The Strength of Phenotypic Selection in Natural Populations

J. G. Kingsolver,[1,*] H. E. Hoekstra,[1] J. M. Hoekstra,[1,†] D. Berrigan,[1,‡] S. N. Vignieri,[1] C. E. Hill,[1,§] A. Hoang,[1] P. Gibert,[1,‖] and P. Beerli[2]

1. Department of Zoology, University of Washington, Seattle, Washington 98195;
2. Department of Genetics, University of Washington, Seattle, Washington 98195

ABSTRACT: How strong is phenotypic selection on quantitative traits in the wild? We reviewed the literature from 1984 through 1997 for studies that estimated the strength of linear and quadratic selection in terms of standardized selection gradients or differentials on natural variation in quantitative traits for field populations. We tabulated 63 published studies of 62 species that reported over 2,500 estimates of linear or quadratic selection. More than 80% of the estimates were for morphological traits; there is very little data for behavioral or physiological traits. Most published selection studies were unreplicated and had sample sizes below 135 individuals, resulting in low statistical power to detect selection of the magnitude typically reported for natural populations. The absolute values of linear selection gradients $|\beta|$ were exponentially distributed with an overall median of 0.16, suggesting that strong directional selection was uncommon. The values of $|\beta|$ for selection on morphological and on life-history/phenological traits were significantly different: on average, selection on morphology was stronger than selection on phenology/life history. Similarly, the values of $|\beta|$ for selection via aspects of survival, fecundity, and mating success were significantly different: on average, selection on mating success was stronger than on survival. Comparisons of estimated linear selection gradients and differentials suggest that indirect components of phenotypic selection were usually modest relative to direct components. The absolute values of quadratic selection gradients $|\gamma|$ were exponentially distributed with an overall median of only 0.10, suggesting that quadratic selection is typically quite weak. The distribution of γ values was symmetric about 0, providing no evidence that stabilizing selection is stronger or more common than disruptive selection in nature.

Keywords: microevolution, natural populations, natural selection, phenotypic selection, sexual selection.

* Corresponding author. Present address: Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280; e-mail: jgking@bio.unc.edu.

† Present address: Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 20892.

‡ Present address: National Cancer Institute, Bethesda, Maryland 20892.

§ Present address: Department of Biology, Coastal Carolina University, Conway, South Carolina 29528.

‖ Present address: Laboratoire Populations, Genetique, et Evolution, Centre National de la Recherche Scientifique, 91198, Gif Sur Yvette Cedex, France.

How strong is phenotypic selection on quantitative traits in nature? Is strong directional selection rare? Is stabilizing or correlational selection common? Is selection more intense on life history than on morphological traits? What is the relative importance of direct versus indirect selection on correlated traits? The answers to these and similar questions are fundamental to understanding the role of selection in determining evolutionary changes and adaptation in current populations. In his influential book, *Natural Selection in the Wild*, Endler (1986) provided a comprehensive review of published field studies of selection through 1983. His synthesis documented the abundant evidence that selection on quantitative traits occurs frequently in natural populations. However, quantitative estimates of the strength of phenotypic selection on quantitative traits were largely lacking at that time (e.g., Endler's review reports values of directional selection differentials from a total of 25 species).

The early 1980s also saw the development and introduction of new methods for estimating the strength of selection on multiple quantitative traits (Lande 1979; Lande and Arnold 1983; Arnold and Wade 1984*a*, 1984*b*). These methods had two important advantages. First, they provided a means of distinguishing the direct and indirect components of selection on a set of correlated traits. This allowed researchers to use simple statistical methods (partial regression models) to estimate the strength of direct selection on a trait in terms of a common metric, the selection gradient, that represents the relationship of relative fitness to the variation in a quantitative trait measured in standard deviation units. Second, selection gradients were shown to be directly relevant to quantitative genetic models for the evolution of multiple, correlated traits (Lande 1979; Lande and Arnold 1983).

Table 1: Summary of the database of phenotypic selection studies (1984–1997)

	Number of items in the database
Studies	63
Records	1,582
Species	62
Genera	51
Taxon type:	
Invertebrates (I)	534 records (19 studies)
Plants (P)	587 records (18 studies)
Vertebrates (V)	461 records (27 studies)
Study type:	
Cross-sectional (C)	14 studies
Longitudinal (L)	51 studies

These and other advances led to an explosion of field studies of selection on quantitative traits in a variety of systems during the ensuing 15 yr. These studies have been important for exploring the role of selection for evolution of particular traits in many specific model systems of interest. However, while aspects of these studies have been reviewed with respect to specific issues or taxa (Primack and Kang 1989; Travis 1989; Gontard-Danek and Møller 1999), no synthesis of this literature has emerged that examines the strength of selection in natural systems more generally.

In this report, we describe such a synthesis. We have reviewed the published literature from 1984 through 1997 that reported standardized selection gradients and/or selection differentials on natural variation in quantitative traits within field populations. Using the resulting database, we explore how the strength of directional and quadratic selection varies across systems, traits, and fitness components and examine the power of these studies for detecting selection on quantitative traits.

Methods

General Considerations

We started our literature review with 1984 because Endler's monograph (1986) reviewed the literature on selection, including estimates of selection differentials, published by the end of 1983. We examined all studies published from 1984 through 1997 in 15 peer-reviewed journals: *American Journal of Botany, American Naturalist, Annals of the Entomological Society of America, Biological Journal of the Linnean Society, Ecological Entomology, Ecology, Environmental Entomology, Evolution, Evolutionary Ecology, Genetics, Heredity, Journal of Evolutionary Biology, Nature, Philosoph-*

ical Transactions of the Royal Society of London, Proceedings of the Royal Society of London, and *Science.* These represent the major peer-reviewed outlets for original scientific research on phenotypic selection. We did not include unpublished studies or studies published in book chapters or technical reports, for which it is more difficult to assess quality control.

We used four main criteria for including studies in our analyses. First, the studies involved quantitative traits showing continuous phenotypic variation within the study population: studies of discrete and categorical traits were not considered. Second, the studies examined natural phenotypic variation within populations: studies involving genetically or phenotypically manipulated traits, highly inbred or experimental genetic lines, or domesticated species were not considered. Third, the studies were conducted in natural conditions in the field: studies in the lab, greenhouse, experimental plots, or other controlled environmental conditions were not considered.

The fourth criterion concerns the metrics used to estimate selection. We included only those studies that reported one or more of the following metrics: standardized linear selection gradients (β), standardized quadratic selection gradients (γ), standardized linear selection differentials (i), or standardized quadratic selection differentials (Lande and Arnold 1983; Arnold and Wade 1984a, 1984b). These metrics estimate selection on a trait in terms of the effects on relative fitness in units of (phenotypic) standard deviations of the trait, allowing direct comparisons among traits, fitness components, and study systems. Selection gradients estimate the selection directly on the trait of interest, whereas selection differentials estimate total selection on the trait acting both directly and indirectly via selection on other, correlated traits (Lande and Arnold 1983). Linear selection gradients and differentials estimate the strength of directional selection; quadratic selection gradients and differentials estimate the curvature of the selection function. Note that stabilizing selection on a trait implies that the quadratic selection gradient and differential are negative, and disruptive selection on a trait implies that the quadratic selection gradient and differential are positive. However, the converse is not true: a negative quadratic selection gradient or differential is not sufficient to demonstrate stabilizing selection (Mitchell-Olds and Shaw 1987).

These criteria exclude many important and interesting studies of selection—we estimate that more than half of all studies of selection in our target journals were excluded by these criteria. However, this enables us to focus our analyses on the question, What are the patterns of phenotypic selection on natural variation in quantitative traits in natural populations in the wild?

