive characteristics. Those individuals that do have more surviving offspring are said to have higher fitness (note that fitness is a relative, not an absolute, measure). When the characteristics under selection show heredity (i.e., when parents pass on some of their characteristics to their offspring), selection will lead to evolutionary change in these characteristics. Indeed, when populations exhibit variation, heredity, and differential reproduction for a trait, evolution by natural selection will occur. Because these three conditions are met for many traits in many populations, evolution by natural selection is widespread.

Joel G. Kingsolver (e-mail: jgking@bio.unc.edu) and David W. Pfennig (e-mail: dpfennig@email.unc.edu) work in the Department of Biology at the University of North Carolina, Chapel Hill, NC 27599. © 2007 American Institute of Biological Sciences.

The factors in the environment that exert selection—both the biological ones, such as an individual's competitors, predators, and parasites, and the nonbiological ones, such as the weather—are called agents of selection. Traits on which agents act are termed targets of selection.

Regardless of the precise agent or target of selection, phenotypic selection can take several forms. To understand these forms, we first need to clarify the nature of phenotypic variation. Most traits in most organisms show continuous variation. Such traits—termed quantitative traits—are determined by the combined influence of many different genes and the environment.

When selection acts on quantitative traits, three main patterns, or modes of selection, are possible. These three modes can be visualized for a population by mapping (or, more formally, regressing) the fitness associated with a particular phenotype onto the range of all possible phenotypes in that population. This regression provides a statistical estimate of the fitness function. The three modes of selection are defined according to the shape of the fitness function, which describes the relationship between fitness and the phenotype and determines the strength and form of natural selection (figure 1).

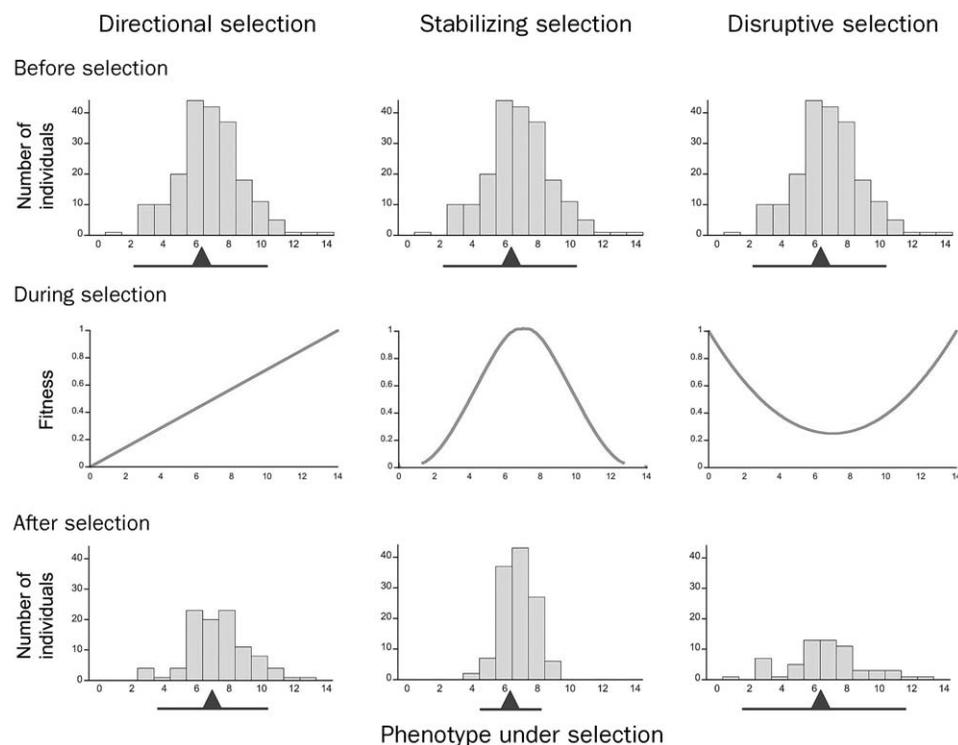


Figure 1. Three different modes of selection (directional, stabilizing, and disruptive), showing the trait distribution of a hypothetical population before selection (top), the fitness function (center), and the trait distribution after selection among the survivors (bottom) for each mode. The triangle under each histogram indicates the mean of each population; the bar under each histogram indicates the variation (± 2 standard deviations) of each population. Modified from Freeman and Herron (2004); data from Cavalli-Sforza and Bodmer (1971).

The first mode, directional selection, is characterized by a linear fitness function (i.e., a straight line). Here, fitness consistently increases (or decreases) with the value of the trait. With positive directional selection, fitness increases with increasing trait values, whereas with negative directional selection, fitness decreases with increasing trait values. Directional selection also tends to reduce variation in a population, although often not dramatically.

The second mode, stabilizing selection, is characterized by a nonlinear fitness function (i.e., a curved line). Here, individuals with intermediate trait values have the highest fitness. Although stabilizing selection does not tend to change the mean trait value, it does tend to reduce variation in a population by disfavoring individuals in the tails of the trait's distribution.

The third mode, disruptive selection, is also characterized by a nonlinear fitness function, but here, individuals with extreme trait values have the highest fitness. As with stabilizing selection, disruptive selection does not tend to change the mean trait value. Unlike stabilizing selection, however, disruptive selection increases variation by favoring individuals in the tails of the trait's distribution.

When the fitness function is described by a straight line (as in directional selection), the slope of the linear regression line measures the strength of selection. When the fitness function has curvature (as in stabilizing and disruptive selection), quadratic regression is required to estimate the strength of selection. In this case, the fitness (w) of a trait (z) can be estimated according to the equation

$$w = \alpha + \beta z + (\gamma/2)z^2,$$

where α is the y -intercept of the fitness function (i.e., of the regression line), β is the fitness function's slope, and γ measures the amount of curvature in the fitness function. Quadratic regression of fitness on phenotype indicates disruptive performance when $\beta = 0$ and γ is significantly positive (Lande and Arnold 1983). By contrast, quadratic regression of fitness on phenotype indicates stabilizing performance when $\beta = 0$ and γ is significantly negative (Lande and Arnold 1983). In both cases, γ measures the strength of quadratic selection.

All three modes of selection drive evolution by eliminating

individuals with low fitness and preserving individuals with high fitness. Moreover, as noted earlier, if the trait of interest is heritable, then evolution will result, but the resulting trait distribution will differ depending on the mode of selection. In particular, for traits under positive directional selection, the population will evolve larger trait values, whereas for those under negative directional selection, the population will evolve smaller trait values. For traits under stabilizing selection, the population will evolve a smaller range of trait values as the average trait value becomes more common in the population. Finally, for traits under disruptive selection, the population will evolve a wider range of trait values, possibly leading to the evolution of alternative phenotypes. Disruptive selection may even promote the formation of new species if the two phenotypic extremes become reproductively isolated from one another.