Table 2: Summary of the studies of phenotypic selection included in the database

Metric	Median	Mode	Range
Sample size	134	…	10–8,088
Total no. of traits	4	1	1–8
Total no. of fitness measures	1	1	1–8
No. of temporal replicates	1	1	1–10
No. of spatial replicates	1	1	1–12

The Database

Five kinds of data were recorded from each included study into a database (for details of fields and variables, see appendix): first, the study system, for example, species name and taxonomic grouping (see below), geographic location, and type of habitat; second, the study, for example, method of analysis (e.g., cross-sectional) and the number of temporal and spatial replicates; third, the traits, for example, number of traits in the study and trait class (e.g., morphological); fourth, fitness, for example, fitness component (e.g., survival) and number of fitness measures in the study; and, finally, selection estimates, for example, estimates of selection gradients and/or differentials, associated standard errors and *P* values, and sample size.

Each record in the database contains the information about linear and quadratic selection on a single trait for one selection episode within one replicate within a study. (Information about correlational selection, which involves a combination of traits, is coded as a separate record.) As a result, a single study may contribute more than one record to the database (see tables 1, 2).

The process of assembling the database proceeded in four steps. First, an initial reviewer was assigned to part or all of each target journal; a total of 15 reviewers were involved. Each reviewer surveyed all the articles published in his/her assigned journal and identified those studies that met all the criteria for inclusion. The reviewer then coded the studies into the database. Second, an independent editor was assigned to a set of studies in the database; a total of nine editors (the authors of this article) were involved. We ensured that the reviewer and the editor were different individuals for each study. The editor independently coded each assigned study and corrected any errors in the database; ambiguities were discussed by the entire group of editors and resolved. Third, all the editors independently coded two studies in the database to assess variability among editors in codings. We used this information to exclude some fields from the database that were not coded consistently among the editors. Fourth, one author (J. G. Kingsolver) searched the assembled database to identify and to correct remaining errors and empty cells. We excluded one study that reported two extreme outlier estimates for linear selection gradients where β exceeded 5. While some errors and ambiguities in the database prob-

Table 3: Number of estimates of linear selection in the database as a function of taxon, trait type, and fitness component

Taxon		Trait		Fitness component	
Estimates of linear selection gradients[a]					
Invertebrates	333	Morphology	815	Mating success	407
Plants	363	Life history/phenology	128	Survival	288
Vertebrates	297	Principal component	33	Fecundity	271
…	…	Behavior	14	Total fitness	19
…	…	Interaction	NA	Net reproductive rate	3
…	…	Other	3	Other	5
Estimates of linear selection differentials[b]					
Invertebrates	233	Morphology	594	Mating success	267
Plants	183	Life history/phenology	125	Survival	293
Vertebrates	337	Principal component	21	Fecundity	142
…	…	Behavior	10	Total fitness	34
…	…	Interaction	NA	Net reproductive rate	12
…	…	Other	3	Other	5

Note: NA = not applicable.

[a] $N = 993$ total estimates.

[b] $N = 753$ total estimates.

Table 4: Number of estimates of quadratic selection in the database as a function of taxon, trait type, and fitness component

Taxon		Trait		Fitness component	
Estimates of quadratic selection gradients, γ_{ii}/γ_{ij}[a]					
Invertebrates	215/44	Morphology	358	Mating success	139/26
Plants	147/59	Life history/phenology	77	Survival	112/19
Vertebrates	103/6	Principal component	9	Fecundity	199/64
…	…	Behavior	15	Total fitness	0/0
…	…	Interaction	109	Net reproductive rate	0/0
…	…	Other	6	Other	0/0
Estimates of quadratic selection differentials[b]					
Invertebrates	56	Morphology	183	Mating success	52
Plants	69	Life history/phenology	28	Survival	110
Vertebrates	104	Principal component	6	Fecundity	65
…	…	Behavior	3	Total fitness	2
…	…	Interaction	NR	Net reproductive rate	0
…	…	Other	3	Other	0

Note: NR = not recorded.

[a] $N = 465/109$ total estimates.

[b] $N = 229$ total estimates.

ably remain and another set of researchers might code these same studies somewhat differently, this process helped to ensure that our codings are internally consistent.

Analysis

The database includes a highly heterogeneous set of studies and study systems with disparate biological characteristics: the overall "average" strength of selection is unlikely to be very informative. We focus our analyses on the distributions of selection strengths and on how methodological (e.g., type of study, sample size) or biological (e.g., trait type, fitness component, taxon) characteristics may contribute to the heterogeneity of selection strengths. We explore these issues graphically. We did not conduct formal
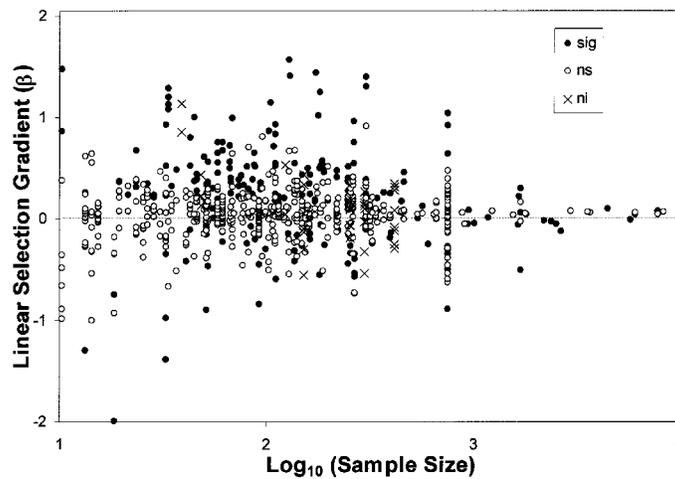

Figure 1: Linear selection gradient estimates (β) as a function of sample size ($\log_{10}$ scale; $N = 993$ estimates). The statistical significance (at the $P = .05$ level) of each estimate is given: filled circles indicate significantly different from 0; open circles indicate not significant; x's indicate significance of the estimate not available.
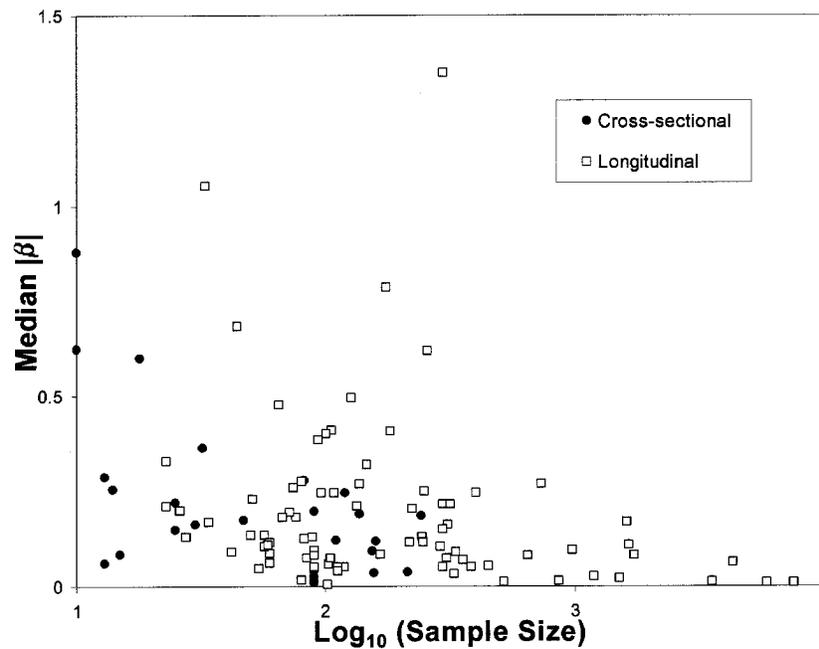
Figure 2: Median value of the linear selection gradient estimates ($|\beta|$) as a function of sample size ($\log_{10}$ scale) for each replicate of each study. Data for both cross-sectional (*filled circles*, $N = 32$) and longitudinal (*open squares*, $N = 85$) are included.

meta-analyses (Hedges and Olkin 1985; Arnqvist and Wooster 1995) of these data for two reasons. First, most of the studies contributed multiple estimates of selection, often involving multiple traits and/or fitness measures estimated from a single experiment; as a result, the data are clearly not independent (Gurevitch and Hedges 1999). Second, because selection gradients represent partial regression coefficients, meta-analysis requires information about the entire phenotypic variance-covariance matrix for each study (L. Hedges, personal communication); this was not available for most studies. Meta-analysis of selection differentials only requires knowledge of variances (or the equivalent), but most studies did not report variances or standard errors for estimates of selection differentials. We recommend that studies of selection report this information so that valid meta-analyses of selection can be performed in the future.

Results

Summary of the Database and Studies

We identified 63 studies of 62 species in 51 genera that met our criteria for inclusion (table 1). The majority of studies (81%) used longitudinal (cohort) methods of analysis for detecting selection. The resulting database had 1,582 total records, in which invertebrates, plants, and vertebrates were

all well represented. The studies were diverse in terms of numbers of traits, number of fitness measures, replication, and sample size (table 2); for example, sample sizes associated with the estimates ranged from 10 to 8,088. However, the majority of studies were not replicated and measured selection on one to four traits; the median sample size across studies was 134. This small sample size results in quite low power to detect selection. For example, consider a trait under directional selection such that $\beta = 0.2$. In a selection study of this trait using a sample size of 134, the probability that one could reject the null hypothesis of no selection ($\beta = 0$) at the 95% level is <50% (see "Discussion").

The database contained 993 estimates of linear selection gradients (β) and 753 estimates of standardized linear differentials (i; table 3). Note that many studies reported both gradients and differentials for the same traits (see fig. 6 below). Invertebrates, plants, and vertebrates were all well represented in these estimates. Over 80% of the estimates were for morphological traits, and 13%–17% were for life-history/phenological traits. Other types of traits were poorly represented. Estimates involving aspects of mating success, fecundity, and survival are all well represented; estimates for more comprehensive aspects of fitness, such as net reproduction rate or intrinsic rate of increase, are rare.

There were 574 estimates of quadratic selection gradients (γ) in the database (table 4); quadratic selection differentials
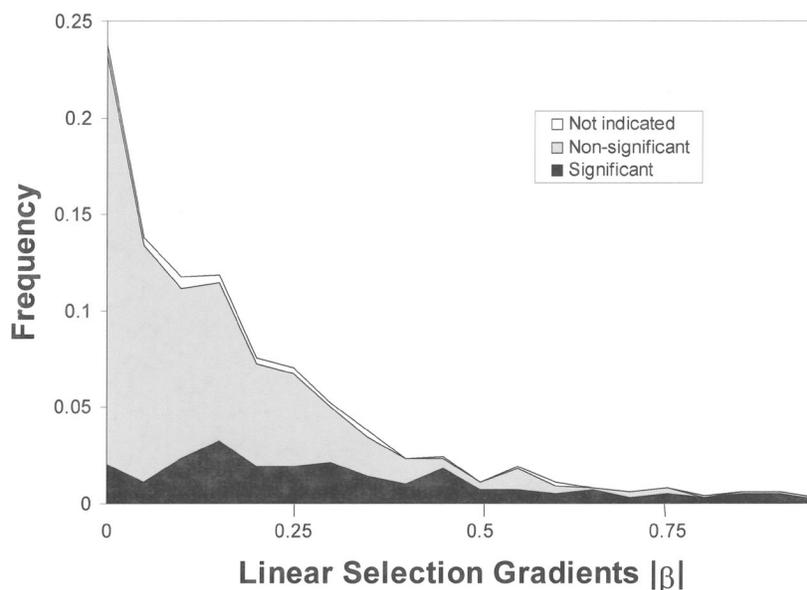
Figure 3: Frequency distribution of the absolute values of the linear selection gradient estimates ($|\beta|$) binned at 0.05 value intervals ($N = 993$). The distributions are stacked according to the statistical significance (at the $P = .05$ level) of each individual estimates: black indicates significantly different from 0; grey indicates not significant; open indicates significance of the estimates not available.

were less commonly reported (table 4). As for linear selection estimates, most estimates of quadratic selection were for morphological traits, with life-history/phenological traits also represented. There were 109 estimates of correlational selection recorded (off-diagonal elements of γ: "Interaction" in table 4); 58% of these involved pairs of morphological traits, and 29% involved combinations of morphological with life-history/phenological traits.

Strength of Linear (Directional) Selection

Most estimates for the linear gradient (β) varied between -1 and 1 (fig. 1). Variation in β decreased with increasing sample size, as expected if sampling variance contributes substantially to the total variance in β (Light and Pillemer 1984; Palmer 1999). For sample sizes >1,000, most estimates cluster between -0.1 and 0.1. Note that, for very small sample sizes (<20), there are few estimates near 0; this may reflect a "file drawer" phenomenon, in which studies with small sample size that do not suggest strong selection are unlikely to be submitted and accepted for publication and are thus not represented in our database (Iyengar and Greenhouse 1988; Palmer 1999). For our purposes, the direction of selection (plus or minus) is not informative, and the distribution of β values is fairly symmetric about 0; thus, it is generally more useful to consider

the absolute value $|\beta|$ as an estimate of the magnitude of selection.

Because most studies report multiple estimates of selection, many of the estimates are not independent. To evaluate this, we plotted the median magnitude of selection $|\beta|$ for each replicate of each study as a function of sample size (fig. 2). Again the median $|\beta|$ is typically between 0.0 and 0.3 at all sample sizes; most values exceeding 0.3 occur at smaller sample sizes, as expected from sampling variance. Note that most studies using cross-sectional methods of analysis had sample sizes smaller than the overall median sample size (134), and all replicates with sample sizes <25 involved cross-sectional methods. This is of interest because cross-sectional methods generally require more stringent assumptions than longitudinal (cohort) methods (Endler 1986).

The frequency distribution of the magnitude of linear selection ($|\beta|$) is approximately exponentially distributed, with a mean value of 0.22 (fig. 3). The exponential distribution implies that the magnitude of $|\beta|$ is typically rather modest (median of 0.16) but with a long tail of larger values. For example, 13% of the estimates of $|\beta|$ were >0.5, and 5% were >0.75. Note that only 25% of the individual estimates of $|\beta|$ are significantly different from 0 at the 95% level; only for $|\beta|$ greater than ~0.3 are the majority of estimates significantly different from 0. This reflects the low
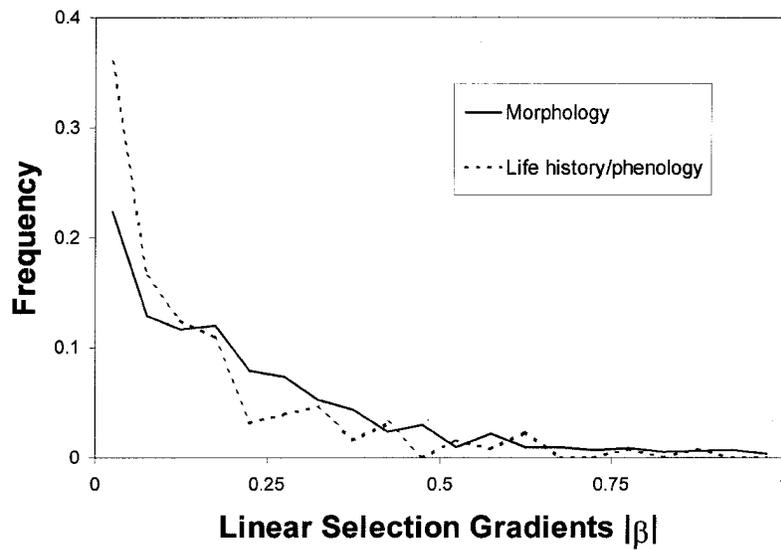
Figure 4: Frequency distribution of the absolute values of the linear selection gradient estimates ($|\beta|$) binned at 0.05 value intervals, for morphological traits (*solid line*, $N = 815$) and life-history/phenological traits (*dashed line*, $N = 128$).

statistical power of most studies to detect "typical" selection (see "Discussion").