Measuring selection

In principle, measuring phenotypic selection is straightforward. Suppose we are interested in possible selection acting on some trait, z , in a population. We can measure the trait values for a sample of individuals from the population and then estimate the fitness associated with different trait values (e.g., by measuring the body size or reproductive condition of individuals with different trait values; box 1). Alternatively, we can follow individuals over time and measure components of fitness, such as survival, mating success, or fecundity. In either case, because the evolutionary consequences of selection depend on relative (not absolute) fitness, the fitness value for an individual should be standardized to the mean fitness of all members of the population. The relationship between variation in relative fitness and variation in the trait values represents selection on the trait (estimated from β for directional selection and from γ for quadratic selection; see “What is selection, and how does it work?” above).

A critical assumption of this approach is that variation in the trait causes the observed variation in fitness. Three factors can complicate this relationship, however. First, rather than acting directly on the trait of interest, selection may be acting on other, unmeasured traits that are correlated with the trait of interest, generating a spurious correlation between the measured trait and fitness. One way to reduce this problem is to estimate directional selection on a set of traits that may influence fitness (box 2). This allows us to distinguish direct selection on the trait from the indirect effects of correlated traits; the strength of direct selection is called the selection gradient (β).

A second complication involves environmental effects. If environmental conditions affect fitness, and individuals with different traits experience different environmental conditions, this can alter the measured relationship between traits and fitness and thus estimates of selection (Rausher 1992, Stinchcombe et al. 2002). A useful experimental solution is to randomize the locations (environments) of individuals with different phenotypes or genotypes (Rausher 1992), but this can be difficult to achieve in many natural environments.

An alternative approach to estimating selection, dubbed “phenotypic engineering,” involves experimentally manipulating phenotypic traits and evaluating the effects of the manipulation on subsequent fitness in natural environments, relative to appropriate controls (Sinervo and Basolo 1996). This method has been used to demonstrate selection on particular phenotypes in a number of systems (Sinervo et al. 1992, Grether 1996). Phenotypic engineering is especially useful for determining whether a trait is under selection and what mode of selection might operate on it, because it can expand the range of phenotypic values and reduce the problem of correlated traits (Travis and Reznick 1998). However, because phenotypic engineering often involves altering trait expression beyond the range of trait values observed in natural populations, such manipulations do not help researchers estimate the strength of selection on natural populations in the wild.

A third complication is that different phenotypic traits have different units and dimensions (e.g., body mass versus age at first reproduction), and changes in a single trait have different consequences in different organisms (e.g., a 1-gram change in body mass is a much greater increase in relative size in mice than in whales). To compare selection across different traits and systems, we need to standardize selection. One common approach is to standardize the selection gradient relative to the standard deviation (σ) of the phenotypic trait. The standardized selection gradient β_σ has a natural interpretation: It is the change in relative fitness that results from 1 standard deviation of change in a trait. Thus, if $\beta_\sigma = 0.5$, moving 1 standard deviation away from the population mean increases relative fitness by 50%.

With these statistical tools in hand, we ask: How strong is selection? Is selection on traits associated with survival stronger than on traits associated with mating success? How common are stabilizing selection and disruptive selection in nature?

How strong is selection in nature?

Numerous studies have measured phenotypic selection in natural populations using the methods described above (Endler 1986). We are therefore in a position to synthesize these studies and look for more general patterns of selection. Such a synthesis has been undertaken recently. Kingsolver and colleagues (2001) reviewed selection studies published between 1984 and 1998 and identified 63 studies of 62 species involving a wide range of taxa, geographic areas, and types of traits. These studies yielded 993 estimates of directional selection (β_σ). Positive and negative values of β_σ occur with equal frequency, so it is more informative to consider the absolute value, $|\beta_\sigma|$, as an indicator of the magnitude of directional selection.

A frequency distribution of $|\beta_\sigma|$ shows a wide range of values, with small values most common but with a long “tail” of higher values (figure 2; Kingsolver et al. 2001). For example, the median value was 0.16, and 13% of the values were greater than 0.5, indicating very strong selection. To put this in perspective, imagine a population in which a heritable trait ($h^2 = 0.5$; see box 2) experiences persistent directional se-

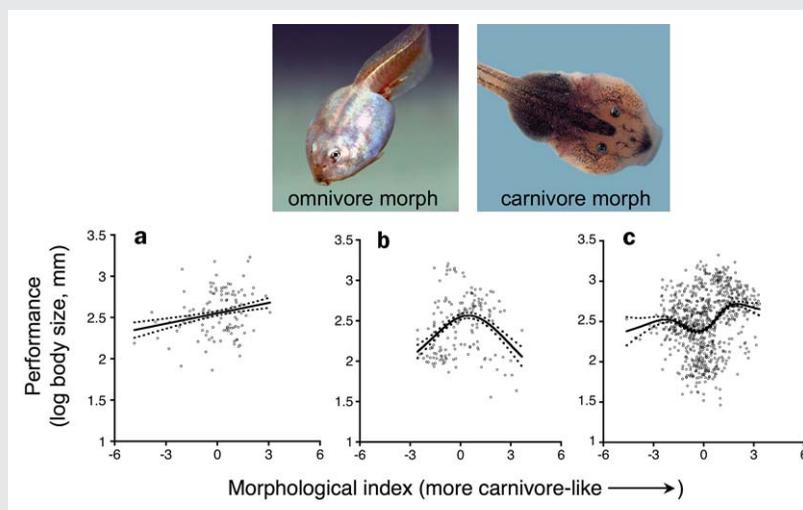
Box 1. Different modes of selection in natural populations: A case study from amphibians.

Populations confronting different ecological circumstances can undergo different modes of selection. For example, Mexican spadefoot toads (*Spea multiplicata*) and Plains spadefoot toads (*Spea bombifrons*) co-occur in the southwestern United States. Their tadpoles are highly variable in resource use and trophic morphology, as represented by two extreme morphotypes (Orton 1954, Pomeroy 1981, Pfennig 1992): (1) the omnivore morph, a round-bodied tadpole with a long intestine, small jaw muscles, and smooth mouthparts used for feeding on detritus (60% by gut volume; Pomeroy 1981) and anostracan fairy shrimp (38% by gut volume; Pomeroy 1981); and (2) the carnivore morph, a narrow-bodied tadpole with a short intestine, greatly enlarged jaw muscles, and notched mouthparts used for feeding on larger anostracan fairy shrimp (85% by gut volume; Pomeroy 1981) and detritus (see figure). In some ponds, there is clear dimorphism in trophic morphology; in other ponds, intermediates—both in morphology and in resource use—may be the most common phenotype present (Pomeroy 1981, Pfennig 1990).

Using body size as a proxy for fitness (body size correlates with several important fitness components in larval amphibians), Pfennig and colleagues (2007) found that the mode of selection operating on trophic morphology varies for different species and populations. Specifically, in mixed-species ponds, the most carnivore-like *S. bombifrons* tadpoles were the largest (see the figure, panel a; cubic splines [solid lines] are bracketed by 95% confidence intervals [dashed lines] estimated from 1000 bootstrap replicates). This observation suggests that directional selection favors more carnivore-like *S. bombifrons*. Presumably, this pattern reflects selection on *S. bombifrons* to express resource-use phenotypes that minimize their overlap with *S. multiplicata* for food; *S. multiplicata* tend to be more omnivore-like than *S. bombifrons*.