For morphological and for life-history/phenological traits (table 3), there are enough estimates to compare the frequency distributions of the magnitude of linear selection ($|\beta|$; fig. 4). The values of $|\beta|$ for morphological and life-history/phenological traits are significantly different (Wilcoxon rank sum test: $Z = 4.25$, $P < .001$): the magnitude of selection on morphological traits (median $|\beta| = 0.17$) is typically greater than on life-history/phenological traits (median $|\beta| = 0.08$; fig. 4). This basic difference between morphological and life-history/phenological traits also holds within taxa (e.g., within vertebrates and within plants). It is interesting that >61% of the estimates of β for life-history/phenological traits were for plants; few studies have reported estimates of selection strengths on life history in animals, particularly for invertebrates.

The magnitudes of linear selection ($|\beta|$) are also significantly different (Kruskal-Wallis rank sum test: $\chi^2 = 49.22$, df $= 2$, $P < .001$) for selection via different components of fitness (fig. 5). In particular, values of $|\beta| < 0.10$ are much more frequent for selection via survival (median $|\beta| = 0.09$) than via fecundity (median $|\beta| = 0.16$) or mating success (median $|\beta| = 0.18$). This basic difference between selection on survival and mating success holds for vertebrates and for plants but is less clear for selection within invertebrates; it also holds when only morphological traits are considered. This pattern suggests that sexual selection is stronger than viability selection.

Our analyses thus far consider standardized selection

gradients (β), which estimate the strength of selection directly on the trait of interest (Lande and Arnold 1983). In contrast, standardized selection differentials (i) estimate the total strength of selection on a trait, resulting both from direct selection and from indirect effects of selection on correlated traits. Because many studies reported estimates of both β and i for the same traits, we can evaluate the relationship between direct and total components of linear selection (fig. 6). Similar values for β and i indicate that direct and total selection are similar, implying that indirect effects of selection on correlated traits are small. Figure 6 shows that most of the estimates for β and i do fall on or near the 1 : 1 line, suggesting that indirect selection is usually small relative to direct selection. In addition, there are only a handful of cases (upper left and lower right quadrants of fig. 6) where both selection gradients and differentials are individually significant and where the gradients and differentials are of opposite sign (i.e., where estimates of direct and total selection are in opposite directions).

Strength of Quadratic Selection

The quadratic selection gradient (γ) estimates the curvature of the relative fitness surface about the mean trait value. Most estimates for the quadratic gradient (γ) varied between -1 and 1 (fig. 7). Variation in γ decreased with increasing sample size for sample sizes above 200–300, as expected if sampling variance contributes substantially to the total variance (Light and Pillemer 1984). As for the
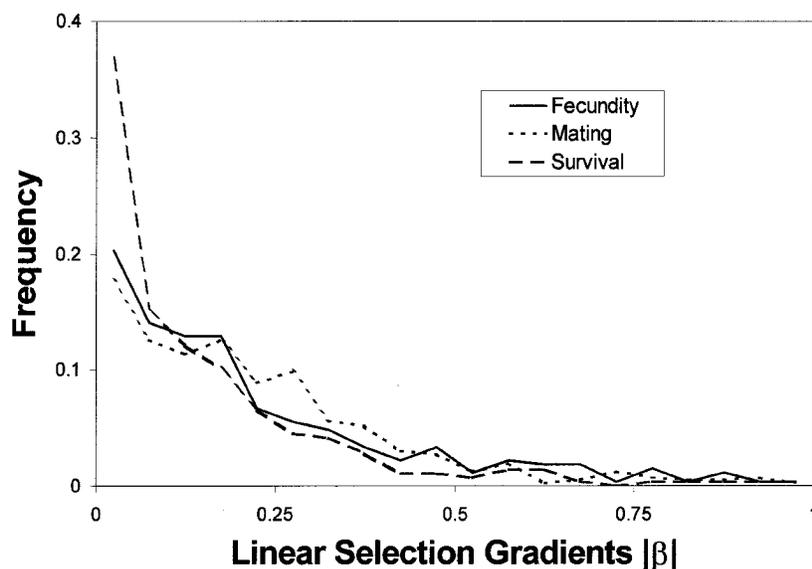
Figure 5: Frequency distribution of the absolute values of the linear selection gradient estimates ($|\beta|$) binned at 0.05 value intervals, for selection via three different components of fitness: fecundity (*solid line*, $N = 271$), mating success (*short dashes*, $N = 407$), and survival (*long dashes*, $N = 288$).

linear gradients, for sample sizes above 1,000, most estimates of γ cluster between -0.1 and 0.1. Note that, for very small sample sizes (<30), there are few small or nonsignificant estimates of γ reported, suggesting a strong file drawer effect in which studies with small sample sizes that do not suggest strong selection are unlikely to be submitted and accepted for publication (Iyengar and Greenhouse 1988; Palmer 1999). Indeed, some studies stated that estimates of γ were initially computed but were not significant and were therefore not reported in the published paper.

The quadratic selection gradients (γ) represent the curvature of fitness with respect to the traits: stabilizing selection implies negative curvature ($\gamma < 0$), whereas disruptive selection implies positive curvature ($\gamma > 0$). The frequency distribution of γ is approximately double exponential (fig. 8), with most values concentrated near 0. The magnitude of quadratic selection is typically small, with a median $|\gamma|$ of 0.10, and 84% of the individual estimates are not significantly different from 0. In addition, the distribution of γ_{ii} is symmetrical about 0. These results suggest that quadratic selection is typically quite weak, and they provide no evidence that stabilizing selection is stronger or more common than disruptive selection. Although the data are much sparser (109 estimates), the frequency distribution of off-diagonal elements of γ, representing correlational selection, is qualitatively similar to that for the diagonal elements.

The magnitudes of quadratic selection strengths also varied significantly among fitness components (fig. 9; Kruskal-Wallis rank sum test: $\chi^2 = 112.04$, df $= 2$, $P < .001$). As for linear selection, the magnitude of quadratic selection via survival is typically weaker (median $|\gamma| = 0.02$) than quadratic selection via fecundity (median $|\gamma| = 0.14$) or mating success (median $|\gamma| = 0.16$). This pattern appears to hold within taxa, though the number of estimates available are rather small (table 4).

Discussion: What Have We Learned about Selection?

Limitations of the Studies, Database, and Analyses

With over 2,500 published estimates of selection gradients and differentials, our information about the strength of phenotypic selection in natural populations has expanded by more than fivefold since the time of Endler's (1986) review. In addition, these estimates are only a subset, perhaps a minority, of the published evidence on phenotypic selection, since we excluded many interesting published studies of selection in the field because they did not satisfy our criteria for inclusion (see "Methods"). The primary reasons for exclusion were the use of alternative methods for estimating or analyzing selection and the use of experimentally generated variation in phenotypes or in environmental conditions. What have we learned from this
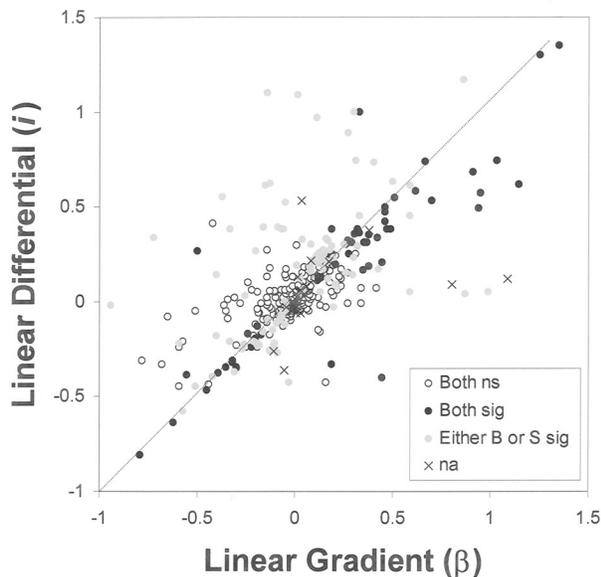
Figure 6: Linear selection differential (*i*, vertical axis) and linear selection gradients (β, horizontal axis) for traits where both are estimated (*N* = 200). The 1 : 1 line is indicated on the plot (*dashed line*). The statistical significance (at the *P* = .05 level) of each pair of estimates is given: black circles indicate that both *i* and β estimates for a trait are significantly different from 0; grey circles indicate one estimate is significantly different from 0; open circles indicate neither estimate is significantly different from 0; *x*'s indicate that the significance of the estimates is not available.

large mass of new data about phenotypic selection on quantitative traits in the wild?