A different mode of selection was detected among *S. multiplicata* in these mixed-species ponds. Here, stabilizing selection appears to favor individuals with intermediate phenotypes (see panel b). Presumably, carnivore phenotypes in these individuals are selectively disfavored because they are competitively inferior to *S. bombifrons* (Pfennig and Murphy 2002). Yet why does selection not favor omnivores, which are as distinct as possible from *S. bombifrons*? Pfennig and colleagues (2007) hypothesize that selection acts against *S. multiplicata* omnivores in mixed-species ponds because omnivores metamorphose later and at a smaller body size than carnivores. Because mixed-species ponds typically contain relatively high shrimp densities, *S. multiplicata* that express an intermediate trophic phenotype—and that can thereby supplement their detritus diet with, but not specialize on, the more nutritious shrimp resource—may be selectively favored. Thus, in mixed-species ponds, selection appears to favor *S. multiplicata* that are as carnivore-like as possible while simultaneously minimizing resource overlap with *S. bombifrons*.

Finally, a third mode of selection was detected among *S. multiplicata* in single-species ponds (see panel c). Here, disruptive selection favors extreme trophic phenotypes. In these ponds, individuals expressing trophic phenotypes on either end of a resource-use spectrum would most likely have fewer (and, in the case of extreme omnivores, perhaps lower-quality) resources available. Nevertheless, compared with the majority of the population that may be intermediate in phenotype (and in resource use), individuals on opposite ends of the resource spectrum would also most likely have fewer competitors with which to share those resources. Thus, relative to intermediate individuals, the overall fitness of extreme omnivores and carnivores may be high. Additional evidence that such density-dependent disruptive selection favors extreme phenotypes comes from field experiments demonstrating that the two morphs are maintained within ponds by negative frequency-dependent selection (Pfennig 1992), which is a hallmark of competitively mediated disruptive selection (Day and Young 2004).



lection of median magnitude ($\beta_{\sigma} = 0.16$). In fewer than 50 generations, the population mean would shift by 3 standard deviations, thereby exceeding the initial range of variation in the population. Thus, phenotypic selection in many natural populations is strong enough to cause substantial evolutionary changes in tens to hundreds of generations, which is a very short timescale in evolutionary terms (Reznick et al. 1997, Hendry and Kinnison 1999, Hoekstra et al. 2001).

Several complications temper this important conclusion, however (Kingsolver et al. 2001, Hereford et al. 2004, Hersch and Phillips 2004). First, studies that fail to detect strong or significant selection are less likely to be published, particularly if the study has a small sample size. This leads to a publication bias, in which studies with larger effects are more likely to be reported than those with smaller effects. There is some indication of such publication biases in the selection data, slightly inflating the average magnitude of selection detected (figure 2; Kingsolver et al. 2001, Hersch and Phillips 2004). Second, many selection studies have small sample sizes that limit their statistical power. For example, as illustrated in figure 2, only 25% of the individual values of β_{σ} are significantly different from zero at the 95% significance level (one would expect 5% of the values to be significant as a result of chance alone). Consequently, most studies have insufficient statistical power to detect selection of typical magnitude (Hersch and Phillips 2004). Thus, selection is potentially potent, albeit typically difficult to detect. A third limitation is that most studies measure selection in terms of one or more components of fitness (e.g., aspects of an individual's survival, mating success, or fecundity) rather than total lifetime fitness (e.g., the total number of surviving offspring that an individual produces). Indeed, less than 5% of the available measurements of phenotypic selection involve total lifetime fitness, which is difficult to measure in most natural field populations (Kingsolver et al. 2001). This is important because the magnitude and even the direction of selection on a trait may differ for different components of fitness. On the other hand, a recent statistical analysis by Knapczyk and Conner (forthcoming) indicates that sampling error does not bias estimates of the average strength of phenotypic selection, and suggests that publication bias is detectable only for selection estimates with very small sample sizes.

A recent alternative approach to assessing the magnitude of selection is to standardize the selection gradient using the mean value of the trait rather than the standard deviation (Hereford et al. 2004). The mean-standardized gradient, β_{μ} , has a useful and natural interpretation: Selection on fitness itself would produce a β_{μ} of 1. A recent survey of selection studies from 1984 through 2003 reported a bias-corrected median value for β_{μ} of 0.31, and more than 20% of the values exceeded 1, indicating that selection on these traits was stronger than selection on fitness itself (Hereford et al. 2004). As Hereford and colleagues (2004) note, such large values "cannot be representative of selection on all traits." However, there are a number of limitations to the use of mean-standardized measures of selection. First, the inter-

Box 2. Selection and evolution of multiple traits.

Evolution by natural selection requires three conditions: variation, inheritance, and selection (differential reproduction). We can describe quantitatively how evolution proceeds from these conditions. Suppose we have a trait z that is experiencing directional selection, with a selection gradient β . The evolutionary change in the mean trait value of the population per generation is given by

$$\Delta\bar{z} = G\beta,$$

where G is the additive genetic variance for the trait (Lande 1979). Thus the amount of evolutionary change per generation is simply the product of the genetic variation and the strength of selection on the trait. A useful alternative way to consider genetic variation is in terms of heritability (h^2): $h^2 = G/P$, where P is the phenotypic variance in the trait. Heritability indicates the fraction of the total population variation in a trait that is due to the additive effects of genes.

This relationship can readily be extended to multiple correlated traits (Lande and Arnold 1983). Consider two traits, z_1 and z_2 , that experience directional selection gradients β_1 and β_2 . Then the evolutionary change in the two traits per generation is expressed as

$$\Delta\bar{z}_1 = G_{11}\beta_1 + G_{12}\beta_2$$

and

$$\Delta\bar{z}_2 = G_{21}\beta_1 + G_{22}\beta_2,$$

where G_{11} and G_{22} are the additive genetic variances for traits 1 and 2, and $G_{12} = G_{21}$ is the genetic covariance between traits 1 and 2. The new feature here is the effect of genetic covariance on evolution. Genetic covariances can arise when some of the same genes affect multiple traits: For example, some genes can affect both body weight and brain weight (Lande 1979). Suppose that there is positive directional selection on trait 1 but not on trait 2 (i.e., $\beta_2 = 0$) and that there is a positive genetic covariance between the traits (i.e., $G_{12} > 0$). Then trait 2 will evolve even though there is no direct selection on it, as a result of selection on trait 1 and of the genetic covariance between the two traits. Correlations among traits can have important effects on how organisms evolve in response to selection (Lande and Arnold 1983).

pretation of β_{μ} is valid only for traits that represent true ratios and where the zero value is not arbitrary. This limitation excludes many interesting phenotypic traits, such as