One striking pattern is that >80% of the selection estimates are for morphological traits; phenological traits are also well represented, especially for plants. In contrast, there are very few data for selection on behavioral, physiological, or performance traits. In part, this may reflect the possibility that behavioral phenotypes are often measured as discrete rather than quantitative traits. While the selection estimates represent a diversity of aspects of fitness associated with survival, fecundity, or mating success, only a handful of studies consider more integrated measures of fitness, such as net reproductive rate or intrinsic rate of increase.

It is important to recognize some of the major sources of biases in this database that will influence our interpretations about the strength of selection. Study systems and traits are not chosen "at random" for selection studies: they are often chosen at least in part because investigators suspect the possibility that selection may be operating. By considering only studies published in core peer-reviewed journals, we have tended to exclude more poorly designed studies and "negative" studies that failed to detect selection. Indeed, there is clear evidence of some publication

bias in our data set, in which estimates of weak and non-significant selection in studies with small sample sizes are relatively underrepresented (see figs. 1, 7). Each of these biases will tend to inflate our estimation of the strength of selection for some "random" trait or study system.

Another important feature of the studies taken collectively is their low statistical power (Gurevitch and Hedges 1999). Most of the studies were not replicated in either time or space. In addition, the sample size associated with estimates is quite low, with a median sample size across studies of only *N* = 134 for the estimates. The median sample size among estimates is even lower (*N* = 92), implying that studies with smaller sample sizes tended to report relatively more estimates of selection. As a result, standard errors associated with the estimates are quite large relative to the magnitude of selection, and most studies lacked the statistical power to detect selection of "average" magnitude. Thus, only 25% of the linear gradients and differentials and only 16% of the quadratic gradients and differentials were individually significant at the *P* = .05 level. Moreover, nearly all studies estimated multiple selection gradients or differentials in each experiment, and very few corrected for multiple significance tests (e.g., by adjusting the critical α to maintain an appropriate level of experiment-wide Type I error; Fairbairn and Reeve 2001). For example, a "typical" study (see table 2) that estimated linear and quadratic gradients on four traits from a single longitudinal experiment would generate 20 individual significance tests. As a result, the results reported here almost certainly overestimate the frequency of statistically significant selection and overestimate the power of these studies to detect selection of "average" magnitude (Fairbairn and Reeve 2001). It is sobering that, for sample sizes exceeding ~10^3, most estimates of linear (fig. 1) and quadratic (fig. 7) selection gradients cluster between −0.1 and 0.1: our most powerful studies indicate that selection is weak or absent.

Patterns of Directional Selection

One consistent pattern that emerges from our analyses is that the distribution of selection strengths is approximately exponential (figs. 3–5, 8, 9). The median magnitude of linear selection gradients in our database was |β| = 0.16. There are two interesting consequences of this exponential distribution. First, there is a long "tail" of values indicating that very strong directional selection (|β| > 0.5) may occur. Second, directional selection on most traits and in most systems is quite weak. Both of these patterns are detectable in Endler's (1986) analysis of linear selection differentials based on a much smaller set of estimates (his fig. 7.2*A*), though Endler emphasized the first point more strongly. The theoretical reasons why the magnitude of
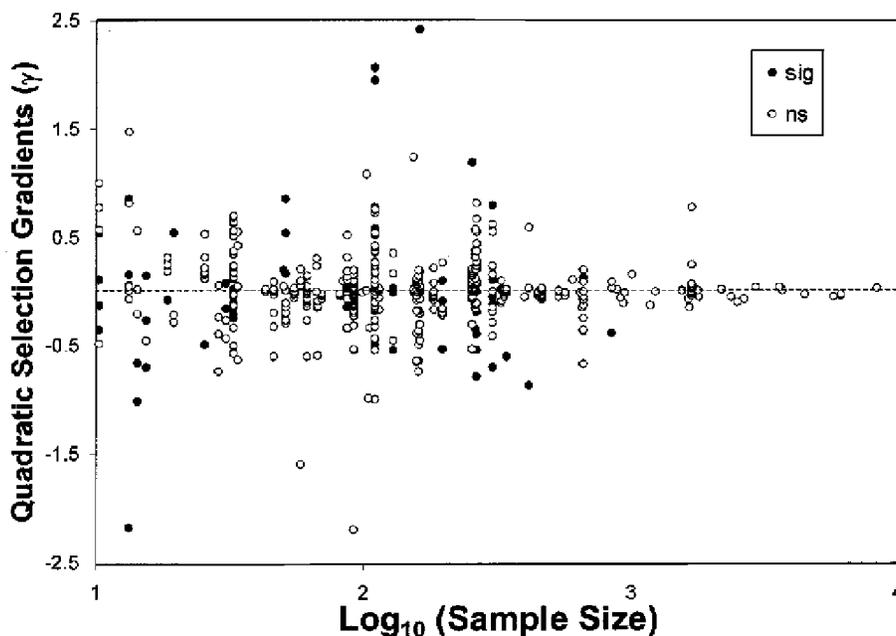
Figure 7: Quadratic selection gradient estimates (γ) as a function of sample size ($\log_{10}$ scale; $N = 574$ estimates). The statistical significance (at the $P = .05$ level) of each estimate is given: filled circles indicate significantly different from 0; open circles indicate not significant.

selection strengths might be distributed exponentially are unclear. For example, if most populations are near adaptive peaks or if variation in selection estimates is due largely to sampling error, one would expect a normal distribution of directional selection gradients centered about 0 rather than a double exponential distribution (J. Felsenstein, M. Kirkpatrick, and M. Lynch, personal communication).

Our analyses also revealed several interesting sources of heterogeneity in the strengths of linear selection. For example, we found the magnitude of linear selection was on average twice as great for morphological traits than for life-history/phenological traits (fig. 4), a result that held qualitatively for both vertebrates and for plants (comparisons for invertebrates were precluded because of a lack of estimates for life-history/phenological traits). Because life history is often closely associated with fitness, one might expect a priori that selection on life history would be stronger than on morphological traits—the reverse of the observed pattern (Mousseau and Roff 1987). However, most of the life-history/phenological traits considered in our studies involved the timing of developmental events (e.g., date of germination, date of first flowering) rather than age-specific mortality or reproduction. One possible statistical explanation for the apparent differences in selection strength between morphological and life-history/phenological traits involves measurement error. In selection gradient (and other standard regression) analyses, measurement error in the

traits is assumed to be negligible: measurement error will reduce the estimated magnitude of the selection gradients. Because measurement error may be substantially greater for life-history and phenological traits than for morphological traits, this could generate an apparent difference in the estimated strength of selection.

Our analyses also indicate that the average strength of selection varies among fitness components: in particular, selection via survival tends to be weaker than selection via fecundity or mating success (fig. 5; H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver, unpublished results). This result appears to hold for different kinds of traits and in different taxa. Endler (1986) found a similar difference between mortality and nonmortality components of selection for polymorphic (discrete) traits but not for quantitative traits based on the limited data then available. There are several possible interpretations of this interesting result (H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver, unpublished results). First, sexual selection (mating success) may in fact tend to be stronger than viability selection (survival). This interpretation is consistent with theoretical models suggesting that sexual selection and mate choice may be important to rapid adaptive divergence (Lande 1981; Kirkpatrick 1982). Second, researchers studying sexually selected traits may be better at identifying
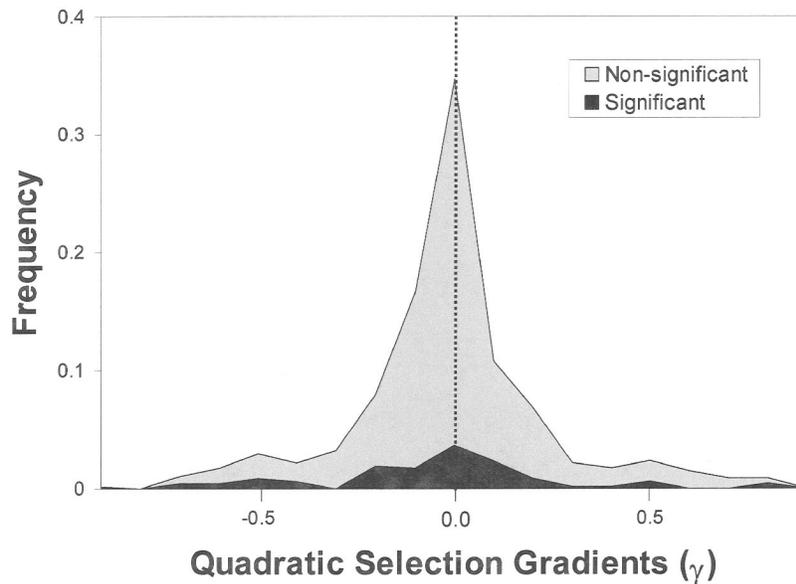
Figure 8: Frequency distribution (in %) of the quadratic selection gradient estimates (γ) binned at 0.10 value intervals ($N = 465$ estimates). The distributions are stacked according to the statistical significance (at the $P = .05$ level) of each individual estimates: black indicates significantly different from 0; grey indicates not significant.

study systems in which strong sexual selection is occurring. It is noteworthy that many traits important in mating success are sexually dimorphic, providing an important clue to researchers interested in selection. Third, because mating success can often be measured over short time periods compared with survival, this difference may reflect the effects of measurement timescale on the estimated strength of selection. H. Hoekstra and colleagues (unpublished results) found that the average magnitude of directional selection was significantly greater for selection episodes measured over shorter intervals (days) than those measured over long intervals (months or years). This result is consistent with analogous studies of effects of measurement scale on estimated rates of microevolution. However, this effect of timescale does not fully account for the differences in the magnitude of sexual and viability selection (H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, A. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver, unpublished results).

Finally, estimates of selection via different fitness components may be biased statistically. For example, most selection studies measure mating success as a discrete count, assuming implicitly that fitness increases linearly with the number of mates. This approach can lead to overestimates of selection gradients (Brodie and Janzen 1996). However, this effect does not explain the apparent differences in mean selection strength between mating success and fecundity,

which is also typically measured as a discrete count. Conversely, most selection studies measure survival as a binomial variable (dead or alive) but use standard regression methods that assume normality for estimating selection gradients; this can produce incorrect estimates of β (Janzen and Stern 1998). However, despite its flaws, this approach does not consistently bias selection estimates toward smaller values (F. Janzen, personal communication).

Our comparisons of linear selection gradients (which estimate only direct selection on a trait) and selection differentials (which include direct effects as well as indirect effects through correlated traits) provide insight into the contributions of indirect selection (fig. 6). Most estimates cluster within ± 0.1 of the 1 : 1 line in figure 6, indicating that direct and total selection on traits are usually similar—that is, that indirect selection is typically small in effect. For example, there are only a handful of cases where both selection gradients and differentials are individually significant and where the gradients and differentials are of opposite sign (i.e., where estimates of direct and total selection are in opposite directions). This does not necessarily imply that correlated traits and indirect selection are unimportant. For example, many researchers may choose traits in ways that reduce the degree of colinearity among traits, and some used principal components (which are by definition independent) as traits in their analyses (note that data for principal components were excluded from
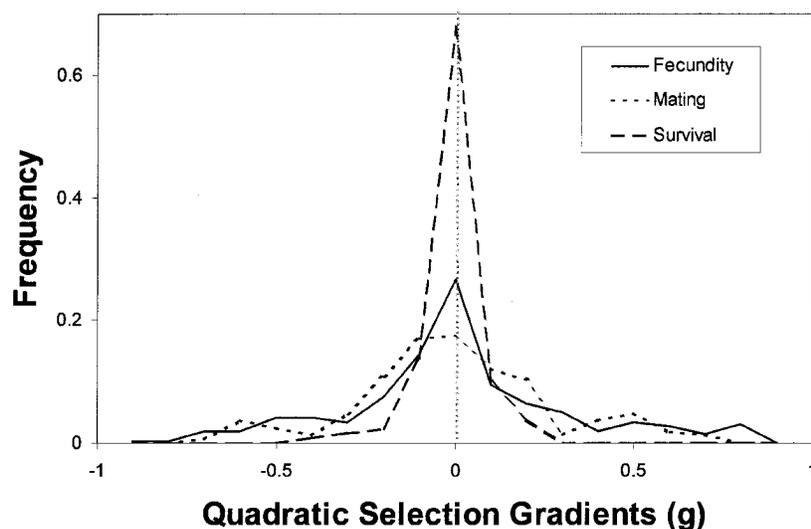
Figure 9: Frequency distribution (in %) of the quadratic selection gradient estimates (γ) binned at 0.10 value intervals, for selection via three different components of fitness: fecundity (*solid line*, $N = 199$), mating success (*short dashes*, $N = 139$), and survival (*long dashes*, $N = 112$).

fig. 6). However, our results do not indicate that indirect effects of correlated traits frequently mask or reverse the direct effects of selection, within the precision of the selection estimates now available.

Patterns of Quadratic Selection

Estimates of quadratic selection provide information about the curvature of the fitness surface: stabilizing selection requires negative curvature ($\gamma < 0$), whereas disruptive selection requires positive curvature ($\gamma > 0$). Estimates of quadratic selection gradients vary widely between -2.5 and 2.5, but most values cluster between -0.1 and 0.1, particularly for studies with larger sample sizes (fig. 7). Only 16% of the individual estimates were significantly different from 0 at the 0.05 level (5% would be expected from Type I errors alone); and the median magnitude of quadratic selection $|\gamma|$ was only 0.10. Most notably, the distribution is symmetric about 0, with negative and positive curvatures similar in frequency and magnitude: this implies that stabilizing selection in no more common or intense than disruptive selection for quantitative traits in the field. This pattern is quite different from that described by Endler (1986), whose analysis of quadratic differentials indicated a strong preponderance of negative values (his fig. 7.4*A*).

There are several interpretations of this striking result. First, quadratic selection may indeed be quite weak, and stabilizing selection may be uncommon. If true, this will require a major rethinking of the roles of stabilizing selection

in adaptation and adaptive landscapes, the maintenance of genetic variation in quantitative traits, and the rationale for optimization approaches to evolution. In addition, future studies will need sample sizes of 500–1,000 or more reliably to detect quadratic selection of this magnitude. Second, most study systems and traits were chosen with some expectation that directional selection might be operating: typically quadratic selection was not the primary motivation for the study. The failure to find substantial quadratic selection in systems and traits where none was anticipated may not be surprising. Third, patterns of environmental variation may alter phenotypic covariances in ways that obscure optimizing selection in nature (Price et al. 1988; Travis et al. 1999). In any case, there remains an urgent need for well-designed field studies of stabilizing selection in appropriate systems, for example, where optimizing selection might be expected (Travis 1989).

Finally, only a handful of studies have studied the strength of correlational selection (the off-diagonal elements of γ). As a result, we still know almost nothing about correlational selection in the field, despite abundant evidence that functional interactions among traits are important in determining organismal performance.

Future Studies of Selection

This review demonstrates that our information about the strength of phenotypic selection in natural populations has increased dramatically in the past 2 decades, but many

important issues about selection remain unresolved. Our analyses suggest some specific directions for future study.

First, higher methodological standards are needed for future studies of selection (Gurevitch and Hedges 1999; Fairbairn and Reeve 2001). Most studies to date are unreplicated and have sample sizes too small to detect selection of typical magnitude. To reduce problems with Type I errors, significance testing of selection gradients and differentials should adjust for multiple tests within studies. Future studies should also report phenotypic variances and covariances among traits and standard errors on selection differentials so that appropriate meta-analyses can be conducted.

Second, we have abundant information about directional selection on morphological traits. By contrast, selection on quantitative behavioral and physiological traits remains largely unknown and should be the focus of future studies.