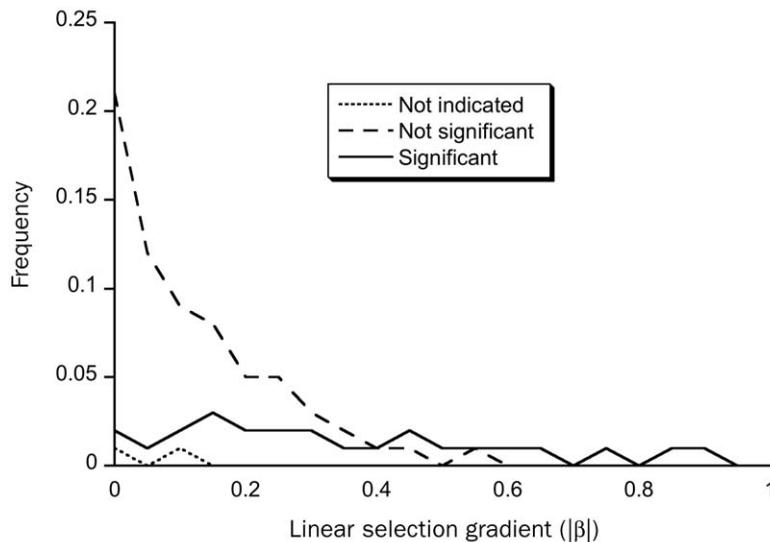


Figure 2. Frequency distribution of the magnitude of directional selection ($|\beta|$). Different distributions are shown according to the statistical significance of each individual estimate. Modified from Kingsolver and colleagues (2001).

phenology and seasonal timing, and composite traits, such as principal components (Kingsolver et al. 2001). A second, practical issue is that because the information needed to compute β_{μ} is not always reported in published studies, this approach excludes up to 70% of the available data on phenotypic selection. Third, analyses indicate that large values of β_{μ} are consistently associated with small values of the coefficient of variation (CV), the ratio of the standard deviation to the mean of the trait. For example, for values of β_{μ} greater than 1, the median value of CV was 0.10—a mean 10 times greater than the standard deviation. In contrast, for values of β_{μ} less than 1, the median value of CV was 0.26. There is no obvious biological reason for very strong selection to be associated with small CV values (i.e., with traits that show small variation relative to the mean), and a statistical explanation for this pattern is more likely.

Given the enormous diversity of organisms, we are usually interested not in average selection but rather in differences in selection among different components of fitness, agents of selection, and targets of selection. One important issue to resolve is whether the relative magnitude of phenotypic selection due to variation in survival or fecundity (natural selection) is greater than that due to variation in mating success (sexual selection). The data on directional selection gradients (β_{σ}) indicate that sexual selection is significantly stronger than natural selection (figure 3). For example, the median magnitude of sexual selection is more than twice as great as that of natural selection, a pattern that holds for diverse plant and animal taxa. This result

suggests that competition for mates may be important for rapid evolution in nature. Many people view evolution as a “struggle for existence.” Yet the struggle for existence may often be less important than the struggle to mate.

Selection on size: Cope’s rule

Body size is an especially common target of selection (box 1). This is perhaps not surprising, given that an organism’s body size affects nearly every aspect of its biology, from its biochemistry to its ecology (Bonner 2006). A striking pattern that has emerged from investigations into the evolution of body size is a tendency for species within a taxonomic group to evolve toward larger body size, a pattern known as Cope’s rule. While exceptions are known, Cope’s rule has been documented in numerous plant and animal taxa (box 3; Hone and Benton 2004). Many explanations for Cope’s rule have been proposed, ranging from statistical artifact to differences in extinction rates. We were interested in whether phenotypic selection on body size within natural

populations could account for Cope’s rule.

To address this question, we considered studies of the strength of directional selection (β_{σ}) on body size compared with other morphological traits (Kingsolver and Pfennig 2004). We identified 42 studies that measured selection on morphological traits including size, and 20 studies that measured selection both on body size and on other morphological traits within the same study. When we plotted the frequency distribution of selection strengths (β_{σ}) from these studies, a clear pattern emerged (figure 4a). For morphological traits excluding size, this frequency distribution is sym-

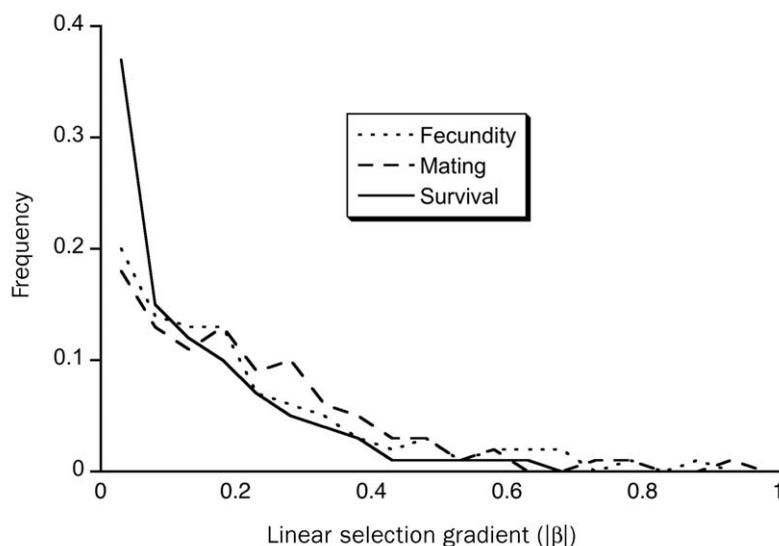


Figure 3. Frequency distribution of the magnitude of directional selection ($|\beta|$) for selection via three different components of fitness (fecundity, mating, and survival). Modified from Kingsolver and colleagues (2001).

metric about zero, with 50% of the values greater than zero and a median value for β_{σ} of 0.02. This is not surprising: Sometimes there is positive selection and sometimes negative selection on various morphological traits in different studies. In contrast, the distribution of directional selection values for body size is strongly skewed toward positive values: 79% of the values exceed zero, and the median value of $\beta_{\sigma} = 0.15$ (figure 4a).

Selection appears to favor larger size, regardless of whether increased size is thought to increase survival (figure 4b), fecundity (figure 4c), or mating success (figure 4d). In most studies of natural populations to date, larger individuals have higher survival, greater fecundity, and greater mating success—that is, bigger is generally fitter.

Does it follow, then, that organisms will evolve larger size? Recall that directional selection for a trait, such as increased size, will lead to evolutionary change only if there is heritable variation for the trait. Heritable variation for body size exists in most natural populations that have been studied. There may also be opposing selection on traits that are correlated with size. For example, longer development time (time to reach adulthood or sexual maturity) is frequently genetically correlated with larger body size, but there may be selection for shorter development time that opposes selection favoring larger size. The available data indicate some evidence for selection favoring shorter development times, but this is not sufficient to counterbalance selection on size (Kingsolver and Pfennig 2004).