Third, to address the apparent differences in selection strength among fitness components, studies are needed that estimate the strength of both sexual and viability selection at similar timescales on the same traits. Only a handful of such studies are currently available.

Fourth, much more effort is needed to estimate the strength of quadratic selection in systems and traits where optimizing or correlational selection is anticipated based on functional analyses (Travis 1989). Because the magnitude of quadratic selection may be relatively small, larger sample sizes than in past studies may be required.

Finally, our analyses and comparisons were only possible because selection data have been reported in a common metric—that is, in standardized selection differentials and gradients. However, in many cases estimation and visualization of fitness surfaces may be more informative for understanding selection in particular systems (Schluter and Nychka 1994; Brodie et al. 1995). We suggest that future studies report both standardized metrics (differentials or gradients) and fitness surfaces, so that both specific patterns of selection in particular systems and general patterns of selection across systems can be evaluated.

APPENDIX

Table A1: Fields included for each record included in the selection database for this study

Field	Field name	Description	Values
1	Species	Genus and species	…
2	Taxon.group	Taxonomic group	I = invertebrate, P = plant, V = vertebrate
3	Studytype	Type of study	C = cross-sectional, L = longitudinal
4	#tmp_reps	No. of temporal replicates	…
5	#spc_reps	No. of spatial replicates	…
6	Dur_rep	Duration of the replicate	In days
7	RepID	ID no. of the replicate	…
8	Traitclass	Trait type	MO = morphology, LH = life history or phenology, BE = behavior, PC = principal component, I = interaction
9	Traitname	Name of trait	…
10	Ttl#trait	Total no. of traits	…
11	Fit.comp	Fitness component	F = fecundity/fertility, M = mating success, S = survival, N = net reproductive rate, T = aggregate
12	Ttl#meas	Total no. of fitness measures	…
13	Measure	Name of fitness measure	…
14	Grad.linear.value	Linear gradient, β	…
15	Grad.linear.StErr	SE of β	…
16	Grad.linear.p-value	P value of β	…
17	Grad.quad.value	Quadratic gradient, γ	…

Table A1 (*Continued*)

Field	Field name	Description	Values
18	Grad.quad.StErr	SE of γ	…
19	Grad.quad.p-value	*P* value of γ	…
20	Diff.linear.value	Linear differential, *i*	…
21	Diff.linear.StErr	SE of *i*	…
22	Diff.linear.p-value	*P* value of *i*	…
23	Diff.quad.value	Quadratic differential, *j*	…
24	Diff.quad.StErr	SE of *j*	…
25	Diff.quad.p-value	*P* value of *j*	…
26	*N*	Sample size	…
27	Subset_popn	Subset of population	E.g., M = males, F = females, J = juveniles, A = adults
28	Authors	Authors of study	…
29	Year	Year of publication	…
30	Journal	Journal name	…
31	Vol:p–p	Volume and page no.	…
32	StudyID	Study ID	…

Note: See "Methods." Values common to all fields: NA = not available or not applicable; O = other; NS = not significant.

Studies Included in the Selection Database

Alatalo, R. V., and A. Lundberg. 1986. Heritability and selection on tarsus length in the pied flycatcher (*Ficedula hypoleuca*). Evolution 40:574–583.

Alatalo, R. V., L. Gustafsson, and A. Lundberg. 1990. Phenotypic selection on heritable size traits—environmental variance and genetic response. American Naturalist 135:464–471.

Anholt, B. R. 1991. Measuring selection on a population of damselflies with a manipulated phenotype. Evolution 45:1091–1106.

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection: applications. Evolution 38:720–734.

Arnqvist, G. 1992. Spatial variation in selective regimes: sexual selection in the water strider, *Gerris odontogaster*. Evolution 46:914–929.

Bjorklund, M., and M. Linden. 1993. Sexual size dimorphism in the great tit (*Parus major*) in relation to history and current selection. Journal of Evolutionary Biology 6:397–415.

Brodie, E. D. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake, *Thamnophis ordinoides*. Evolution 46:1284–1298.

Campbell, D. R. 1991. Effects of floral traits on sequential components of fitness in *Ipomopsis aggregata*. American Naturalist 137:713–737.

Campbell, D. R., N. M. Waser, M. V. Price, E. A. Lynch, and R. J. Mitchell. 1991. Components of phenotypic selection: pollen export and lower corolla width in *Ipomopsis aggregata*. Evolution 45:1458–1467.

Campbell, D. R., N. M. Waser, and E. J. Melendez-Ackerman. 1997. Analyzing pollinator-mediated selection in a plant hybrid zone: hummingbird visitation patterns on three spatial scales. American Naturalist 149:295–315.

Conner, J. 1988. Field measurements of natural and sexual selection in the fungus beetle, *Bolitotherus cornutus*. Evolution 42:736–749.

Conner, J. K., and S. Rush. 1997. Measurements of selection of floral traits in black mustard, *Brassica nigra*. Journal of Evolutionary Biology 10:327–335.

Conner, J. K., S. Rush, and P. Jennetten. 1996. Measurement of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). I. Selection through lifetime female fitness. Evolution 50:1127–1136.

Conner, J. K., S. Rush, S. Kercher, and P. Jennetten. 1996. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). II. Selection through lifetime male and total fitness. Evolution 50:1137–1146.

Downhower, J. F., L. S. Blumer, and L. Brown. 1987. Seasonal variation in sexual selection in the mottled sculpin. Evolution 41:1386–1394.

Fairbairn, D. J., and R. F. Preziosi. 1996. Sexual selection and the evolution of sexual size dimorphism in the water strider, *Aquarius remigis*. Evolution 50:1549–1559.

Gibbs, H. L. 1988. Heritability and selection on clutch size in Darwin's medium ground finches (*Geospiza fortis*). Evolution 42:750–762.

Grant, B. R. 1985. Selection on bill characteristics in a population of Darwin finches, *Geospiza conirostris*, on Isla Genovesa, Galapagos. Evolution 39:523–532.

Grant, B. R., and P. R. Grant. 1989. Natural selection in

a population of Darwin finches. American Naturalist 133:377–393.

———. 1993. Evolution of Darwin finches caused by a rare climatic event. Proceedings of the Royal Society of London B, Biological Sciences 251:111–117.

Grether, G. F. 1996. Intersexual competition alone favors a sexually dimorphic ornament in the rubyspot damselfly *Hetaerina americana*. Evolution 50:1949–1957.

———. 1996. Sexual selection and survival selection on wing coloration and body size in the rubyspot damselfly *Hetaerina americana*. Evolution 50:1939–1948.

Harvey, A. W. 1990. Sexual differences in contemporary selection acting on size in the hermit crab, *Clibanarius digueti*. American Naturalist 136:292–304.

Hews, D. K. 1990. Examining hypotheses generated by field measures of sexual selection on male lizards, *Uta palmeri*. Evolution 44:1956–1966.

Hoglund, J., and L. Saterburg. 1989. Sexual selection in common toads—correlates with age and body size. Journal of Evolutionary Biology 2:367–372.

Johnston, M. O. 1991. Natural selection on floral traits in two species of *Lobelia* with different pollinators. Evolution 45:1468–1479.

Jones, R., and D. C. Culver. 1989. Evidence for selection on sensory structures in a cave population of *Gammarus minus* (Amphipoda). Evolution 43:689–693.

Kaar, P., J. Jokela, T. Helle, and I. Kojola. 1996. Direct and correlative phenotypic selection on life-history traits in three pre-industrial human populations. Proceedings of the Royal Society of London B, Biological Sciences 263:1475–1480.

Kalisz, S. 1986. Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). Evolution 40:479–491.

Kelly, C. A. 1992. Spatial and temporal variation in selection on correlated life history traits and plant size in *Chamae crista fasciculata*. Evolution 46:1658–1673.

King, R. B. 1993. Color pattern variation in Lake Erie water snakes: prediction and measurements of natural selection. Evolution 47:1818–1833.