What are the evolutionary consequences of consistent directional selection for larger size? A selection gradient of 0.15 and a modest heritability ($h^2 = 0.33$) would lead to an evolutionary increase in the mean size in a population by 0.05 standard deviations each generation. This rate of evolution falls well within the range of microevolutionary change observed in some populations within the past century (Hendry and Kinnison 1999). If extrapolated over a longer time period, this could translate into substantial increases in body size in a species or evolutionary lineage. Directional selection on size of the magnitude we have documented would increase the mean size of individuals by 5 standard deviations in only 100 generations—much faster than the rates measured from the fossil record that illustrate Cope's rule. As a result,

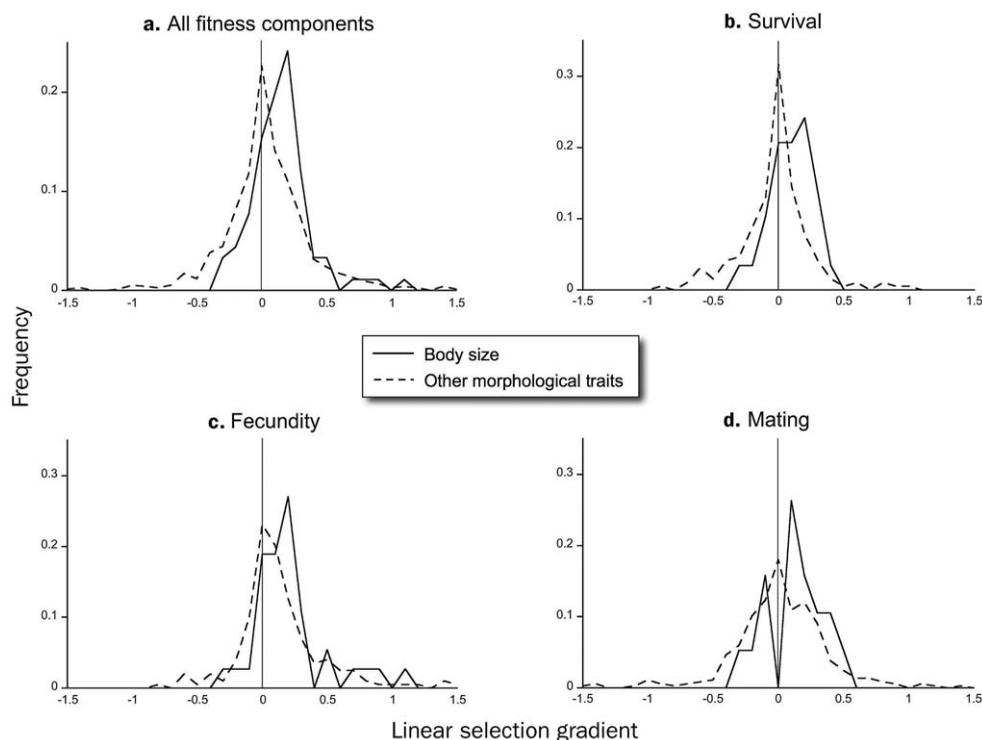


Figure 4. Frequency distribution of the magnitude of directional selection ($|\beta|$) for body size and for other morphological traits for (a) all fitness components, (b) traits related to survival, (c) traits related to fecundity, and (d) traits related to mating success. Modified from Kingsolver and Pfennig (2004).

the positive directional selection observed in contemporary populations is more than sufficient to account for Cope's rule.

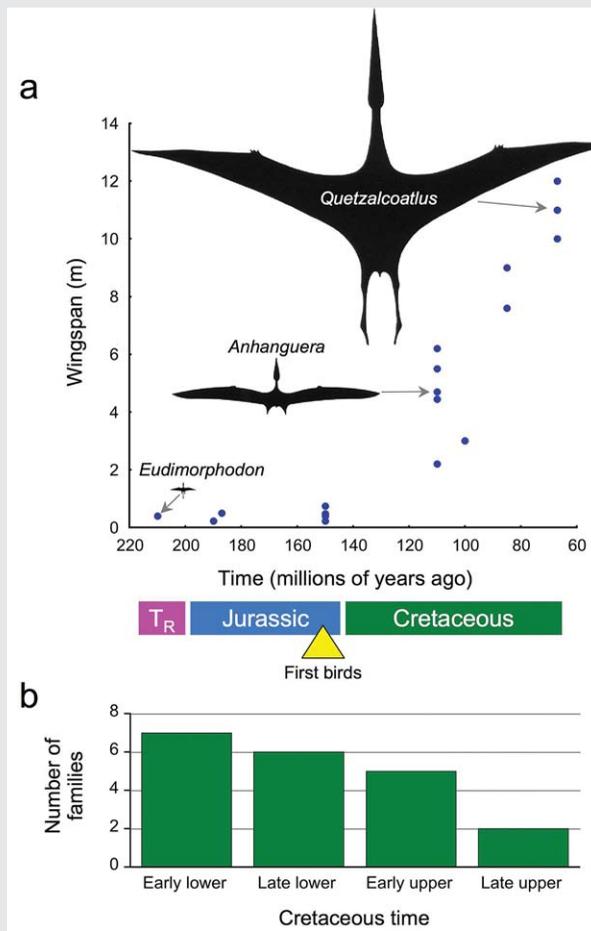
But our proposed explanation for Cope's rule also presents a paradox. If selection generally favors larger size, why aren't more contemporary species near their maximum potential size? Indeed, the largest known species of arthropods, insects, amphibians, reptiles, birds, and land mammals lived millions of years ago; the largest present-day representatives of these groups are much smaller. What prevents organisms from evolving toward ever-increasing size? The most likely explanations involve extinction. Species with larger body sizes generally have smaller population sizes, have longer generation times, and require larger areas of habitat (Bonner 2006), all of which increase the likelihood of species extinction during periods of environmental change. Many of the world's most threatened and endangered species of vertebrates have relatively large body size. For example, during the widespread extinctions of mammals in North America that followed the end of the last ice age, large-bodied species were particularly hard hit: Mammoths and mastodons, American horses and camels, giant ground sloths, cave bears, and saber-toothed cats all went extinct. More generally, studies of mass extinctions of diverse taxa throughout life's history reveal that large species are often more likely to go extinct than their smaller relatives. As a result, extinction may help to explain why

Box 3. Cope's rule in pterosaurs.

Pterosaurs were flying diapsid reptiles (other diapsids include ichthyosaurs, plesiosaurs, lizards, crocodiles, and dinosaurs). From the time they first appeared 220 million years ago to the time they went extinct 65 million years ago (during the end-Cretaceous mass extinction), pterosaurs increased dramatically in overall body size. (Panel a in the figure presents the estimated maximum wingspan for 18 genera of pterosaurs, based on data in Lawson 1975, Maisey 1991, Hazelhurst and Rayer 1992, Company et al. 2001, Buffetaut et al. 2002, Chiappe et al. 2004, and Unwin 2006.) In fact, during their 155-million-year reign, pterosaurs increased in size by a remarkable 3000%.

Pterosaurs underwent their most impressive increase in size during the Cretaceous period (144 million to 65 million years ago), shortly after birds first appeared (about 150 million years ago). One hypothesis for this increase in size is that it may have been driven by competition from birds. Possible evidence of such competition is provided by fossil assemblages in China, which reveal that birds were more common in terrestrial, inland areas, whereas pterosaurs were more abundant in coastal areas (Wang et al. 2005).

Regardless of the precise selective agent, if any, that may have favored larger size, during the time they were undergoing their most dramatic size increase (i.e., during the Cretaceous) pterosaurs were also becoming less diverse (panel b; data from Unwin 2006). Therefore, by becoming larger, pterosaurs may have paid a cost in terms of increased vulnerability to extinction, a pattern observed in many taxa.



most organisms remain relatively small in the face of continuing natural and sexual selection for larger size within populations.