Kingsolver, J. G. 1995. Viability selection on seasonally polyphenic traits: wing melanin pattern in western white butterflies. Evolution 49:932–941.

Koenig, W. D., and S. S. Albano. 1987. Lifetime reproductive success, selection, and the opportunity for selection in the white-tailed skimmer, *Plathemis lydia* (Odonata, Libellulidae). Evolution 41:22–36.

Linden, M., L. Gustafsson, and T. Part. 1992. Selection on fledging mass in the collared flycatcher and the great tit. Ecology 73:336–343.

Martin, T. E. 1998. Are microhabitat preferences of coexisting species under selection and adaptive? Ecology 79:656–670.

Mitchell-Olds, T., and J. Bergelson. 1990. Statistical genetics of an annual plant, *Impatiens capensis*. II. Natural selection. Genetics 124:417–421.

Møller, A. P. 1991. Viability is positively related to degree of ornamentation in male swallows. Proceedings of the Royal Society of London B, Biological Sciences 243:145–148.

Moore, A. J. 1990. The evolution of sexual dimorphism by sexual selection: the separate effects of intrasexual selection and intersexual selection. Evolution 44:315–331.

Morgan, M. T., and D. J. Schoen. 1997. Selection on reproductive characters: floral morphology in *Asclepias syriaca*. Heredity 79:433–441.

Nunez-Farfan, J., and R. Dirzo. 1994. Evolutionary ecology of *Datura stramonium* L. in central Mexico: natural selection for resistance to herbivorous insects. Evolution 48:423–436.

O'Connell, L. M., and M. O. Johnston. 1998. Male and female pollination success in a deceptive orchid, a selection study. Ecology 79:1246–1260.

O'Neil, P. 1997. Natural selection on genetically correlated phenological characters in *Lythrum salicaria* L. (Lythraceae). Evolution 51:267–274.

Preziosi, R. F., and D. J. Fairbairn. 1996. Sexual size dimorphism and selection in the wild in the waterstrider *Aquarius remigis*: body size, components of body size and male mating success. Journal of Evolutionary Biology 9:317–336.

———. 1997. Sexual size dimorphism and selection in the wild in the water strider, *Aquarius remigis*: lifetime fecundity selection on female total length and its components. Evolution 51:467–474.

Price, D. K. 1996. Sexual selection, selection load and quantitative genetics of zebra finch bill colour. Proceedings of the Royal Society of London B, Biological Sciences 263:217–221.

Price, T. 1984. Sexual selection on body size, territory, and plumage variables in a population of Darwin's finches. Evolution 38:327–341.

Price, T. D. 1984. The evolution of sexual dimorphism in Darwin's finches. American Naturalist 123:500–518.

Sabat, A. M., and J. D. Ackerman. 1996. Fruit set in a deceptive orchid: the effect of flowering phenology, display size, and local floral abundance. American Journal of Botany 83:1181–1186.

Santos, M., A. Ruiz, J. E. Quezadadiaz, A. Barbadilla, and A. Fontdevila. 1992. The evolutionary history of *Drosophila buzzatii*. 20. Positive phenotypic covariance between field and adult fitness components and body size. Journal of Evolutionary Biology 5:403–422.

Schemske, D. W., and C. C. Horvitz. 1989. Temporal variation in selection on a floral character. Evolution 43:461–465.

Schluter, D., and J. N. M. Smith. 1986. Natural selection on beak and body size in the song sparrow. Evolution 40:221–231.

Simms, E. L. 1990. Examining selection on the multivariate phenotype: plant resistance to herbivores. Evolution 44:1177–1188.

Smith, T. B. 1990. Natural selection on bill characteristics in the two bill morphs of the African finch, *Pyrenestes ostrinus.* Evolution 44:832–842.

Thessing, A., and J. Ekman. 1994. Selection on the genetical and environmental components of tarsal growth in juvenile willow tits (*Parus montanus*). Journal of Evolutionary Biology 7:713–726.

Van den Berghe, E. P., and M. R. Gross. 1989. Natural selection resulting from female breeding competition in a Pacific salmon (Coho: Oncorhynchuskisutch). Evolution 43:125–140.

Ward, P. I. 1988. Sexual selection, natural selection, and body size in *Gammarus pulex* (Amphipoda). American Naturalist 131:348–359.

Weatherhead, P. J., and R. G. Clark. 1994. Natural selection and sexual size dimorphism in red-winged blackbirds. Evolution 48:1071–1079.

Weatherhead, P. J., H. Greenwood, and R. G. Clark. 1987. Natural selection and sexual selection on body size in redwinged blackbirds. Evolution 41:1401–1403.

Willis, J. H. 1996. Measures of phenotypic selection are biased by partial inbreeding. Evolution 50:1501–1511.

Wilson, P. 1995. Selection for pollination success and the mechanical fit of *Impatiens* flowers around bumblebee bodies. Biological Journal of the Linnean Society 55:355–383.

Zeh, D. W., J. Zeh, A. Bermingham, and E. Bermingham. 1997. The evolution of polyandry. 2. Post-copulatory defences against genetic incompatibility. Proceedings of the Royal Society of London B, Biological Sciences 264:119–125.

Zuk, M. 1988. Parasite load, body size, and age of wild caught male field crickets (Orthoptera: Gryllidae): effects on sexual selection. Evolution 42:969–976.

Literature Cited

Arnold, S. J., and M. J. Wade. 1984*a*. On the measurement of natural and sexual selection: applications. Evolution 38:720–734.

———. 1984*b*. On the measurement of natural and sexual selection: theory. Evolution 38:709–719.

Arnqvist, G., and D. Wooster. 1995. Meta-analysis—synthesizing research findings in ecology and evolution. Trends in Ecology & Evolution 10:236–240.

Brodie, E. D., and F. J. Janzen. 1996. On the assignment of fitness values in statistical analyses of selection. Evolution 50:437–442.

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

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

Fairbairn, D. J., and J. P. Reeve. 2001. Natural selection. *In* C. Fox, D. A. Roff, and D. J. Fairbairn, eds. Evolutionary ecology: concepts and case studies. Oxford University Press, London (in press).

Gontard-Danek, M. C., and A. P. Møller. 1999. The strength of sexual selection: a meta-analysis of bird studies. Behavioral Ecology 10:476–486.

Gurevitch, J., and L. V. Hedges. 1999. Statistical issues in ecological meta-analyses. Ecology 80:1142–1149.

Hedges, V. L., and I. Olkin. 1985. Statistical methods for meta-analysis. Academic Press, Orlando, Fla.

Iyengar, S., and J. B. Greenhouse. 1988. Selection models and the file drawer problem. Statistical Science 3:109–135.

Janzen, F. J., and H. S. Stern. 1998. Logistic regression for empirical studies of multivariate selection. Evolution 52:1564–1571.

Kirkpatrick, M. 1982. Sexual selection and the evolution of female choice. Evolution 36:1–12.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain : body size allometry. Evolution 33:402–416.

———. 1981. Models of speciation by sexual selection on polygenic traits. Proceedings of the National Academy of Sciences of the USA 78:3721–3725.

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

Light, R. J., and D. B. Pillemer. 1984. Summing up: the science of reviewing research. Harvard University Press, Cambridge, Mass.

Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. Evolution 41:1149–1151.

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59:181–197.

Palmer, A. R. 1999. Detecting publication bias in meta-analyses: a case study of fluctuating asymmetry and sexual selection. American Naturalist 154:220–233.

Price, T. M., M. Kirkpatrick, and S. J. Arnold. 1988. Directional selection and the evolution of breeding date in birds. Science (Washington, D.C.) 240:798–799.

Primack, R. B., and H. Kang. 1989. Measuring fitness and natural selection in wild plant populations. Annual Review of Ecology and Systematics 20:367–396.

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

Travis, J. 1989. The role of optimizing selection in natural populations. Annual Review of Ecology and Systematics 20:279–296.

Travis, J., M. G. McManus, and C. F. Baer. 1999. Sources of variation in physiological phenotypes and their evolutionary significance. American Zoologist 39:422–433.

Editor: Joseph Travis