Patterns of quadratic selection

So far, we have emphasized the importance of directional selection in generating evolutionary adaptation and evolutionary change. As noted earlier, however, nonlinear modes of selection are also possible. In quadratic selection, which affects variation rather than the mean trait value in a population, the relationship between fitness and the phenotype is curved (box 1, figure 1). Recall that we can quantify the strength of quadratic selection in terms of the quadratic selection gradient γ , which reflects the curvature of the regression between the trait and fitness. If most populations are well adapted to their current environment, we would expect stabilizing selection to be common and most γ values to be negative. Conversely, disruptive selection, in which γ is positive, is thought to be relatively rare.

What patterns of quadratic selection are observed in natural populations? Kingsolver and colleagues (2001) identified 574 measures of γ . The frequency distribution of γ is sym-

metric about zero, with negative and positive values equally common (figure 5). Fifty percent of the γ values are between -0.1 and $+0.1$, indicating that the magnitude of quadratic selection is rather small; only 16% of the values are significantly different from zero. Thus, stabilizing selection appears to be no more common than disruptive selection, a surprising result that we will return to shortly.

What about the magnitude of quadratic selection? For illustration, a value of -0.1 for γ indicates that individuals 2 standard deviations away from the mean phenotype (about 5% of the population) will have levels of fitness that are 40% below the maximum fitness, a substantial effect. However, the vast majority of γ values of this magnitude are not significantly different from zero (figure 5). This suggests that most studies of quadratic selection do not have the sample size or statistical power to quantify selection of the magnitude that may be typical in natural populations.

Several other factors complicate our interpretation of these results. There is clear evidence for publication bias, in which studies with small sample sizes are more likely to be published if the γ values are larger or statistically significant. Such biases will inflate the magnitude of selection reported in the liter-

ature. Yet when multiple traits are involved, estimating quadratic selection one trait at a time can result in underestimating the magnitude of selection (Blows and Brooks 2003). Environmental biases can also cause underestimates of quadratic selection (Stinchcombe et al. 2002). Moreover, few studies have focused specifically on quadratic selection (Brodie et al. 1991, Blows et al. 2003, Brodie and Ridenhour 2003, Blows 2007; but see Bolnick 2004, Pfennig et al. 2007), so perhaps the paucity of evidence for strong quadratic selection is not surprising. In sum, there is an urgent need for well-designed field studies to measure selection in populations where either form of quadratic selection might be anticipated.

From selection to adaptive landscapes

Phenotypic selection involves the relationship between the trait values and the relative fitness of individuals within a population (box 1). A related concept is the adaptive landscape, which connects the mean trait value of a population to the population's mean fitness (Wright 1932, Lande and Arnold 1983, Phillips and Arnold 1989). The adaptive landscape can be thought of as a surface, consisting of adaptive peaks (mean trait values associated with high mean fitness) and valleys (mean trait values associated with low mean fitness), over which a population can potentially move. Any given population resides at a point on the adaptive landscape, representing the mean phenotype of the individuals that comprise the population. The slope of the landscape at that point indicates the strength of directional selection on the population. Selection should tend to drive the population "uphill" toward the nearest adaptive peak. Once the population reaches the peak, stabilizing selection should keep it there. Because it is generally thought that most organisms are well adapted to their environment, it is commonly assumed that most populations reside at adaptive peaks. If most populations are indeed at or near adaptive peaks, then we would expect that most populations would experience stabilizing rather than directional selection, and that disruptive selection should be uncommon.

The data on selection in natural populations do not match these predictions. Based on the available measurements of γ , stabilizing selection appears to be no more common than disruptive selection (figure 5), and many populations experience at least moderate levels of directional selection. This finding suggests that most populations are not currently at local peaks in the adaptive landscape. An interesting recent analysis uses values of β and γ to compute how far populations are currently from nearby adaptive peaks (Estes and Arnold 2007). When $\gamma < 0$ (as in approximately 50% of the cases), the typical population is only 1 standard deviation about from a fitness peak (Estes and Arnold 2007). However, the same analysis implies that many populations (when $\gamma < 0$) are only

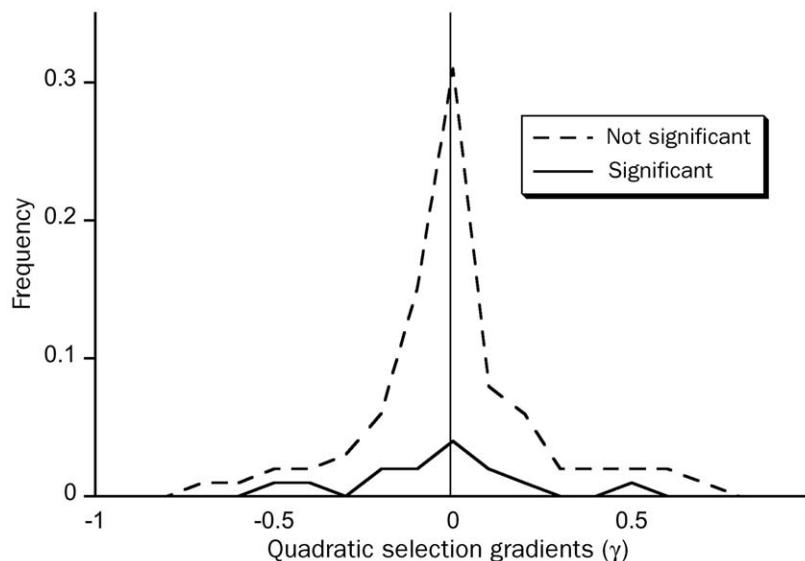


Figure 5. Frequency distribution of the strength of quadratic selection (γ). Stabilizing selection requires a value for γ of less than zero, whereas disruptive selection requires a value greater than zero. Different distributions are shown according to the statistical significance of each estimate. Modified from Kingsolver and colleagues (2001).

1 to 2 standard deviations from fitness valleys, where mean fitness is at a minimum.

Why don't more populations appear to reside at adaptive peaks (Price et al. 1988)? One possibility, discussed earlier, is that published studies do not represent an unbiased estimate of the true frequency or strength of stabilizing selection in natural populations. Another possibility is that random environmental change causes adaptive peaks to fluctuate over time. For example, just as directional selection moves a population close to an adaptive peak, the environment may change, causing the peak (and the entire landscape) to shift to a different range of trait values (figure 6). A shifting adaptive landscape would preclude the population from experiencing stabilizing selection; instead, the population would tend to experience directional selection that fluctuates in both sign (positive or negative) and magnitude. Such a pattern of shifting directional selection has been documented in several systems (Gibbs and Grant 1987, Losos et al. 2006).

Although this analysis can explain why many populations experience at least moderate levels of directional selection, it does not explain why disruptive selection may be as common as stabilizing selection. This result is surprising, because disruptive selection is generally thought to be relatively rare in nature (e.g., Endler 1986). Of course, one possible explanation for the apparent commonness of disruptive selection is that it is an artifact of sampling bias. Recall that most studies of quadratic selection do not have the sample size or statistical power to quantify selection of the magnitude that may be typical in natural populations. Alternatively, disruptive selection may be relatively common, and its widespread occurrence may reflect a ubiquitous agent of selection in nature: competition for re-

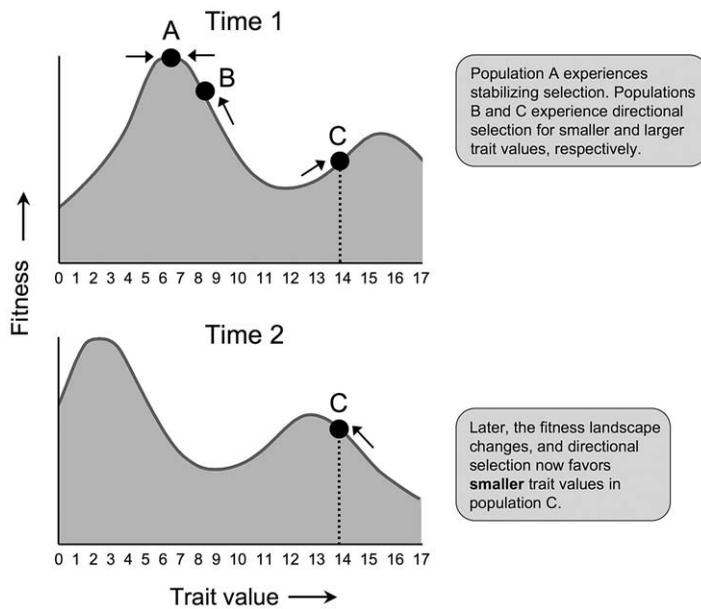


Figure 6. A diagrammatic representation of how a shifting fitness landscape might prevent a population from experiencing stabilizing selection on a particular trait.

sources, such as food. Because competition tends to decrease individual fitness, natural selection is generally thought to favor traits that lessen competition's intensity. One way for selection to do so is to favor evolutionary divergence between initially similar phenotypes through density-dependent or frequency-dependent disruptive selection (Sinervo and Calsbeek 2006). In a population that exploits a continuously varying resource, those individuals that utilize the most common resource (e.g., intermediate-size prey) will initially have a fitness advantage. As more individuals begin to exploit this resource, however, competition will become increasingly severe, and the fitness of these individuals will begin to decline (Day and Young 2004). As long as there is a broad range of resource types, individuals that specialize on less common resources on either end of the resource-use spectrum (e.g., very small or very large prey) will have fewer competitors. Eventually, the fitness of these divergent individuals may exceed that of individuals with intermediate phenotypes, as disruptive selection, driven by resource competition, favors less common, more extreme phenotypes. Evolution resulting from such frequency-dependent disruptive selection may explain the prevalence within many natural populations of alternative morphs for resource use or mating tactics (e.g., box 1; Gross 1996).

Conclusions and future directions

As we have seen, phenotypic selection has now been quantified in numerous organisms and in a broad range of ecological contexts. An analysis of these studies has revealed some interesting, and occasionally unexpected, patterns. We summarize four such patterns here:

First and foremost, phenotypic selection in nature is common and can be measured in the field in real time (figure 2). In particular, directional selection is often sufficiently strong to cause substantial evolutionary change in a relatively short period.

Second, selection acting on traits that influence mating success (e.g., elaborate displays in males) appears to be stronger than selection acting on traits that influence survival or fecundity (i.e., sexual selection tends to be stronger than natural selection; figure 3). Thus, competition for mates may be more important in evolution than is generally assumed.

Third, in most species studied, directional selection favors larger body size (figure 4a). This pattern contrasts with the pattern for other morphological traits, which tend to experience positive and negative directional selection with equal frequency (figure 4a). Moreover, bigger organisms are generally fitter, regardless of whether larger body size enhances survival (figure 4b), fecundity (figure 4c), or mating success (figure 4d). In fact, directional selection favoring larger body size is sufficiently strong to explain Cope's rule, the widespread tendency for lineages to evolve toward larger body size.

Finally, we have little evidence that stabilizing selection is more common than disruptive selection (figure 5). This unexpected result may reflect statistical biases, lack of statistical power, the tendency for environments and adaptive landscapes to change frequently, or the widespread tendency for organisms to compete for scarce resources.

Many questions remain unanswered, however. Here we highlight four such questions that are likely to be fruitful areas for future research.

First, how does phenotypic selection acting on a particular trait change over time? Although phenotypic selection is sometimes strong, it is not clear whether it remains so for long. Environments may change so frequently that the magnitude and direction of selection may also vary frequently. We urgently need more long-term studies of selection in the wild to determine whether the magnitude, the direction, and even the mode of selection tend to vary over time (e.g., Grant and Grant 2006) and space (e.g., box 1).

Second, how common and how strong is stabilizing selection? As we have seen, the available evidence suggests that disruptive selection is as common as stabilizing selection. Does this pattern reflect the true pattern of selection in nature, or does it merely reflect publication bias or some other distortion in the data available?

Third, what component or components of fitness provide the most complete picture of the strength and pattern of selection in nature? A good operational definition of fitness is that it is the total number of offspring that an individual produces in its lifetime. Yet, because it is often not practical to measure the lifetime number of offspring produced, most studies of selection focus on only one component of fitness, such as survival (or even, more indirectly, traits that correlate with survival; see, e.g., box 1). It is often not known, however,

how reliably the measured fitness component predicts true lifetime fitness.

Finally, what role has phenotypic selection played in generating the amazing diversity of life-forms in the world around us? Our review of phenotypic selection in natural populations suggests that selection is often sufficiently potent to account for large-scale phenotypic change over relatively short periods of evolutionary time. Therefore, if selection persists, long-term trends may result from selection acting at the level of the individual. One such trend is Cope's rule (box 3). Do other macroevolutionary trends, such as the increase in diversity over geological time, also emerge from phenotypic selection acting on individuals within populations?

In sum, modern analyses of phenotypic selection reveal a dynamism and complexity that Darwin and his contemporaries probably never imagined. Understanding the patterns and power of phenotypic selection is central to evolutionary biology.

Acknowledgments

We thank Jeff Conner, Joe Hereford, Tom Martin, Joe Travis, and an anonymous referee for helpful comments on an earlier version of the paper. This work was supported in part by NSF grants IBN-0212798 to J. G. K. and DEB-0234714 to D. W. P.

References cited

- Blows MW. 2007. A tale of two matrices: Multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology* 20: 1–8.
- Blows MW, Brooks R. 2003. Measuring nonlinear selection. *American Naturalist* 162: 815–820.
- Blows MW, Brooks R, Kraft PG. 2003. Exploring complex fitness surfaces: Multiple ornamentation and polymorphism in male guppies. *Evolution* 57: 1622–1630.
- Bolnick DI. 2004. Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution* 58: 608–618.
- Bonner JT. 2006. *Why Size Matters*. Princeton (NJ): Princeton University Press.
- Brodie ED, Ridenhour BJ. 2003. Reciprocal selection at the phenotypic interface of coevolution. *Integrative and Comparative Biology* 43: 408–418.
- Brodie ED Jr, Formanowicz DRJ, Brodie ED III. 1991. Predator avoidance and antipredator mechanisms: Distinct pathways to survival. *Ethology Ecology and Evolution* 3: 73–77.
- Buffetaut E, Grigorescu D, Csiki Z. 2002. A new giant pterosaur with a robust skull from the latest Cretaceous of Romania. *Naturwissenschaften* 89: 180–184.
- Cavalli-Sforza LL, Bodmer WF. 1971. *The Genetics of Human Populations*. San Francisco: W. H. Freeman.
- Chiappe LM, Codorniu L, Grellet-Tinner G, Rivarola D. 2004. Argentinian unhatched pterosaur fossil. *Nature* 432: 571.
- Company J, Unwin DM, Ruiz-Omenaca JJ, Pereda-Suberbiola X. 2001. A giant azhdarchid pterosaur from the latest Cretaceous of Valencia, Spain—the largest flying creature ever? *Journal of Vertebrate Paleontology* 21: 41A–42A.
- Darwin C. 1859. *On the Origin of Species by Means of Natural Selection*. London: John Murray.
- Darwin C, Wallace AR. 1858. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the Linnean Society of London* 2: 45–62.
- Day T, Young KA. 2004. Competitive and facilitative evolutionary diversification. *BioScience* 54: 101–109.
- Endler JA. 1986. *Natural Selection in the Wild*. Princeton (NJ): Princeton University Press.
- Estes S, Arnold SJ. 2007. Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence on all timescales. *American Naturalist* 169: 227–244.
- Freeman S, Herron JC. 2004. *Evolutionary Analysis*. 3rd ed. Upper Saddle River (NJ): Pearson Education.
- Gibbs HL, Grant PR. 1987. Oscillating selection in Darwin's finches. *Nature* 327: 511–513.
- Grant PR, Grant BR. 2006. Evolution of character displacement in Darwin's finches. *Science* 313: 224–226.
- Grether GF. 1996. Sexual selection and survival selection on wing coloration and body size in the rubyspot damselfly *Hetaerina americana*. *Evolution* 50: 1939–1948.
- Gross MR. 1996. Alternative reproductive strategies and tactics: Diversity within sexes. *Trends in Ecology and Evolution* 11: 92–98.
- Hazelhurst G, Rayer JMV. 1992. Flight characteristics of Jurassic and Triassic Pterosauria: An appraisal based on wing shape. *Paleobiology* 18: 447–463.
- Hendry AP, Kinnison MT. 1999. The pace of modern life: Measuring rates of contemporary microevolution. *Evolution* 53: 1637–1653.
- Hereford J, Hansen TF, Houle D. 2004. Comparing strengths of directional selection: How strong is strong? *Evolution* 58: 2133–2143.
- Hersch E, Phillips PC. 2004. Power and potential bias in the detection of selection in natural populations. *Evolution* 58: 479–485.
- Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hoang A, Hill CE, Beerli P, Kingsolver JG. 2001. Strength and tempo of directional selection in the wild. *Proceedings of the National Academy of Sciences* 98: 9157–9160.
- Hone DWE, Benton MJ. 2004. The evolution of large size: How does Cope's rule work? *Trends in Ecology and Evolution* 20: 4–6.
- Kettlewell HBD. 1973. *The Evolution of Melanism: The Study of a Recurring Necessity; with Special Reference to Industrial Melanism in the Lepidoptera*. Oxford (United Kingdom): Oxford University Press.
- Kingsolver JG, Pfennig DW. 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution* 58: 1608–1612.
- Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CH, Hoang A, Gibert P, Beerli P. 2001. The strength of phenotypic selection in natural populations. *American Naturalist* 157: 245–261.
- Knapczyk FN, Conner JK. Estimates of the average strength of natural selection are not inflated by sampling error or publication bias. *American Naturalist*. Forthcoming.
- Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33: 402–416.
- Lande R, Arnold SJ. 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210–1226.
- Lawson DA. 1975. Pterosaur from the latest Cretaceous of West Texas: Discovery of the largest flying creature. *Science* 187: 947–948.
- Losos JB, Schoener TW, Langerhans RB, Spiller DA. 2006. Rapid temporal reversal in predator-driven natural selection. *Science* 314: 1111.
- Maisey JG. 1991. *The Santana Fossils: An Illustrated Atlas*. Neptune City (NY): TFH.
- Orton GL. 1954. Dimorphism in larval mouthparts in spadefoot toads of the *Scaphiopus hammondi* group. *Copeia* 1954: 97–100.
- Pfennig DW. 1990. The adaptive significance of an environmentally cued developmental switch in an anuran tadpole. *Oecologia* 85: 101–107.
- . 1992. Polyphenism in spadefoot toads as a locally adjusted evolutionarily stable strategy. *Evolution* 46: 1408–1420.
- Pfennig DW, Murphy PJ. 2002. How fluctuating competition and phenotypic plasticity mediate species divergence. *Evolution* 56: 1217–1228.
- Pfennig DW, Rice AM, Martin RA. 2007. Field and experimental evidence for competition's role in phenotypic divergence. *Evolution* 61: 257–271.
- Phillips PC, Arnold SJ. 1989. Visualizing multivariate selection. *Evolution* 43: 1209–1222.
- Pomeroy LV. 1981. Developmental polymorphism in the tadpoles of the spadefoot toad *Scaphiopus multiplicatus*. PhD dissertation, University of California, Riverside.

- Price TM, Kirkpatrick M, Arnold SJ. 1988. Directional selection and the evolution of breeding date in birds. *Science* 240: 798–799.
- Rausher MD. 1992. The measurement of selection on quantitative traits: Biases due to the environmental covariances between traits and fitness. *Evolution* 46: 616–625.
- Reznick DN, Shaw FH, Rodd FH, Shaw RG. 1997. Evaluation of the rate of evolution in natural populations of guppies. *Science* 275: 1935–1937.
- Sinervo B, Basolo AL. 1996. Testing adaptation using phenotypic manipulations. Pages 149–185 in Rose MR, Lauder GV, eds. *Adaptation*. New York: Academic Press.
- Sinervo B, Calsbeek R. 2006. The developmental, physiological, neural and genetic causes and consequences of frequency-dependent selection in the wild. *Annual Review of Ecology and Systematics* 37: 581–610.
- Sinervo B, Doughty P, Huey RB, Zamudio K. 1992. Allometric engineering: A causal analysis of natural selection on offspring size. *Science* 258: 1927–1930.
- Stinchcombe JR, Rutter MT, Burdick DS, Rausher MD, Mauricio R. 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: Theory and practice. *American Naturalist* 160: 511–523.
- Travis J, Reznick DN. 1998. Experimental approaches to the study of evolution. Pages 437–459 in Resitarris J, Bernardo J, eds. *Issues and Perspectives in Experimental Ecology*. New York: Oxford University Press.
- Unwin DM. 2006. *The Pterosaurs from Deep Time*. New York: Pi Press.
- Wang XA, Kellner WA, Zhou Z, de Almeida Campos D. 2005. Pterosaur diversity and faunal turnover in Cretaceous terrestrial ecosystems in China. *Nature* 237: 875–879.
- Wright S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics* 1: 356–366.

doi:10.1641/B570706

Include this information when citing this material.